

STEREOLOGICAL ANALYSIS  
OF  
SPATIAL STRUCTURES



PHD DISSERTATION  
LINDA VADGÅRD HANSEN  
lvh@imf.au.dk

SUPERVISOR: EVA B. VEDEL JENSEN

DEPARTMENT OF  
MATHEMATