

J. Dvořák and E.B.V. Jensen

On semi-automatic estimation of surface area

On semi-automatic estimation of surface area

J. Dvořák¹ and E.B.V. Jensen²

¹Department of Probability and Mathematical Statistics, Charles University in Prague

²Department of Mathematics, Centre for Stochastic Geometry and Advanced Bioimaging,
Aarhus University

Summary

In the present paper, we propose a semi-automatic procedure for estimation of particle surface area. It uses automatic segmentation of the boundaries of the particle sections and applies different estimators depending on whether the segmentation was judged by a supervising expert to be satisfactory. If the segmentation is correct the estimate is computed automatically, otherwise the expert performs the necessary measurements manually.

In case of convex particles we suggest to base the semi-automatic estimation on the so-called flower estimator, a new local stereological estimator of particle surface area. For convex particles, the estimator is equal to four times the area of the support set (flower set) of the particle transect. We study the statistical properties of the flower estimator and compare its performance to that of two discretizations of the flower estimator, namely the pivotal estimator and the surfactor.

For ellipsoidal particles, it is shown that the flower estimator is equal to the pivotal estimator based on support function measurements along four perpendicular rays. This result makes the pivotal estimator a powerful approximation to the flower estimator. In a simulation study of prolate and oblate ellipsoidal particles, the surfactor also performs well for particles which are not extremely elongated. In particular, the surfactor is not very much affected by the singularity in the surfactor formula or by possible inaccuracies in the necessary angle measurements.

We also assess the performance of the semi-automatic procedure in a study of somatostatin positive inhibitory interneurons from mice hippocampi. Only 35% of the cells needed to be analysed manually and an important decrease in workload was obtained by using the semi-automatic approach.

Keywords: Local stereology, surface area, flower estimator, pivotal estimator, surfactor, computerized image analysis.

Corresponding author: Jiří Dvořák, dvorak@karlin.mff.cuni.cz

1 Introduction

In Cruz-Orive (2005), a new local stereological principle, the invariator, for generating test lines with motion invariant density in three-dimensional space was introduced (for an introduction to local stereology, see Jensen (1998)). Using the invariator principle, a new unbiased estimator of the particle surface area was derived. In the case of a convex particle, this estimator takes the simple form of 4 times the area of the support set (flower set) of the particle transect, generated by an isotropic random section through a fixed reference point of the particle. For this reason, the estimator s_{flo} will be called the flower estimator. A discretized version of the estimator, called the pivotal estimator s_{piv} , was also provided in Cruz-Orive (2005). The pivotal estimator involves determination of the values of the support function of the particle transect in the section plane.

In a subsequent series of papers by Cruz-Orive, further aspects of these new estimators were studied. In Cruz-Orive (2008), the variance of the pivotal estimator was compared with the variance of another unbiased estimator of surface area, the so-called surfactor s_{sur} that had been suggested much earlier in Jensen & Gundersen (1987) and further discussed in Jensen & Gundersen (1989). For a review of other early methods of surface area estimation, see Kubínová & Janáček (1998).

In the comparison performed by Cruz-Orive (2008), a discretization along a single ray in the section plane, emanating from the reference point, was used for both estimators. It was concluded that for spherical particles with arbitrarily positioned reference point in their interior the pivotal estimator has smaller variance than the surfactor. In Cruz-Orive (2011), exact formulae for the area of the support set were derived for a wide range of shapes of the particle transect, thus enabling the practical use of the flower estimator.

Very recently, it was shown in Cruz-Orive (2012) that for a convex particle the mean surfactor estimator (also called integrated or average surfactor) in the plane section is in fact identical to the flower estimator, under rather non-restrictive assumptions about the boundary of the particle. As a result, two approximations of the flower estimator are now available – the pivotal estimator and the surfactor. In fact these are two discretizations of the same integral, see Section 2.

We discuss the above-mentioned estimators in detail in Section 2. We present a number of new results concerning their variances and other statistical properties. Moreover, we assess and compare their performance in a broader setting than the one of Cruz-Orive (2008) – we consider ellipsoids of revolution for which we derive exact formulae for the variance of s_{flo} , under the assumption that the reference point is positioned in the centre of the ellipsoid. The most surprising result concerns the decrease in the relative variance of s_{flo} with increasing elongation of the ellipsoid, in the case of oblate ellipsoids. Another surprising result of the comparison indicates that for ellipsoidal particles choosing a centrally positioned reference point may not be optimal in the sense of minimizing the variance of s_{flo} . Other positions of the reference point may lead to lower variability of the estimates.

In the present paper we also propose a semi-automatic version of the flower estimator that uses automatic segmentation of the boundaries of the particle sections. An expert supervises the process. If the segmentation is judged to be satisfactory,

an estimate of the surface area of the particle is calculated, using the area of the support set of the particle section. In the remaining cases, the expert intervenes and uses a discretized version of the flower estimator, based on a small number of manual measurements in the section plane. The resulting estimator is called the semi-automatic flower estimator s_{semi} . A similar semi-automatic approach has recently been proposed for volume estimation, see Hansen, Nyengaard, Andersen & Jensen (2011). We study the precision of the semi-automatic flower estimator obtained by applying a pivotal estimator respectively a surfactor on particle sections with unsatisfactory segmentation. The resulting two estimators are called $s_{\text{semi}}^{\text{piv}}$ and $s_{\text{semi}}^{\text{sur}}$, respectively.

The application of the semi-automatic flower estimator is illustrated in a study of somatostatin positive inhibitory interneurons from mice hippocampi which have been genetically labelled with green fluorescent protein (GFP) and observed by optical fluorescent microscopy. These data have recently also been studied in Hansen, Nyengaard, Andersen & Jensen (2011) for semi-automatic volume estimation.

The composition of the paper is as follows. In Section 2, we give the necessary background and notation for the flower estimator, the pivotal estimator and the surfactor. We also give new results concerning the variance of these estimators in this section. The performance of the estimators is compared for ellipsoidal particles in Section 3. The semi-automatic approach is described in Section 4 and its performance in the data set of interneurons is studied in Section 5. Conclusions of our investigations are presented in Section 6. Mathematical derivations are deferred to two appendices.

2 The estimators

Let Y be a particle (compact subset of \mathbb{R}^3). The aim is to estimate the surface area of Y . Throughout this paper we assume that Y is convex and has a piecewise smooth boundary so that the set of points on the boundary of Y where a tangent plane is not uniquely determined has two-dimensional area (Hausdorff) measure zero.

Let $S(Y)$ be the surface area of Y and O an arbitrary reference point in its interior. We assume that O is the origin of the coordinate system. We consider a random isotropic section plane L_2 through O and denote by $\partial(Y \cap L_2)$ the boundary of the particle section. Let $h(\omega)$ denotes the support function of $Y \cap L_2$ with respect to O , i.e.

$$h(\omega) = \max(\langle x, u_\omega \rangle; x \in Y \cap L_2), \omega \in [0, 2\pi), \quad (2.1)$$

where u_ω is a unit vector making an angle ω with a fixed axis in L_2 , x is a vector joining O with a point of $Y \cap L_2$ and $\langle x, u_\omega \rangle$ denotes the inner product of the two vectors, see Cruz-Orive (2005).

Finally, let H_{L_2} denote the support set of the (convex) planar section $Y \cap L_2$, i.e. the set which has the graph of the support function $h(\omega)$ as its boundary. The support set of a convex planar set is sometimes called the flower set. Its area will be denoted by $A(H_{L_2})$.

In the following subsections we present estimators of the surface area $S(Y)$ of a fixed particle Y . Note that if we extend this basic setting and consider Y to

be a random particle these estimators can be used to estimate the population mean surface area $\mu = \mathbb{E}S(Y)$. If a sample Y_1, \dots, Y_n from a population of random particles is available and s is one of the estimators in question, $\hat{s} = \frac{1}{n} \sum_{i=1}^n s(Y_i)$ can be employed to estimate μ . The estimators considered in this paper can also easily be adjusted to different sampling schemes and used for size-weighted surface area estimation as discussed in Karlsson & Cruz-Orive (1997).

2.1 The flower estimator

The invariator principle proposed by Cruz-Orive (2005) can be utilized in the estimation of the surface area of the convex particle Y , see also Cruz-Orive (2011). The surface area of Y is

$$S(Y) = 4 \mathbb{E}\{A(H_{L_2})\} = 2 \mathbb{E}\left\{\int_0^{2\pi} h^2(\omega) d\omega\right\}, \quad (2.2)$$

where the expectation is taken with respect to the isotropic random orientation of L_2 .

It follows that four times the area of the support set (the flower set) of the observed section $Y \cap L_2$ is an unbiased estimator of $S(Y)$, called the flower estimator:

$$s_{\text{flo}} = 4A(H_{L_2}) = 2 \int_0^{2\pi} h^2(\omega) d\omega. \quad (2.3)$$

If Y is an ellipsoidal particle with a reference point in its interior an especially simple formula for s_{flo} can be obtained. In that case any planar section of Y through the reference point is an ellipse. The formula for the area of its support set is given in Cruz-Orive (2011). It follows that

$$s_{\text{flo}} = 2\pi(a_{L_2}^2 + b_{L_2}^2 + r^2), \quad (2.4)$$

where a_{L_2} and b_{L_2} denote the semiaxes lengths of the section ellipse and r is the distance of the reference point O from the centre of the section ellipse. We use this result in Section 3 to compare the empirical performance of the estimators in case of ellipsoidal particles.

Concerning the variance of the flower estimator, we have derived exact formulae for the case of oblate and prolate ellipsoids with the reference point in their centre. For an oblate ellipsoid with semiaxes lengths equal to a, a and 1 , $a > 1$, the variance of s_{flo} equals

$$\text{var}(s_{\text{flo}}) = 2\pi^2 a^2 + \frac{2\pi^2 a \operatorname{arcsinh}(\sqrt{a^2 - 1})}{\sqrt{a^2 - 1}} - \frac{4\pi^2 a^2 \operatorname{arcsinh}^2(\sqrt{a^2 - 1})}{a^2 - 1}. \quad (2.5)$$

Similarly, for a prolate ellipsoid with semiaxes lengths equal to $1, 1$ and c , $c > 1$, the variance of s_{flo} equals

$$\text{var}(s_{\text{flo}}) = \frac{2\pi^2 c^2}{c^2 - 1} \left(c^2 - 1 + c^2 \sqrt{c^2 - 1} \arctan(\sqrt{c^2 - 1}) - 2c^2 \arctan(\sqrt{c^2 - 1})^2 \right). \quad (2.6)$$

Full details including derivation of these formulae are given in Appendix A.

A surprising consequence of the formula (2.5) is the fact that for an oblate ellipsoid the relative variance of s_{flo} (i.e. the variance divided by the square of the true surface area) converges to 0 as the elongation of the ellipsoid increases ($a \rightarrow \infty$). This trend is apparent in the numerical results in Section 3 and is not to be considered a mistake. Again, necessary details are given in Appendix A, see especially Figure 9.

2.2 The pivotal estimator

The pivotal estimator s_{piv} was proposed in Cruz-Orive (2005) as an approximation to the flower estimator in case of convex particles. In the most simple form it can be expressed as

$$s_{\text{piv},1} = 4\pi h^2(\theta), \quad (2.7)$$

where the angle θ is random with uniform distribution in $[0, 2\pi)$. This secures the unbiasedness of the estimator. In order to reduce the variance of the estimator (and make the above-mentioned approximation more accurate) one could consider more directions in which the values of the support function are measured. This leads to another version of the pivotal estimator:

$$s_{\text{piv},N} = \frac{4\pi}{N} \sum_{n=0}^{N-1} h^2\left(\theta + n\frac{2\pi}{N}\right), \quad (2.8)$$

where $N \in \mathbb{N}$ and the angle θ is random and uniformly distributed in $[0, 2\pi)$. Note that the distribution of $s_{\text{piv},N}$ remains the same if θ is uniformly distributed in $[0, 2\pi/N)$ instead. Also, $s_{\text{piv},N}$ is an unbiased estimator of $S(Y)$ for any $N \in \mathbb{N}$.

To see that this procedure in fact results in variance reduction in most cases consider $s_{\text{piv},1} = 4\pi h^2(\theta)$ and $s_{\text{piv},2} = 2\pi(h^2(\theta) + h^2(\theta + \pi))$, where θ has a uniform distribution in $[0, 2\pi)$. Since $h(\theta) = h(\theta + 2\pi)$, $h^2(\theta)$ and $h^2(\theta + \pi)$ have the same distribution. As a consequence $s_{\text{piv},2}$ has the form $\frac{1}{2}(X_1 + X_2)$, where X_1 and X_2 have the same distribution as $s_{\text{piv},1}$. Hence, by Cauchy-Schwarz inequality,

$$\begin{aligned} \text{var}(s_{\text{piv},2}) &= \frac{1}{4}(\text{var } X_1 + \text{var } X_2 + 2\text{cov}(X_1, X_2)) \\ &= \frac{1}{2}\text{var}(s_{\text{piv},1}) + \frac{1}{2}\text{cov}(X_1, X_2) \leq \text{var}(s_{\text{piv},1}). \end{aligned} \quad (2.9)$$

Therefore, the variance of $s_{\text{piv},2}$ is never larger than the variance of $s_{\text{piv},1}$. It can be shown in a similar way that $\text{var}(s_{\text{piv},N}) \leq \text{var}(s_{\text{piv},1})$ and $\text{var}(s_{\text{piv},2N}) \leq \text{var}(s_{\text{piv},N})$ for any $N \in \mathbb{N}$.

As a compromise between accuracy and convenience it was suggested in Cruz-Orive (2005) to use $s_{\text{piv},4}$ in practice when manual measurements are required. In fact, it turns out that $s_{\text{piv},4}$ is the optimal approximation of s_{flo} in the case of general (triaxial) ellipsoidal particles with arbitrarily positioned reference point in its interior, in the sense that $s_{\text{piv},4} = s_{\text{flo}}$, no matter the choice of θ in (2.8). Detailed derivation of this result is presented in Appendix B. As a consequence, we expect $s_{\text{piv},4}$ to be very efficient in practice for ellipsoidal or nearly ellipsoidal particles.

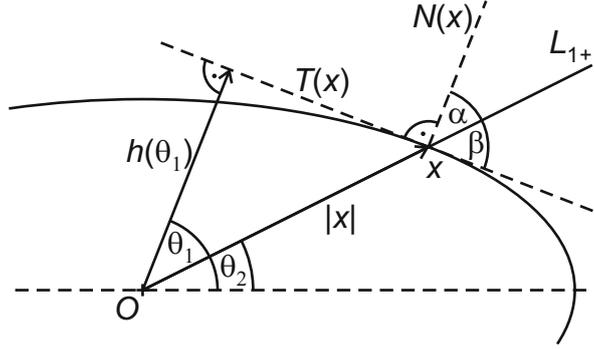


Figure 1: Notation used for the local geometry at a boundary point of a particle section.

2.3 The surfactor

Let $L_{1+} = L_{1+}(\theta)$ be a ray in L_2 emanating from O and uniquely determined by the angle $\theta \in [0, 2\pi)$ from a fixed axis in L_2 . For $x \in \partial(Y \cap L_2) \cap L_{1+}$ we define $\beta(x)$ to be the angle between L_{1+} and the line $T(x) \subset L_2$ tangent to $\partial(Y \cap L_2)$ at x , see Figure 1. Similarly, we define $\alpha(x)$ to be the angle between L_{1+} and the line $N(x) \subset L_2$ normal to $\partial(Y \cap L_2)$ at x . Note that $\alpha(x) + \beta(x) = \pi/2$.

The classical surfactor $s_{\text{sur},1}$ is an unbiased estimator of the particle surface area, see Jensen & Gundersen (1987) and Jensen (1998, Chapter 5.6). It is based on one isotropic ray $L_{1+}(\theta) \subset L_2$ through the reference point. Isotropic orientation of $L_{1+}(\theta)$ is obtained by letting the angle θ be uniformly distributed over $[0, 2\pi)$. There are two equivalent expressions for the classical surfactor $s_{\text{sur},1}$:

$$s_{\text{sur},1} = s_{\text{sur},1}(Y \cap L_2, L_{1+}(\theta)) = 4\pi|x_\theta|^2 [1 + (\pi/2 - \beta(x_\theta)) \cot \beta(x_\theta)], \quad (2.10)$$

$$s_{\text{sur},1} = s_{\text{sur},1}(Y \cap L_2, L_{1+}(\theta)) = 4\pi|x_\theta|^2 [1 + \alpha(x_\theta) \tan \alpha(x_\theta)], \quad (2.11)$$

where $x_\theta \in \partial(Y \cap L_2) \cap L_{1+}(\theta)$. For convex particles there is almost surely a single point in the intersection. Note that in general the distance from the origin O to a boundary point x and the angle $\alpha(x)$ or $\beta(x)$ give complementary information about the size and shape of the section $Y \cap L_2$ and either one cannot be calculated or estimated from the other.

Again, in order to reduce the variance of the estimator we consider another variant of the surfactor which uses measurements along more rays emanating from the reference point:

$$s_{\text{sur},N} = \frac{1}{N} \sum_{n=0}^{N-1} s_{\text{sur},1} \left(Y \cap L_2, L_{1+} \left(\theta + n \frac{2\pi}{N} \right) \right), \quad (2.12)$$

where θ is uniformly distributed over $[0, 2\pi)$. Clearly, $s_{\text{sur},N}$ is unbiased for any $N \in \mathbb{N}$.

Note that traditionally the surfactor is considered to use measurements along section lines passing through the reference point. However, this can be accommodated in our more general description using the number of rays emanating from the reference point. For example, the surfactor using two perpendicular section lines is identical to $s_{\text{sur},4}$ using measurements along four perpendicular rays. We use this description to emphasize the connection to the pivotal estimator.

If the information about the whole particle section $Y \cap L_2$ is available, for instance by automatic segmentation of a digital image, we can take advantage of it and estimate $S(Y)$ by the so-called integrated surfactor:

$$s_{\text{int}} = \frac{1}{2\pi} \int_0^{2\pi} s_{\text{sur},1}(Y \cap L_2, L_{1+}(\theta)) \, d\theta = 2 \int_0^{2\pi} |x_\theta|^2 [1 + \alpha(x_\theta) \tan \alpha(x_\theta)] \, d\theta, \quad (2.13)$$

where again $x_\theta \in \partial(Y \cap L_2) \cap L_{1+}(\theta)$. We can regard the integrated surfactor as a limit of $s_{\text{sur},N}$, $N \rightarrow \infty$. Cruz-Orive (2012) showed that the integrated surfactor and the flower estimator are identical if Y is a convex particle with boundary of class \mathcal{C}^2 or a convex polyhedral particle. More precisely, for any orientation of the section plane L_2 the right-hand sides of (2.3) and (2.13) coincide.

To the best of our knowledge the only studies using the surfactor to estimate surface area of particles are Karlsson & Cruz-Orive (1997) (tungsten particles in cemented carbide) and Tandrup, Gundersen & Jensen (1997) (cells in the dorsal root ganglion of the rat). The limited use of the surfactor is probably due to the singularity in the surfactor formula, see (2.10) and (2.11), and the necessity to perform manual angle measurements. In contrast to what has earlier been believed and also anticipated in Cruz-Orive (2005), our findings suggest that the surfactor is not affected much by the singularity and a possible inaccuracy in the angle measurements, see Section 3.

3 Performance of the surface area estimators – ellipsoidal particles

3.1 Basic setting

In order to investigate properties of the surface area estimators described in the previous section we looked into a situation where the ground truth is known. For this purpose we have chosen oblate and prolate ellipsoids for which the formulae for their surface areas are readily available.

Consider a given oblate or prolate ellipsoid Y with semiaxes lengths denoted by a , b and c and the reference point O in its centre. Given the orientation of the section plane L_2 (with $O \in L_2$) the lengths of the semiaxes of the section ellipse $Y \cap L_2$ are easily determined, see Appendix A. In the general case with reference point placed arbitrarily in the interior of Y another method must be employed, e.g. using the results of Jensen & Møller (1986).

In this basic setting we assume that all the necessary measurements are performed precisely. We generated $M = 1000$ independent isotropic sections through the reference point for each ellipsoid. For each section we calculated s_{flo} using (2.4) and $K = 100$ independent replications of $s_{\text{piv},2}$, $s_{\text{sur},2}$ and $s_{\text{sur},4}$. Note that $s_{\text{piv},4}$ is optimal in case of ellipsoidal particles in the sense that $s_{\text{piv},4} = s_{\text{flo}}$, see Section 2.2 and Appendix B. Hence, these two estimators are not distinguished in the following.

For all the estimators we estimated their relative bias and relative error (square root of the estimated mean squared error, divided by the correct value of the surface

Table 1: Estimated relative error of the estimators for oblate ellipsoids (semiaxes lengths satisfying $c < a = b$) and prolate ellipsoids (semiaxes lengths satisfying $c > a = b$).

Oblate ellipsoids						
$a : c$	1.5 : 1	2 : 1	3 : 1	4 : 1	6 : 1	8 : 1
s_{flo}	0.0887	0.120	0.124	0.117	0.0856	0.0743
$s_{\text{piv},2}$	0.222	0.351	0.488	0.560	0.625	0.655
$s_{\text{sur},2}$	0.234	0.405	0.689	0.923	1.33	1.62
$s_{\text{sur},4}$	0.0891	0.132	0.252	0.412	0.723	0.957
Prolate ellipsoids						
$a : c$	1 : 1.5	1 : 2	1 : 3	1 : 4	1 : 6	1 : 8
s_{flo}	0.149	0.289	0.525	0.729	1.03	1.27
$s_{\text{piv},2}$	0.258	0.457	0.763	1.03	1.39	1.65
$s_{\text{sur},2}$	0.268	0.507	0.970	1.44	2.24	3.07
$s_{\text{sur},4}$	0.149	0.295	0.590	0.911	1.53	2.07

area of the ellipsoid). For $s_{\text{piv},2}$, $s_{\text{sur},2}$ and $s_{\text{sur},4}$ we also estimated the within-section variance by

$$\frac{1}{M(K-1)} \sum_{i=1}^M \sum_{j=1}^K (s_{ij} - \bar{s}_i)^2, \quad (3.1)$$

where s denotes either $s_{\text{piv},2}$, $s_{\text{sur},2}$ or $s_{\text{sur},4}$ and $\bar{s}_i = \frac{1}{K} \sum_{j=1}^K s_{ij}$, s_{ij} being the j th replication in the i th section. Later we express the estimated within-section variance as a fraction of the total estimated variance.

In almost all cases the absolute value of the estimated relative bias was $< 1\%$, except for the prolate ellipsoids with semiaxes length ratios 4 : 1 or higher, where it was $< 2.6\%$. However, using more than 1000 section planes brings the estimated relative bias closer to 0. This is consistent with the theoretical unbiasedness of the estimators.

Table 1 summarizes the estimated relative errors of the estimators. The results indicate that the estimated relative errors tend to grow with the increasing elongation of the ellipsoid, the important exception being s_{flo} in the case of oblate ellipsoids. This observation is explained in detail in Section 3.1.1.

Increasing the number of rays (directions) used in the estimation resulted in significantly reduced variance of the estimators – the estimated relative error was always smaller for $s_{\text{sur},4}$ than for $s_{\text{sur},2}$, see Table 1. The same holds for $s_{\text{piv},4} = s_{\text{flo}}$ and $s_{\text{piv},2}$. Moreover, for slightly elongated ellipsoids $s_{\text{sur},4}$ showed very small values of the within-section variability, see Table 2. Virtually no improvement in precision could be gained by using s_{flo} instead of $s_{\text{sur},4}$ in that case.

Table 2: Estimated fraction of the within-section variance for oblate ellipsoids (semiaxes lengths satisfying $c < a = b$) and prolate ellipsoids (semiaxes lengths satisfying $c > a = b$). (All values are %.)

Oblate ellipsoids						
$a : c$	1.5 : 1	2 : 1	3 : 1	4 : 1	6 : 1	8 : 1
$s_{\text{piv},2}$	84.3	88.3	93.7	95.5	98.1	98.8
$s_{\text{sur},2}$	85.6	91.6	96.8	98.4	99.7	99.8
$s_{\text{sur},4}$	0.909	17.1	75.7	92.0	98.7	99.4
Prolate ellipsoids						
$a : c$	1 : 1.5	1 : 2	1 : 3	1 : 4	1 : 6	1 : 8
$s_{\text{piv},2}$	66.5	59.7	53.3	49.1	44.7	41.7
$s_{\text{sur},2}$	69.4	67.8	70.2	74.2	79.9	82.3
$s_{\text{sur},4}$	0.352	4.15	20.8	35.9	53.2	62.8

3.1.1 Variance of s_{flo}

Table 1 for oblate ellipsoids shows counter-intuitive improvement of the precision of s_{flo} , in terms of relative error or relative variance, with increasing elongation of the ellipsoid. This surprising phenomenon does not occur for prolate ellipsoids and is caused by the fact that the absolute variance of s_{flo} increases at a lower rate than the square of the surface area of the ellipsoid.

We have derived exact formulae for the variance of s_{flo} for both oblate and prolate ellipsoids with centrally positioned reference point, see (2.5) and (2.6). Appendix A provides the necessary details, together with a discussion of the asymptotic behaviour of the relative variance of s_{flo} . A surprising consequence of formula (2.5) is that the relative variance of s_{flo} for oblate ellipsoids goes to 0 in the limit with the semiaxes lengths ratio going to infinity, see Appendix A.

Only one source of variability affects the precision of s_{flo} , namely the random orientation of the section plane L_2 (called L_2 -variability in the following). For oblate ellipsoids it follows the above-mentioned trend – relative L_2 -variability goes to 0 in the limit with increasing elongation of the ellipsoid.

Apart from L_2 -variability, the estimators $s_{\text{piv},N}$ and $s_{\text{sur},N}$, $N \geq 1$, are also influenced by the random orientation of the rays in L_2 emanating from the reference point (called L_1 -variability). To be precise, $s_{\text{flo}} = s_{\text{piv},4k}$, $k \in \mathbb{N}$, and hence these variants of the pivotal estimator are not affected by L_1 -variability. For a given ellipsoid L_1 -variability can be reduced by using more section lines as demonstrated in Table 1 for $s_{\text{sur},2}$ and $s_{\text{sur},4}$. However, L_2 -variability cannot be reduced in any way.

For any given N and increasing elongation of the ellipsoid the L_1 -variability becomes the dominant source of variability of $s_{\text{sur},N}$ at some point, i.e. at some semiaxes length ratio, and L_2 -variability becomes negligible in comparison. This is partially demonstrated in Table 2 as within-section variability corresponds to L_1 -variability. Even for very large N the relative variance of $s_{\text{sur},N}$ for oblate ellipsoid first follows the decreasing trend of L_2 -variability but eventually (possibly for extremely

elongated ellipsoids) the L_1 -variability becomes dominant and the relative variance increases rapidly (results not shown).

In case of prolate ellipsoids the situation is similar but differs in one important aspect: the relative variance of s_{flo} increases with the increasing elongation of the ellipsoid and thus no improvement in precision is observed. Nevertheless, the reasoning about L_1 -variability becoming dominant at some point and the relative variance increasing rapidly still holds.

3.2 Influence of the position of the reference point

To assess the influence of the reference point position we performed the estimation using $s_{\text{flo}} = s_{\text{piv},4}$ and $s_{\text{sur},4}$ for oblate and prolate ellipsoids with semiaxes lengths ratio 2:1 and 8:1, respectively. Note that we focus on s_{flo} and $s_{\text{sur},4}$ rather than $s_{\text{piv},2}$ or $s_{\text{sur},2}$ because these estimators are expected to be preferred in most applications. Equation (2.4) was used to calculate s_{flo} .

Five positions of the reference point were chosen to lie on the minor semiaxis of the ellipsoid, the first one being at the centre of the ellipsoid and the last one at distance 0.95 times the minor semiaxis length from the centre. The latter choice was motivated by the aim to often obtain tangent angles β close to 0 (normal angles α close to $\frac{\pi}{2}$) which could make the surfactor unstable, see (2.10), (2.11). This was considered to be the worst case scenario with respect to the position of the reference point and the performance of $s_{\text{sur},4}$.

For each reference point location we generated 1000 independent isotropic sections through the reference point and computed 100 independent replications of $s_{\text{sur},4}$ in each section. We assume that all the necessary length and angle measurements are performed precisely.

Table 3 summarizes the estimated characteristics of s_{flo} and $s_{\text{sur},4}$. In almost all cases the absolute value of the estimated relative bias was $< 1\%$, except for the prolate ellipsoid with semiaxes length ratio 8 : 1, where the estimated relative bias was $< 3.1\%$ for all considered positions of reference point. However, using more than 1000 section planes brings the estimated relative bias closer to 0. This is consistent with the theoretical unbiasedness of s_{flo} and $s_{\text{sur},4}$, regardless of the position of the reference point in the interior of Y .

Concerning s_{flo} , some improvement in precision while moving the reference point towards the boundary is apparent in each inspected case, see Table 3. In most cases this trend reverses at some point and the relative errors increase again when the reference point gets very close to the boundary.

As discussed in Section 3.1.1, variability of s_{flo} is caused by the random orientation of the section plane only. Recall equation (2.4) and note that the term r^2 is constant for a given ellipsoid and given position of the reference point, and hence does not contribute to $\text{var}(s_{\text{flo}})$. The variance of s_{flo} thus consists of two components arising from $a_{L_2}^2$ and $b_{L_2}^2$. However, a_{L_2} and b_{L_2} are in general strongly correlated. This makes it difficult to assess their delicate interplay which determines $\text{var}(a_{L_2}^2 + b_{L_2}^2)$.

Figure 2 shows a concrete example (oblate ellipsoid with semiaxes lengths equal to 2, 2 and 1), where the changes in the values of the relative error are very prominent. The behaviour of $\text{var}(a_{L_2}^2)$, $\text{var}(b_{L_2}^2)$ and $\text{var}(a_{L_2}^2 + b_{L_2}^2)$ as the reference point

Table 3: Estimated characteristics of s_{flo} and $s_{\text{sur},4}$ for different ellipsoids and 5 reference points located along the minor semiaxis. Minor semiaxis length is 1, major semiaxis length is 2 and 8, respectively. Based on 1000 independent section planes through the reference point and 100 independent replications of $s_{\text{sur},4}$ for each section.

Oblate ellipsoids					
Position of ref. point	[0,0,0]	[0,0,0.50]	[0,0,0.75]	[0,0,0.85]	[0,0,0.95]
2 : 1, s_{flo} , rel. error	0.120	0.0490	0.0576	0.103	0.151
2 : 1, $s_{\text{sur},4}$, rel. error	0.132	0.131	0.168	0.178	0.201
2 : 1, $s_{\text{sur},4}$, w. s. variance (%)	17.1	85.8	87.6	65.8	39.2
8 : 1, s_{flo} , rel. error	0.0743	0.0558	0.0371	0.0352	0.0492
8 : 1, $s_{\text{sur},4}$, rel. error	0.957	0.961	1.07	1.05	0.967
8 : 1, $s_{\text{sur},4}$, w. s. variance (%)	99.4	99.8	99.9	99.9	99.7
Prolate ellipsoids					
Position of ref. point	[0,0,0]	[0.50,0,0]	[0.75,0,0]	[0.85,0,0]	[0.95,0,0]
1 : 2, s_{flo} , rel. error	0.289	0.242	0.206	0.234	0.284
1 : 2, $s_{\text{sur},4}$, rel. error	0.295	0.247	0.245	0.273	0.322
1 : 2, $s_{\text{sur},4}$, w. s. variance (%)	4.15	16.6	31.9	28.4	24.8
1 : 8, s_{flo} , rel. error	1.27	1.22	1.06	1.05	1.01
1 : 8, $s_{\text{sur},4}$, rel. error	2.07	2.00	2.03	1.95	1.88
1 : 8, $s_{\text{sur},4}$, w. s. variance (%)	62.8	64.1	69.5	70.9	72.2

moves from the centre of the ellipsoid towards its boundary is plotted.

Without loss of generality, let a_{L_2} denote the length of the major semiaxis of the section ellipse. If the reference point is in the centre of the ellipsoid, the length of the major semiaxis of the section ellipse is always $a_{L_2} = 2$, no matter the orientation of the section plane, and thus $\text{var}(a_{L_2}^2) = 0$. All the variability of $a_{L_2}^2 + b_{L_2}^2$ is due to $b_{L_2}^2$ in this case, b_{L_2} denoting the length of the minor semiaxis of the section ellipse.

As the reference point is moved along the minor semiaxis towards the boundary the variability of $a_{L_2}^2$ grows and at some point it becomes dominant over the variability of $b_{L_2}^2$. Nevertheless, the strong negative correlation between $a_{L_2}^2$ and $b_{L_2}^2$ (see the right part of Figure 2) reduces $\text{var}(a_{L_2}^2 + b_{L_2}^2)$ to very low values for a certain interval of the reference point positions.

For prolate ellipsoids, a similar trend was not observed. When moving the reference point from the centre of the ellipsoid towards its boundary $\text{var}(a_{L_2}^2 + b_{L_2}^2)$ somewhat decreases, with a small increase in the proximity of the boundary. The main difference from the oblate ellipsoids is that the variance of $b_{L_2}^2$ (squared minor semiaxis length of the section ellipse) is negligible compared to the variance of $a_{L_2}^2$ (squared major semiaxis length of the section ellipse) and hence the behaviour of $\text{var}(a_{L_2}^2 + b_{L_2}^2)$ is almost identical to that of $\text{var}(a_{L_2}^2)$.

Concerning $s_{\text{sur},4}$, the precision of the estimates in terms of the relative error does not increase much or at all while moving the position of the reference point from the centre of the ellipsoid towards its boundary, see Table 3. As discussed in Section 3.1.1, the variance of $s_{\text{sur},4}$ consists of two parts: the variance due to

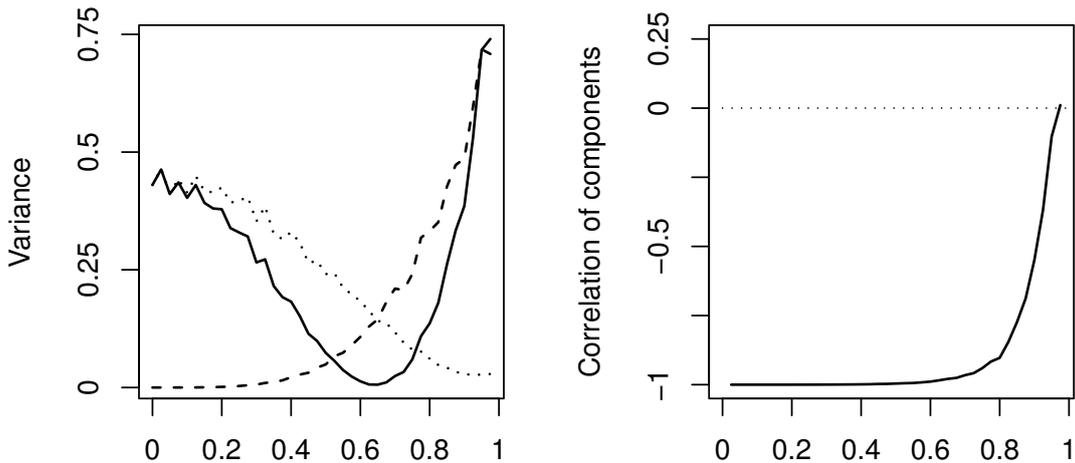


Figure 2: Illustration of the variance of s_{flo} : oblate ellipsoid with semiaxes lengths equal to 2, 2 and 1. Horizontal axis of the graphs – position of the reference point along the minor semiaxis, 0 being at the ellipsoid’s centre and 1 at the boundary. Left: variance of $a_{L_2}^2 + b_{L_2}^2$ (solid line), $a_{L_2}^2$ (dashed line, a_{L_2} denotes the length of the section ellipse’s major semiaxis) and $b_{L_2}^2$ (dotted line, b_{L_2} denotes the length of the section ellipse’s minor semiaxis). Right: estimated coefficient of correlation between $a_{L_2}^2$ and $b_{L_2}^2$. Based on 1000 section planes generated independently for each position of the reference point.

the random orientation of the section plane L_2 (corresponding to $\text{var}(s_{\text{flo}})$) and the additional variability due to the random orientation of the rays emanating from the reference point.

To further investigate the behaviour of the precision of $s_{\text{sur},4}$ we analysed changes in the distribution of the normal angles $\alpha(x)$ and the distance to the boundary $|x|$, caused by the change of the reference point location. For a detailed example consider for instance a prolate ellipsoid with semiaxes lengths 1, 1 and 8. In the $100 \cdot 1000$ replications of $s_{\text{sur},4}$ for each position of the reference point we recorded one of the four necessary angle measurements as well as the corresponding distance to the boundary (the other three measurements do not provide independent information).

Histograms illustrating the distribution of the recorded values of $\alpha(x)$ are shown in Figure 3, together with the graph of $1 + \alpha(x) \tan \alpha(x)$ giving a hint how much the corresponding angle measurements affect the resulting estimate. Even though Figure 3 shows considerable shift towards higher values of $\alpha(x)$ when the reference point moves to the boundary it is not dramatic in terms of increase of $1 + \alpha(x) \tan \alpha(x)$.

Figure 4 shows histograms of the recorded values of $|x|^2(1 + \alpha(x) \tan \alpha(x))$, where $|x|$ is the distance from the reference point to the boundary. Note that the shift towards high values while moving the reference point closer to the boundary is apparent but not substantial. This could be attributed to the fact that even though the normal angles $\alpha(x)$ generally increase the closer the reference point is to the boundary, the elliptic sections get smaller at the same time. Thus, the terms $|x|^2$ and $1 + \alpha(x) \tan \alpha(x)$ compensate each other on average.

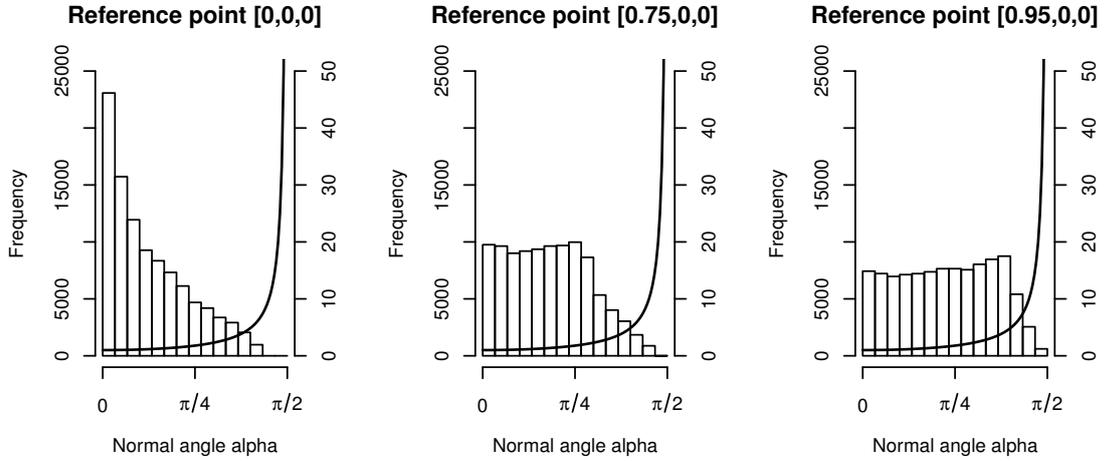


Figure 3: Histograms of recorded values of $\alpha(x)$. Prolate ellipsoid, semiaxes length ratio 1 : 8, different locations of the reference point. Based on 1000 independent section planes and 100 replications of $s_{\text{sur},4}$ for each section plane. Frequencies correspond to the left y -axis. Solid line – graph of function $1 + \alpha \tan \alpha$ with values corresponding to the right y -axis.

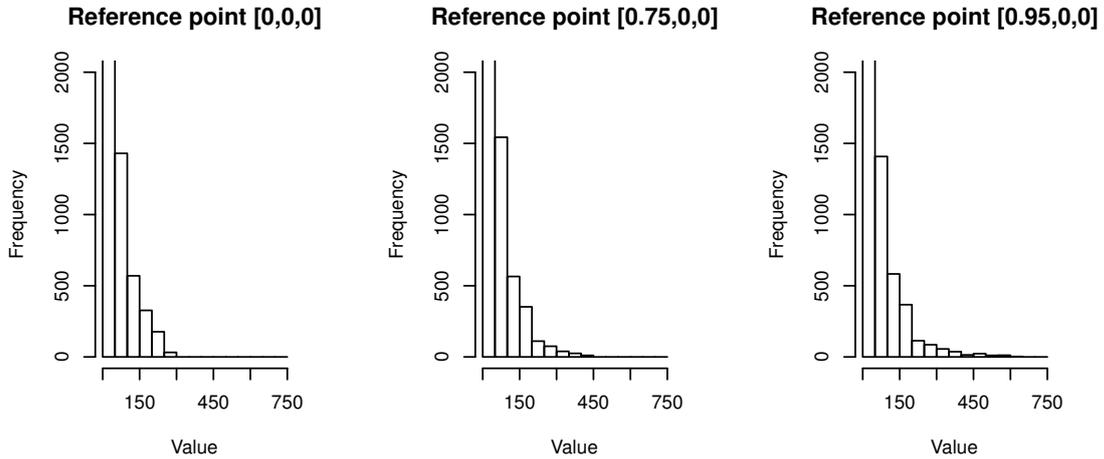


Figure 4: Histograms of recorded values of $|x|^2(1 + \alpha(x) \tan \alpha(x))$, where $|x|$ is the distance from the reference point to the boundary. Prolate ellipsoid, semiaxes length ratio 1 : 8, different locations of the reference point. Based on 1000 independent section planes and 100 replications of $s_{\text{sur},4}$ for each section plane. Vertical axis was truncated – frequencies corresponding to the first column exceed 97 000 in all cases.

3.3 The surfactor – inaccuracy in angle measurement

A characteristic property of the surfactor is that it requires angle measurements in order to determine $\alpha(x)$ or $\beta(x)$, see Figure 1. In practice, additional variability may be introduced to the estimation procedure involving the surfactor by possible inaccuracy of the angle measurements. Note that other estimators considered in this paper do not suffer from this shortcoming.

To investigate the effect of the possible inaccuracy in angle measurements on the resulting surfactor estimates we have added a random error ϵ to the tangent line orientation. This results in corrupted measurement of $\alpha(x)$ and $\beta(x)$. We assume that the additive error term ϵ is a random variable uniformly distributed over $(-\delta, \delta)$, $\delta \geq 0$, and that the necessary length measurements are performed precisely (we focus on inaccuracy of angle measurements as it is supposed to have a larger impact on the resulting estimates due to the singularity in the surfactor formula).

We generated 1000 independent isotropic sections of different ellipsoids with the reference point in their centre and computed 100 independent replications of $s_{\text{sur},4}$ in each section. Table 4 shows the resulting estimates of the characteristics of $s_{\text{sur},4}$.

Table 4: Estimated characteristics of $s_{\text{sur},4}$ with inaccurate angle measurements for oblate and prolate ellipsoids with different semiaxes lengths ratios. The additive error term is a random variable uniformly distributed over $(-\delta, \delta)$. Based on 1000 isotropic sections, each with 100 independent replications of $s_{\text{sur},4}$. Reference point is in the ellipsoid’s centre in each case.

$\delta = 0^\circ$	oblate, 2 : 1	oblate, 8 : 1	prolate, 1 : 2	prolate, 1 : 8
Rel. bias	-0.00107	0.00312	-0.00885	-0.0220
Rel. error	0.132	0.957	0.295	2.07
w.s. variance (%)	17.1	99.4	4.15	62.8
$\delta = 5^\circ$	oblate, 2 : 1	oblate, 8 : 1	prolate, 1 : 2	prolate, 1 : 8
Rel. bias	$1.88 \cdot 10^{-4}$	0.0197	0.00733	0.0659
Rel. error	0.140	1.02	0.301	2.27
w.s. variance (%)	18.4	99.4	4.81	64.3
$\delta = 10^\circ$	oblate, 2 : 1	oblate, 8 : 1	prolate, 1 : 2	prolate, 1 : 8
Rel. bias	0.00867	0.0937	0.00192	0.107
Rel. error	0.140	1.20	0.306	2.62
w.s. variance (%)	24.7	99.8	6.02	68.1
$\delta = 15^\circ$	oblate, 2 : 1	oblate, 8 : 1	prolate, 1 : 2	prolate, 1 : 8
Rel. bias	0.0227	0.790	0.0190	0.607
Rel. error	0.146	23.2	0.317	41.7
w.s. variance (%)	34.7	99.9	8.55	99.2

The results indicate that for slightly elongated ellipsoids the precision of $s_{\text{sur},4}$ is not affected much even by severe inaccuracy of the angle measurements. On the other hand, for considerably elongated ellipsoids the estimates become virtually useless

if $\delta > 10^\circ$. Note also that even though the additive error term has a symmetric distribution, its effect is not because of the shape of the function $1 + \alpha(x) \tan \alpha(x)$. Overestimating the angle $\alpha(x)$ has a larger impact on the resulting estimate than underestimating it. This results in a positive bias of $s_{\text{sur},4}$, see Table 4.

4 Semi-automatic estimation

Studies involving estimation of geometric properties of particles by manual measurement can be very time-consuming – especially if both length and angle measurements are required, as is the case in surface area estimation using the surfactor. If digital images of the particle sections are recorded the workload can be greatly reduced by taking advantage of automated segmentation of particle boundaries by means of digital image analysis. In this section we propose both automatic and semi-automatic procedures for estimating the surface area of a convex particle, using the automatically segmented boundaries.

4.1 The automatic flower estimator

Let \hat{Y}_2 denote the estimate of the true particle section $Y \cap L_2$ provided by automatic segmentation and assume that the reference point O is contained in \hat{Y}_2 . It is natural to use information about the whole particle section and define the automatic flower estimator as

$$s_{\text{aut}} = s_{\text{flo}}(\hat{Y}_2), \quad (4.1)$$

In practice this estimator will often be approximated by $s_{\text{piv},N}(\hat{Y}_2)$ for a suitable number of directions N . We denote the resulting estimator $s_{\text{aut}}^{\text{piv}}$. Moreover, direct or approximate methods of computing the flower estimator given in Cruz-Orive (2011) may be employed.

Naturally, the approximation of the flower estimator using the surfactor may be used, resulting in the estimator denoted $s_{\text{aut}}^{\text{sur}}$ in the following. Note that this estimator could also be used to estimate the surface area of non-convex particles as it does not depend on the assumption of convexity. In that case it should not be considered as an approximation to the flower estimator.

Properties of s_{aut} depend strongly on the quality of the segmentation. The automatic flower estimator s_{aut} provides unbiased estimates if the segmentation is flawless ($\hat{Y}_2 = Y \cap L_2$). However, if the segmentation is incorrect, s_{aut} may be heavily biased.

4.2 The semi-automatic flower estimator

To overcome the possible issues with the bias of the automatic estimator we propose a semi-automatic version s_{semi} , a combination of the automatic flower estimator s_{aut} and the pivotal estimator $s_{\text{piv},4}$ or the surfactor $s_{\text{sur},4}$. We denote the resulting estimators $s_{\text{semi}}^{\text{piv}}$ and $s_{\text{semi}}^{\text{sur}}$, respectively.

Let the process of automatic segmentation and estimation using s_{aut} be supervised by an expert who intervenes only if the segmentation of particle boundary

is not satisfactory. In that case the expert performs manually the necessary measurements required to calculate $s_{\text{piv},4}$ or $s_{\text{sur},4}$. If the segmentation is judged to be satisfactory, $s_{\text{semi}} = s_{\text{aut}}$. Otherwise $s_{\text{semi}} = s_{\text{piv},4}$ or $s_{\text{sur},4}$, respectively.

Unlike the flower estimator, the semi-automatic flower estimator s_{semi} might be biased. Let A be the event that the automatic segmentation is judged satisfactory by the expert. Using the same technique as in Hansen, Nyengaard, Andersen & Jensen (2011) it can be shown that the bias of s_{semi} can be expressed as

$$\text{bias}(s_{\text{semi}}) = p \mathbb{E}(s_{\text{aut}} - s_{\text{flo}}|A). \quad (4.2)$$

Here p is the probability of the segmentation being accepted as satisfactory and $\mathbb{E}(s_{\text{aut}} - s_{\text{flo}}|A)$ is the mean difference of s_{aut} and s_{flo} among the particles with satisfactory segmentation. Note that the bias will be small because $s_{\text{aut}} - s_{\text{flo}}$ is small for particles with satisfactory segmentation.

Similarly, it can be shown that the mean squared error (MSE) of s_{semi} can be expressed as

$$\text{MSE}(s_{\text{semi}}) = p \text{MSE}(s_{\text{aut}}|A) + (1 - p) \text{MSE}(s_{\cdot,4}|A^c), \quad (4.3)$$

where $\text{MSE}(s_{\text{aut}}|A)$ is the mean squared error of s_{aut} among the particles with segmentation judged as satisfactory and $\text{MSE}(s_{\cdot,4}|A^c)$ is the mean squared error of $s_{\text{piv},4}$ or $s_{\text{sur},4}$, respectively, among the cells for which the segmentation is not satisfactory.

Estimates of bias and mean squared error of s_{semi} can be obtained in a straightforward manner by plugging in the empirical estimates of p , $\mathbb{E}(s_{\text{aut}} - s_{\text{flo}}|A)$, $\text{MSE}(s_{\text{aut}}|A)$ and $\text{MSE}(s_{\cdot,4}|A^c)$ into equations (4.2) and (4.3), respectively.

5 A comparative study

In order to assess the performance of automatic and semi-automatic estimators proposed in Section 4 we analysed data from a study of somatostatin positive inhibitory interneurons from mice hippocampi, observed by optical fluorescent microscopy. A characteristic feature of these cells are dendrites which were not regarded as part of the cell body in this study and do not contribute to the cell surface area. The study was originally presented and described in detail in Hansen, Nyengaard, Andersen & Jensen (2011).

The data constitute of points on the boundaries of central cell sections of 91 individual cells. The boundaries were segmented automatically and the xy -coordinates of points on the individual boundaries were recorded. Points not originating from dendrites or from other cells visible in other layers of the thick section were considered to lie on the true boundary of the cell.

To represent the true boundary of each sampled cell section, a polynomial smoothing spline was fitted to the xy -coordinates of points on the true cell section boundary. Because these nerve cells do not have a natural reference point, the centre of mass of the recorded segmented true boundary points was used as a reference point. It approximates the centre of mass of the entire cell and has nearly central position in the cell section.

We estimated the mean cell surface area by $s_{\text{piv},4}$, $s_{\text{sur},4}$ and $s_{\text{sur},200}$ using the true boundaries of the cell sections. 100 independent replications of $s_{\text{piv},4}$ and $s_{\text{sur},4}$ were calculated for each cell section. The most reliable results are obtained by $s_{\text{sur},200}$ since a closer look reveals that some of the true cell sections include very small concave boundary elements and hence are not perfectly convex, see Figure 5. The surfactor is not affected by this problem but the pivotal estimator in fact uses the convex hull of the cell section instead of the true (non-convex) cell section. Using 200 rays in $s_{\text{sur},200}$ turned out to be sufficient for the purpose of approximating s_{flo} since using even more section lines yielded virtually identical results.

To assess the performance of the semi-automatic flower estimator proposed in Section 4 we computed the estimates using $s_{\text{semi}}^{\text{piv}}$ and $s_{\text{semi}}^{\text{sur}}$. Approximation of s_{aut} by $s_{\text{piv},200}$ was used to analyse cell sections with satisfactory segmentation (these exhibit only negligible deviations from convexity).

We also approximated $s_{\text{aut}}^{\text{sur}}$ by $s_{\text{sur},200}$ applied to the whole segmented boundary of the cell section described by a polynomial smoothing spline fitted to all segmented boundary points, i.e. including points originating from dendrites or other cells. To avoid confusion we denote this approximation of the automated surfactor by $s_{\text{aut}}^{\text{sur}}$ in the following. We do not present results obtained using $s_{\text{aut}}^{\text{piv}}$ due to the non-convexity of badly segmented cell sections. We assume that all the necessary measurements were performed precisely.

To decide if the automatic boundary segmentation was satisfactory we used a distance d on the set of subsets of L_2 . The segmentation was judged satisfactory if

$$d(Y \cap L_2, \tilde{Y}) < d_0, \quad (5.1)$$

where \tilde{Y} is the automatically segmented boundary, $Y \cap L_2$ is the true cell section boundary and $d_0 \geq 0$. This enables determining if s_{semi} takes the value of s_{aut} or $s_{\text{sur},4}$ for each cell section. The distance considered in this study was

$$d(B_1, B_2) = \frac{|s_{\text{sur},200}(B_1) - s_{\text{sur},200}(B_2)|}{s_{\text{sur},200}(B_1)}. \quad (5.2)$$

The motivation was that $s_{\text{sur},200}$ provides a very good approximation of s_{flo} while $s_{\text{piv},N}$ disregards the possible non-convexity of the cell sections. Alternatively, a distance using the difference of areas of B_1 and B_2 could be employed.

With $d_0 = 0.15$ a total of 59 cell boundaries had a satisfactory segmentation according to this criterion. Hence the probability p of satisfactory segmentation was estimated to be $\hat{p} = 59/91 \doteq 0.648$. Figure 5 shows examples of both correctly and incorrectly segmented cell section boundaries together with the true cell section boundaries approved by an expert.

Figure 6 shows the estimates using $s_{\text{sur},4}$, $s_{\text{piv},4}$, $s_{\text{aut}}^{\text{sur}}$ and $s_{\text{semi}}^{\text{piv}}$ for each sampled cell plotted against $s_{\text{sur},200}$ in a double logarithmic scale. For clarity only 1 estimate using $s_{\text{sur},4}$ and $s_{\text{piv},4}$ from the 100 repetitions is plotted for each cell. The figures show that both $s_{\text{sur},4}$ and $s_{\text{piv},4}$ are in most cases nearly identical to $s_{\text{sur},200}$ which we regard as the most precise and reliable estimator in this study. On the other hand, $s_{\text{aut}}^{\text{sur}}$ yields very high estimates for some cell sections due to incorrect boundary segmentation (see Figure 5 for typical examples of unsatisfactory boundary segmentation). This results in a considerable positive bias of s_{aut} .

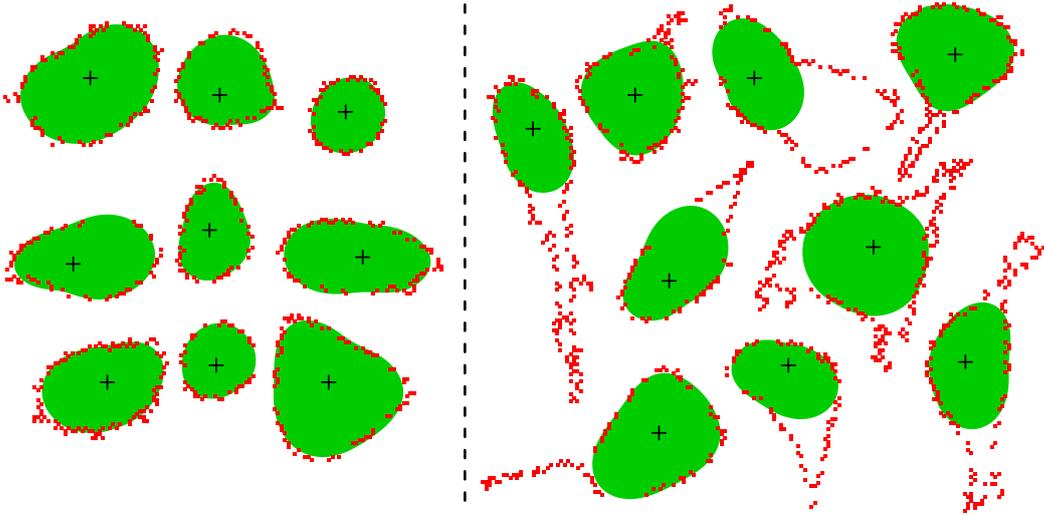


Figure 5: Hippocampus cells data: examples of cell sections with correctly segmented boundary (left) and with boundary segmentation judged as unsatisfactory (right). Red points – segmented boundary points, black crosses – the reference points, green areas – true cell sections approved by an expert.

Estimates of the characteristics of the estimators are summarized in Table 5. To obtain these estimates we used $\hat{\mu} \doteq 645.5 \mu\text{m}^2$, the mean value of $s_{\text{sur},200}$, as the true value of the mean cell surface area μ . The results are in agreement with the observation that the cell sections have very regular boundaries with no considerable elongation, unless the boundary was incorrectly segmented (which affects $s_{\text{aut}}^{\text{sur}}$ only). Thus, low values of estimated relative errors are to be expected for all estimators except $s_{\text{aut}}^{\text{sur}}$. Note also the small positive bias of both variants of s_{semi} caused by the fact that even though the segmentation was judged to be satisfactory for a given cell section, it may still be somewhat different from the true cell section boundary used by $s_{\text{sur},200}$.

Table 5: Hippocampus cells data: estimated characteristics of the estimators.

	$s_{\text{piv},4}$	$s_{\text{sur},4}$	$s_{\text{sur},200}$	$s_{\text{aut}}^{\text{sur}}$	$s_{\text{semi}}^{\text{piv}}$	$s_{\text{semi}}^{\text{sur}}$
Mean (μm^2)	645.2	645.6	645.5	868.4	655.0	655.5
Relative bias	<0.01	<0.01	0	0.3453	0.01667	0.01670
Relative error	0.3574	0.3593	0.3579	0.8712	0.3643	0.3649
w.s. variance (%)	0.395	1.05				

Using formulae for the relative error of the estimators similar to those given in Hansen, Nyengaard, Andersen & Jensen (2011) we can estimate the number of cells needed to be analyzed in order to obtain a given relative error of the estimated mean cell surface area. The results are shown in Figure 7 for $s_{\text{piv},4}$ and $s_{\text{semi}}^{\text{piv}}$, respectively (the curves corresponding to $s_{\text{sur},4}$ and $s_{\text{sur},200}$ cannot be visually distinguished from the one corresponding to $s_{\text{piv},4}$ and are not shown; similarly, the curve corresponding

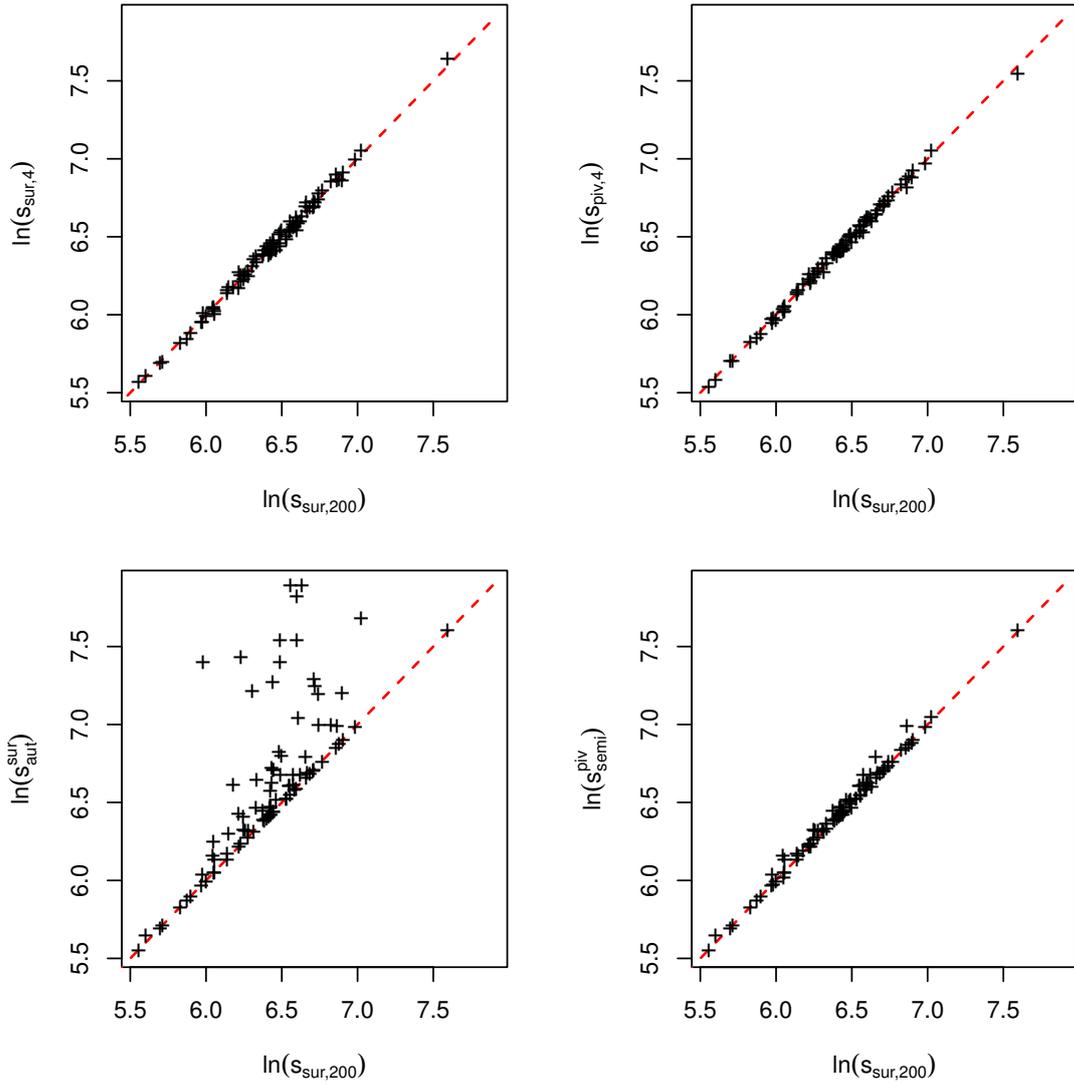


Figure 6: Hippocampus cells data: surface area estimates $s_{\text{sur},4}$, $s_{\text{piv},4}$, $s_{\text{sur}}^{\text{sur}}_{\text{aut}}$ and $s_{\text{piv}}^{\text{piv}}_{\text{semi}}$ plotted against $s_{\text{sur},200}$. The scale is $\ln(\mu\text{m}^2)$.

to $s_{\text{semi}}^{\text{sur}}$ is not shown as it coincides with the curve corresponding to $s_{\text{semi}}^{\text{piv}}$). For example, to obtain the relative error of 0.05 it is needed to analyze 51, 51, 52, 60 and 60 cell sections using $s_{\text{piv},4}$, $s_{\text{sur},200}$, $s_{\text{sur},4}$, $s_{\text{semi}}^{\text{piv}}$ or $s_{\text{semi}}^{\text{sur}}$, respectively.

Since both variants of s_{semi} require analysis of more cells than $s_{\text{piv},4}$ or $s_{\text{sur},4}$ it appears to be inferior to them. However, only about 35 % of the cells in this study needed to be analysed by $s_{\text{piv},4}$ or $s_{\text{sur},4}$ due to unsatisfactory boundary segmentation. Hence using the semi-automatic estimator resulted in significant reduction of the manual workload.

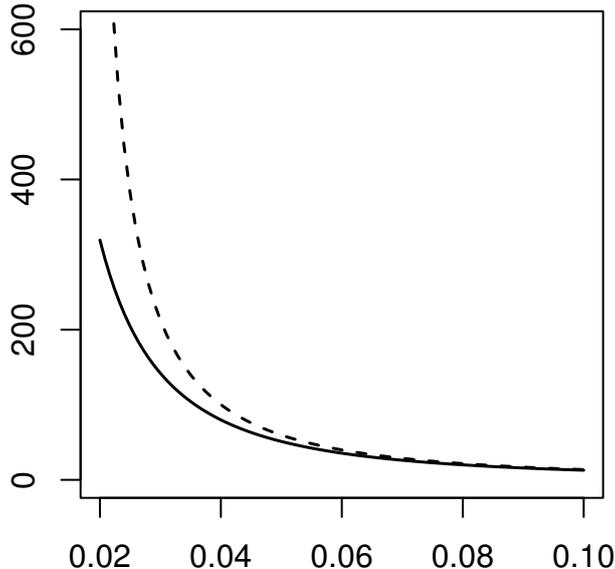


Figure 7: Hippocampus cells data: number of cells needed to be analyzed for $s_{\text{piv},4}$ (solid line) and $s_{\text{semi}}^{\text{piv}}$ (dashed line) in order to achieve a given level of the relative error. The curves corresponding to $s_{\text{sur},4}$ and $s_{\text{sur},200}$ cannot be visually distinguished from the one corresponding to $s_{\text{piv},4}$ (solid line) and are not shown. Similarly, the curve corresponding to $s_{\text{semi}}^{\text{sur}}$ is not shown as it coincides with the curve corresponding to $s_{\text{semi}}^{\text{piv}}$ (dashed line).

6 Conclusions

In the present paper we have considered the task of estimating the surface area of convex particles and discussed two different approximations of the flower estimator, namely the pivotal estimator and the surfactor. We have compared their performance both for artificial data (ellipsoidal particles with known surface area) and real data coming from a study of somatostatin inhibitory interneurons from mice hippocampi. The results indicate that the pivotal estimator using measurements of the support function in four perpendicular directions is very efficient and that it is a powerful approximation of the flower estimator.

We have presented a number of new results concerning the variance of the flower estimator in the form of exact formulae (see equations (2.5) and (2.6)) and empirical observations (e.g. that choosing the centrally positioned reference point is not optimal for ellipsoidal particles in the sense of minimizing the variance of the estimator, see Section 3.2). Moreover, we have shown that the pivotal estimator $s_{\text{piv},4}$

is optimal for ellipsoidal particles with arbitrarily positioned reference point, in the sense that it is equal to the flower estimator, no matter the directions at which the measurements of the support function are performed (as long as the four directions in the section plane are perpendicular to each other, see Appendix B).

Concerning the surfactor, our findings suggest that it performs well in case of particles which are not extremely elongated, it is not very sensitive to the location of the reference point within the particle and it is not affected much by possible inaccuracy of the necessary angle measurements. The singularity introduced to the surfactor formula (2.10) by the term $\cot \beta$, where β is the angle between a random line passing through the reference point and a tangent to the boundary of the particle at the intersection point, should not prevent the surfactor from being used. Values of the angle β close to 0 are not frequently encountered even in case of very elongated particles and reference points close to the boundary.

We have also presented an automatic version of the flower estimator and a semi-automatic procedure of surface area estimation supervised by an expert who intervenes only if the automatically detected boundary is judged to be incorrect. It was found that in the somatostatin inhibitory interneurons study the number of cells needed to be analysed for obtaining, for instance, a 5% precision of the estimate of the mean cell surface area is 51, 51, 52, 60 and 60, depending on whether the estimator s_{flo} , $s_{\text{piv},4}$, $s_{\text{sur},4}$, $s_{\text{semi}}^{\text{piv}}$ or $s_{\text{semi}}^{\text{sur}}$ is used, respectively.

In this concrete study we can conclude that in order to obtain the required precision, it is needed to sample (roughly) the same number of cells, irrespectively of whether the flower estimator s_{flo} , the pivotal estimator $s_{\text{piv},4}$ or the surfactor $s_{\text{sur},4}$ is used. In manual analysis, the pivotal estimator is to be preferred to the surfactor because even though the two estimators have roughly the same precision, $s_{\text{piv},4}$ requires a factor two smaller workload.

For the same reason, the pivotal estimator is to be preferred in the semi-automatic setting. Whether a semi-automatic approach is to be preferred to a manual approach depends on the reduction in workload associated with the semi-automatic approach. Note that the number of cells needed for obtaining a 5% precision was somewhat higher in the semi-automatic case than in the manual case due to a small bias of the semi-automatic flower estimator. It was, however, only needed to manually analyze a fraction of these cells, about 35% in the above-mentioned study of somatostatin positive inhibitory interneurons.

In case of non-convex particles the surfactor will find its use (the pivotal or the flower estimator would estimate the surface area of the convex hull of the particle, thus underestimating the actual surface area). In analogy to the semi-automatic flower estimator, a semi-automatic version of the surfactor may be employed in that case, using the integrated surfactor for particle sections with satisfactory segmentation of the boundary and the classical version of the surfactor for the sections with unsatisfactory segmentation. Alternatively, the method of surface area estimation based on the invariator principle without the assumption of convexity, given in Cruz-Orive (2005), could be used. The upcoming paper Thórisdóttir & Kiderlen (2012) will deal with this topic.

7 Acknowledgement

This project has been supported by the Czech Science Foundation, project no. P201/10/0472, the grant SVV 265 315 and by Centre for Stochastic Geometry and Advanced Bioimaging, funded by a grant from The Villum Foundation.

References

- [1] Cruz-Orive, L.M. (2005). A new stereological principle for test lines in three-dimensional space. *J. Microsc.* 219, 18–28.
- [2] Cruz-Orive, L.M. (2008). Comparative precision of the pivotal estimators of particle size. *Image Anal. Stereol.* 27, 17–22.
- [3] Cruz-Orive, L.M. (2011). Flowers and wedges for the stereology of particles. *J. Microsc.* 243, 86–102.
- [4] Cruz-Orive, L.M. (2012). Uniqueness properties of the invariator, leading to simple computations. *Image Anal. Stereol.*, in press.
- [5] Hansen, L.V., Nyengaard, J.R., Andersen, J.B. & Jensen, E.B.V. (2011). The semi-automatic nucleator. *J. Microsc.* 242, 206–215.
- [6] Jensen, E.B.V. (1998). *Local Stereology*, World Scientific, Singapore.
- [7] Jensen, E.B.V. & Gundersen, H.J.G. (1987). Stereological estimation of surface area of arbitrary particles. *Acta Stereol.* 6, 25–30.
- [8] Jensen, E.B.V. & Gundersen, H.J.G. (1989). Fundamental stereological formulae based on isotropically orientated probes through fixed points with applications to particle analysis. *J. Microsc.* 153, 249–267.
- [9] Jensen, E.B.V. & Møller, J. (1986). Stereological versions of integral geometric formulae for n -dimensional ellipsoids. *J. Appl. Probab.* 23, 1031–1037.
- [10] Karlsson, L.M. & Cruz-Orive, L.M. (1997). Estimation of mean particle size from single sections. *J. Microsc.* 186, 121–132.
- [11] Kubínová, L. & Janáček, J. (1998). Estimating surface area by the isotropic fakir method from thick slices cut in an arbitrary direction. *J. Microsc.* 191, 201–211.
- [12] Noumeir, R. (2000). Fitting of ellipses from support function measurements. *ICASSP* 4, 2294–2297.
- [13] Tandrup, T., Gundersen, H.J.G. & Jensen, E.B.V. (1997). The optical rotator. *J. Microsc.* 186, 108–120.
- [14] Thórisdóttir, Ó. & Kiderlen, M. (2012). Estimating the surface area of non-convex particles from central planar sections. In preparation.

A Variance of the flower estimator – ellipsoids of revolution

We are interested in the behaviour of $\text{var}(s_{\text{flo}})$ with increasing elongation of the ellipsoid. Assume that Y is an oblate or prolate ellipsoid and the reference point O is the centre of the ellipsoid. Consider an isotropic section plane L_2 through O . The planar section $Y \cap L_2$ is an ellipse with semiaxes lengths a_{L_2} and b_{L_2} . The integrated surfactor is given by the formula $s_{\text{flo}} = 2\pi(a_{L_2}^2 + b_{L_2}^2)$, see (2.4). Hence the task of determining $\text{var}(s_{\text{flo}})$ reduces to determining $\text{var}(a_{L_2}^2 + b_{L_2}^2)$ under the assumption of isotropic orientation of the section plane.

A.1 Determining the elliptic section

First consider an oblate ellipsoid with semiaxes lengths $a = b > c$ and fix $c = 1$. Any planar section of this ellipsoid through its centre is an ellipse with the major semiaxis length $a_{L_2} = a$. The minor semiaxis length b_{L_2} is variable and depends on the orientation of the section plane L_2 . Hence, a_{L_2} being constant, $\text{var}(a_{L_2}^2 + b_{L_2}^2) = \text{var}(b_{L_2}^2)$.

For simplicity assume that the orientation of the ellipsoid is such that the minor axis is parallel to the coordinate z -axis. The orientation of the section plane L_2 is determined by its normal vector which is given by a point on the unit hemisphere. Due to rotational symmetry of the ellipsoid b_{L_2} is fully determined by γ , the angle between the normal vector of L_2 and the coordinate z -axis. Figure 8 illustrates the dependence of b_{L_2} on $\gamma \in [0, \pi/2]$.

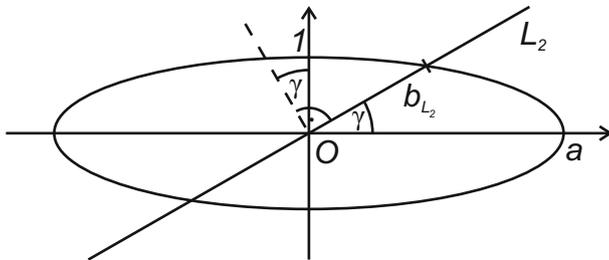


Figure 8: Oblate ellipsoid with semiaxes lengths equal to 1 and $a = b > 1$. The figure shows the planar section through the centre of the ellipsoid parallel to the minor axis of the ellipsoid and perpendicular to L_2 . This elliptic section is sufficient to determine b_{L_2} , the minor semiaxis length of the section ellipse in L_2 .

Points on the boundary of the auxiliary ellipse with semiaxes lengths equal to 1 and $a > 1$ (see Figure 8) can be parametrized by $x = a \cos t$, $y = \sin t$, $t \in [0, 2\pi)$. The parameter t for the point corresponding to the polar angle γ is given by

$$t(\gamma) = \arctan(a \tan \gamma).$$

Hence, $b_{L_2}^2$ can be expressed as

$$b_{L_2}^2 = a^2 \cos^2 t(\gamma) + \sin^2 t(\gamma) = 1 + (a^2 - 1) \cos^2 t(\gamma).$$

Under the assumption of isotropic orientation of L_2 , the angle γ can be expressed as $\gamma = \arccos V$, where V is a random variable with uniform distribution in $[0, 1]$. We get, using $B_{L_2}^2$ to denote the random variable corresponding to $b_{L_2}^2$,

$$B_{L_2}^2 = 1 + (a^2 - 1)Y, \text{ where } Y = \cos^2\{\arctan[a \tan(\arccos V)]\}.$$

Using forward-inverse trigonometric identities we can express Y as

$$Y = \phi(V) = \frac{V^2}{a^2 - (a^2 - 1)V^2}.$$

We are interested in $\text{var}(Y)$ as $\text{var}(s_{\text{flo}}) = 4\pi^2 \text{var}(B_{L_2}^2) = 4\pi^2(a^2 - 1)^2 \text{var}(Y)$. Since ϕ is an increasing, differentiable function $(0, 1) \rightarrow (0, 1)$, we can easily determine the probability density function of Y :

$$f_Y^O(y) = I_{(0,1)}(y) \frac{a}{2\sqrt{y}} (1 + y(a^2 - 1))^{-3/2}, \quad y \in \mathbb{R},$$

and thus the first two moments of Y and thereby $\text{var}(Y)$.

Now consider a prolate ellipsoid with semiaxes lengths $a = b < c$ and fix $a = b = 1$. Any planar section of this ellipsoid through its centre is an ellipse with the minor semiaxis length $b_{L_2} = 1$. The major semiaxis length a_{L_2} is variable and depends on the orientation of the section plane L_2 .

By a procedure similar to the one presented above we get that $\text{var}(s_{\text{flo}}) = 4\pi^2 \text{var}(a_{L_2}^2) = 4\pi^2(c^2 - 1)^2 \text{var}(Y)$, where Y is an auxiliary random variable with probability density function

$$f_Y^P(y) = I_{(0,1)}(y) \frac{c^2}{2\sqrt{1-y}} (1 + y(c^2 - 1))^{-3/2}, \quad y \in \mathbb{R}.$$

From this we can calculate the first two moments of Y and thus $\text{var}(Y)$.

A.2 Calculating $\mathbb{E}Y$ – oblate ellipsoid

For ease of exposition we denote $p = a^2 - 1$. We assume that $a > 1$ and thus $p > 0$. Then

$$\mathbb{E}Y = \int_0^1 y f_Y^O(y) dy = \frac{\sqrt{p+1}}{2} \int_0^1 y^{1/2} (1 + py)^{-3/2} dy.$$

By substituting $t = \sqrt{py}$ we get

$$\mathbb{E}Y = \frac{\sqrt{p+1}}{p^{3/2}} \int_0^{\sqrt{p}} t^2 (1 + t^2)^{-3/2} dt.$$

Then we substitute $t = \tan u$ and write $P = \arctan(\sqrt{p})$:

$$\begin{aligned} \int_0^{\sqrt{p}} t^2 (1 + t^2)^{-3/2} dt &= \int_0^P \frac{1}{\cos u} du - \int_0^P \cos u du \\ &= \left[\log \left(\tan u + \frac{1}{\cos u} \right) \right]_{u=0}^P - [\sin u]_{u=0}^P \\ &= \log \left(\sqrt{p} + \sqrt{p+1} \right) - \frac{\sqrt{p}}{\sqrt{p+1}} = \text{arcsinh}(\sqrt{p}) - \frac{\sqrt{p}}{\sqrt{p+1}}. \end{aligned}$$

In the last line we used forward-inverse trigonometric identities again. Finally,

$$\mathbb{E}Y = \frac{\sqrt{p+1}}{p^{3/2}} \operatorname{arcsinh}(\sqrt{p}) - \frac{1}{p} = \frac{a \operatorname{arcsinh}(\sqrt{a^2-1})}{(a^2-1)^{3/2}} - \frac{1}{a^2-1}.$$

A.3 Calculating $\mathbb{E}Y^2$ – oblate ellipsoid

Again we denote $p = a^2 - 1$ and assume that $a > 1$ and $p > 0$. Then

$$\mathbb{E}Y^2 = \int_0^1 y^2 f_Y^O(y) dy = \frac{\sqrt{p+1}}{2} \int_0^1 y^{3/2} (1+py)^{-3/2} dy.$$

By substituting $t = \sqrt{py}$ we get

$$\mathbb{E}Y^2 = \frac{\sqrt{p+1}}{p^{5/2}} \int_0^{\sqrt{p}} t^4 (1+t^2)^{-3/2} dt.$$

Then we substitute $t = \tan u$ and write $P = \arctan(\sqrt{p})$:

$$\int_0^{\sqrt{p}} t^4 (1+t^2)^{-3/2} dt = \int_0^P \frac{1}{\cos^3 u} du - 2 \int_0^P \frac{1}{\cos u} du + \int_0^P \cos u du.$$

By integration by parts we get $\int_0^P \frac{1}{\cos^3 u} du = \frac{1}{2} \left[\frac{\tan u}{\cos u} \right]_{u=0}^P + \frac{1}{2} \int_0^P \frac{1}{\cos u} du$ and thus

$$\begin{aligned} \int_0^{\sqrt{p}} t^4 (1+t^2)^{-3/2} dt &= \frac{1}{2} \left[\frac{\tan u}{\cos u} \right]_{u=0}^P - \frac{3}{2} \left[\log \left(\tan u + \frac{1}{\cos u} \right) \right]_{u=0}^P + [\sin u]_{u=0}^P \\ &= \frac{\sqrt{p}\sqrt{p+1}}{2} - \frac{3}{2} \operatorname{arcsinh}(\sqrt{p}) + \frac{\sqrt{p}}{\sqrt{p+1}}. \end{aligned}$$

In the last line we used forward-inverse trigonometric identities again. We conclude that

$$\mathbb{E}Y^2 = \frac{p+1}{2p^2} - \frac{3\sqrt{p+1}}{2p^{5/2}} \operatorname{arcsinh}(\sqrt{p}) + \frac{1}{p^2} = \frac{a^2+2}{2(a^2-1)^2} - \frac{3a \operatorname{arcsinh}(\sqrt{a^2-1})}{2(a^2-1)^{5/2}}.$$

A.4 Variance of Y and s_{flo} – oblate ellipsoid

Combining the previous results we get

$$\begin{aligned} \operatorname{var}(s_{\text{flo}}) &= 4\pi^2 (a^2-1)^2 \operatorname{var}(Y) \\ &= 2\pi^2 a^2 + \frac{2\pi^2 a \operatorname{arcsinh}(\sqrt{a^2-1})}{\sqrt{a^2-1}} - \frac{4\pi^2 a^2 \operatorname{arcsinh}^2(\sqrt{a^2-1})}{a^2-1}. \end{aligned}$$

The surface area of an oblate ellipsoid with minor semiaxis length equal to 1 and major semiaxis length equal to $a > 1$ is given by the formula

$$s = 2\pi a^2 \left[1 + \frac{1}{a\sqrt{a^2-1}} \operatorname{arctanh} \left(\frac{\sqrt{a^2-1}}{a} \right) \right].$$

Using l'Hospital's rule it can be shown that the expression in square brackets has a limit 1 as $a \rightarrow \infty$. Recalling that $\operatorname{arcsinh}(x) = \log(x + \sqrt{x^2 + 1})$, it is easily seen that

$$\lim_{a \rightarrow \infty} \frac{\operatorname{var}(s_{\text{flo}})}{s^2} = 0,$$

i.e. that the relative variance of s_{flo} converges to 0 as $a \rightarrow \infty$.

The dependence of the relative variance of s_{flo} for oblate ellipsoids on the major semiaxis length a is shown in Figure 9. The shape of the curve is consistent with our previous empirical findings.

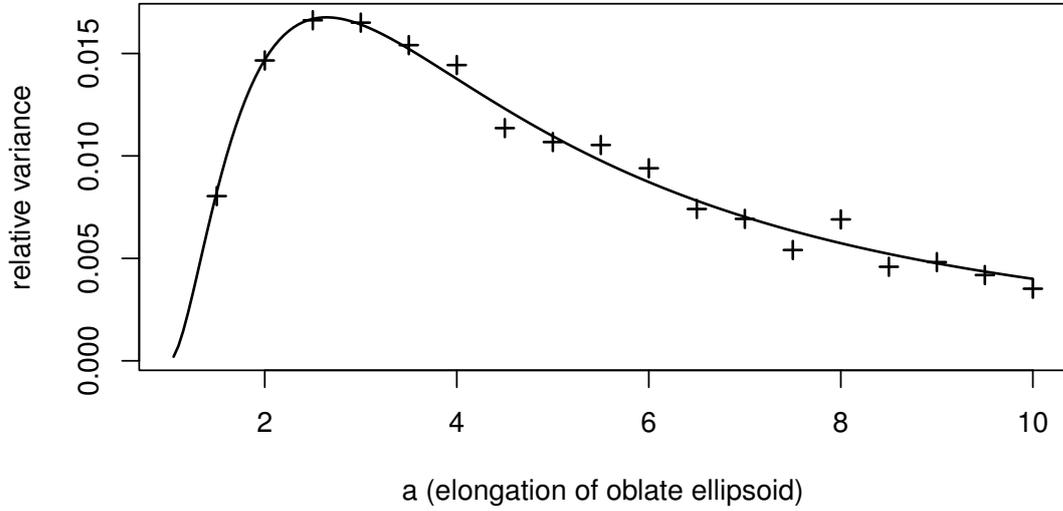


Figure 9: Oblate ellipsoids: dependence of the relative variance of s_{flo} on the major semiaxis length a of the ellipsoid. Solid line – theoretical values, crosses – empirical estimates obtained by using 10 000 section planes through the centre of the ellipsoid. New, independent section planes were used for each ellipsoid.

A.5 Calculating $\mathbb{E}Y$ – prolate ellipsoid

For ease of exposition we denote $p = c^2 - 1$. We assume that $c > 1$ and thus $p > 0$. Then

$$\mathbb{E}Y = \int_0^1 y f_Y^P(y) dy = \frac{p+1}{2} \int_0^1 \frac{y}{\sqrt{1-y}} (1+py)^{-3/2} dy.$$

By substituting $x = \sqrt{1-y}$ we get

$$\int_0^1 \frac{y}{\sqrt{1-y}} (1+py)^{-3/2} dy = 2 \int_0^1 \frac{1-x^2}{(-px^2+p+1)^{3/2}} dx.$$

Then, using the second Euler substitution $\sqrt{-px^2 + p + 1} = xt - \sqrt{p + 1}$, we get

$$2 \int_0^1 \frac{1 - x^2}{(-px^2 + p + 1)^{3/2}} dx = \frac{4}{p + 1} \int_{1+\sqrt{p+1}}^{\infty} \frac{(t^2 + p)^2 - 4(p + 1)t^2}{(t^2 + p)(t^2 - p)^2} dt.$$

Since the integrand is a rational function of t we can use partial fraction decomposition and integrate the individual terms to get the result

$$\frac{4}{p + 1} \int_{1+\sqrt{p+1}}^{\infty} \frac{(t^2 + p)^2 - 4(p + 1)t^2}{(t^2 + p)(t^2 - p)^2} dt = \frac{2 \arctan(\sqrt{p})}{p^{3/2}} - \frac{2}{p(p + 1)}.$$

The simple form of the result is due to specific identities for inverse trigonometric functions.

Finally,

$$\mathbb{E}Y = \frac{p + 1}{p^{3/2}} \arctan(\sqrt{p}) - \frac{1}{p} = \frac{c^2}{(c^2 - 1)^{3/2}} \arctan(\sqrt{c^2 - 1}) - \frac{1}{c^2 - 1}.$$

A.6 Calculating $\mathbb{E}Y^2$ – prolate ellipsoid

Again we denote $p = c^2 - 1$ and assume that $c > 1$ and $p > 0$. Then

$$\mathbb{E}Y^2 = \int_0^1 y^2 f_Y^P(y) dy = \frac{p + 1}{2} \int_0^1 \frac{y^2}{\sqrt{1 - y}} (1 + py)^{-3/2} dy.$$

By substituting $x = \sqrt{1 - y}$ we get

$$\int_0^1 \frac{y^2}{\sqrt{1 - y}} (1 + py)^{-3/2} dy = 2 \int_0^1 \frac{(1 - x^2)^2}{(-px^2 + p + 1)^{3/2}} dx.$$

Then, using the second Euler substitution $\sqrt{-px^2 + p + 1} = xt - \sqrt{p + 1}$, we get

$$2 \int_0^1 \frac{(1 - x^2)^2}{(-px^2 + p + 1)^{3/2}} dx = \frac{4}{p + 1} \int_{1+\sqrt{p+1}}^{\infty} \frac{((t^2 + p)^2 - 4(p + 1)t^2)^2}{(t^2 + p)^3(t^2 - p)^2} dt.$$

Since the integrand is a rational function of t we can use partial fraction decomposition and integrate the individual terms to get the result

$$\frac{4}{p + 1} \int_{1+\sqrt{p+1}}^{\infty} \frac{((t^2 + p)^2 - 4(p + 1)t^2)^2}{(t^2 + p)^3(t^2 - p)^2} dt = \frac{(p - 3) \arctan(\sqrt{p})}{p^{5/2}} + \frac{p + 3}{p^2(p + 1)}.$$

The simple form of the result is due to specific identities for inverse trigonometric functions.

We conclude that

$$\begin{aligned} \mathbb{E}Y^2 &= \frac{(p + 1)(p - 3)}{2p^{5/2}} \arctan(\sqrt{p}) + \frac{p + 3}{2p^2} \\ &= \frac{c^2(c^2 - 4)}{2(c^2 - 1)^{5/2}} \arctan(\sqrt{c^2 - 1}) + \frac{c^2 + 2}{2(c^2 - 1)^2}. \end{aligned}$$

A.7 Variance of Y and s_{flo} – prolate ellipsoid

Combining the previous results and keeping for the moment the simplified notation $p = c^2 - 1$ we get

$$\begin{aligned}\text{var}(s_{\text{flo}}) &= 4\pi^2 p^2 \text{var}(Y) \\ &= \frac{2\pi^2(p+1)}{p} (p + \sqrt{p}(p+1) \arctan(\sqrt{p}) - 2(p+1) \arctan(\sqrt{p})^2).\end{aligned}$$

The surface area of a prolate ellipsoid with minor semiaxis length equal to 1 and major semiaxis length equal to $c > 1$ is given by the formula

$$s = 2\pi \left[1 + \frac{c^2}{\sqrt{c^2-1}} \arcsin\left(\frac{\sqrt{c^2-1}}{c}\right) \right] = 2\pi \left[1 + \frac{p+1}{\sqrt{p}} \arcsin\left(\frac{\sqrt{p}}{\sqrt{p+1}}\right) \right].$$

It can be shown that

$$\lim_{c \rightarrow \infty} \frac{\text{var}(s_{\text{flo}})}{s^2} = \infty,$$

i.e. the relative variance of s_{flo} converges to infinity as $c \rightarrow \infty$. The increase of the relative variance is asymptotically linear in c as

$$\lim_{c \rightarrow \infty} \frac{1}{c} \frac{\text{var}(s_{\text{flo}})}{s^2} = \frac{1}{\pi}.$$

The dependence of the relative variance of s_{flo} for prolate ellipsoids on the major semiaxis length c is shown in Figure 10, together with empirical estimates of the relative variance.

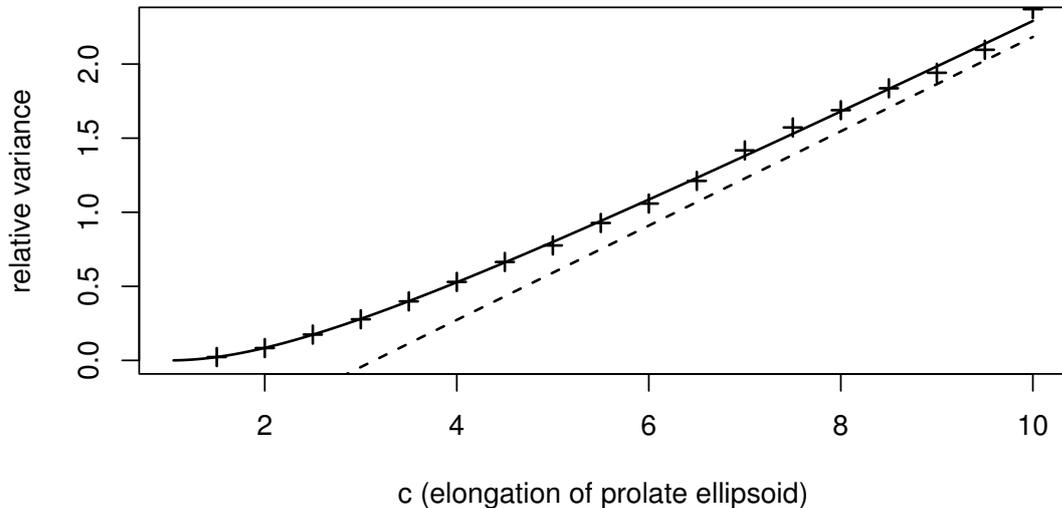


Figure 10: Prolate ellipsoids: dependence of the relative variance of s_{fl_0} on the major semiaxis length c of the ellipsoid. Solid line – theoretical values, crosses – empirical estimates obtained by using 10 000 section planes through the centre of the ellipsoid. New, independent section planes were used for each ellipsoid. Dashed line – the asymptote for the relative variance of s_{fl_0} .

B Optimality of the pivotal estimator – ellipsoids

We show that for an arbitrary (triaxial) ellipsoid with a fixed reference point in its interior the pivotal estimator $s_{piv,4}$ and the flower estimator s_{fl_0} are identical. Hence $s_{piv,4}$ is optimal for estimation of the surface area in the sense that it captures all the information available in each section with the minimal effort (it requires measurement of the support function in only four directions).

Any planar section of the ellipsoid through the reference point O is an ellipse – we denote it Y and its semiaxes lengths a and b . Let $z = [x_0, y_0]$ be the vector of coordinates of the centre of the ellipse in the section plane, with the reference point O being the origin of the coordinate system. According to Noumeir (2000), if $z = [0, 0]$, the support function of the ellipse Y can be expressed as

$$h_Y(\omega) = \sqrt{a^2 \cos^2 \omega + b^2 \sin^2 \omega}, \quad (\text{B.1})$$

where ω is the angle with respect to the first coordinate axis, chosen such that it coincides with the major axis of the ellipse. Note that due to the specific form of the support function (B.1) $h_Y(\omega + \pi) = h_Y(\omega)$. Let u_ω denote the unit vector making the angle ω with the first coordinate axis. Then the support function of the translated set $Y + z, z \in \mathbb{R}^2$, takes the form

$$h_{Y+z}(\omega) = h_Y(\omega) + \langle z, u_\omega \rangle, z \in \mathbb{R}^2, \quad (\text{B.2})$$

where $\langle x, y \rangle$ denotes the inner product of the vectors x and y .

First we consider the case of $z = [0, 0]$, i.e. the reference point is in the centre of the ellipse (this would be the case e.g. if the original reference point was in the centre of the ellipsoid). We show that the estimator

$$\tilde{s}_{\text{piv},2} = 2\pi \left(h_Y^2(\omega) + h_Y^2\left(\omega + \frac{\pi}{2}\right) \right) \quad (\text{B.3})$$

is optimal, no matter the value of ω , since $\tilde{s}_{\text{piv},2} = s_{\text{flo}} = 2\pi(a^2 + b^2)$, see (2.4). Note that the estimator $\tilde{s}_{\text{piv},2}$ slightly differs from the estimator $s_{\text{piv},2}$ defined in Section 2 as the difference in the arguments of h_Y is only $\frac{\pi}{2}$ and not π . The reasoning goes as follows:

$$h_Y^2(\omega) = a^2 \cos^2 \omega + b^2 \sin^2 \omega, \quad (\text{B.4})$$

$$\begin{aligned} h_Y^2\left(\omega + \frac{\pi}{2}\right) &= a^2 \cos^2\left(\omega + \frac{\pi}{2}\right) + b^2 \sin^2\left(\omega + \frac{\pi}{2}\right) \\ &= a^2 \sin^2 \omega + b^2 \cos^2 \omega, \end{aligned} \quad (\text{B.5})$$

$$h_Y^2(\omega) + h_Y^2\left(\omega + \frac{\pi}{2}\right) = a^2 + b^2. \quad (\text{B.6})$$

Hence, $\tilde{s}_{\text{piv},2} = 2\pi(a^2 + b^2) = s_{\text{flo}}$.

In the general case where $z = [x_0, y_0]$ the estimator $s_{\text{piv},4}$ is optimal in the sense that $s_{\text{piv},4} = s_{\text{flo}}$. The estimator can be expressed as

$$s_{\text{piv},4} = \pi \sum_{n=0}^3 h_{Y+z}^2\left(\omega + n \frac{2\pi}{N}\right), \quad (\text{B.7})$$

see (2.8). Since $h_{Y+z}^2(\omega) = h_Y^2(\omega) + \langle z, u_\omega \rangle^2 + 2h_Y(\omega)\langle z, u_\omega \rangle$, it is convenient to decompose $s_{\text{piv},4}$ into three parts: $s_{\text{piv},4} = \pi(A + B + C)$, where

$$A = h_Y^2(\omega) + h_Y^2\left(\omega + \frac{\pi}{2}\right) + h_Y^2(\omega + \pi) + h_Y^2\left((\omega + \pi) + \frac{\pi}{2}\right) = 2(a^2 + b^2), \quad (\text{B.8})$$

$$B = \sum_{n=0}^3 \langle z, u_{\omega+n\frac{\pi}{2}} \rangle^2 = 2|z|^2, \quad (\text{B.9})$$

$$C = \sum_{n=0}^3 h_Y\left(\omega + n\frac{\pi}{2}\right) \langle z, u_{\omega+n\frac{\pi}{2}} \rangle = 0. \quad (\text{B.10})$$

The values of B and C can be easily obtained using $\langle z, u_\omega \rangle = x_0 \cos \omega + y_0 \sin \omega$, $h_Y(\omega + \pi) = h_Y(\omega)$ and simple trigonometric identities. Here $|z|^2 = x_0^2 + y_0^2$ denotes the squared distance from the reference point to the centre of the ellipse. As a result, $s_{\text{piv},4} = 2\pi(a^2 + b^2 + |z|^2) = s_{\text{flo}}$.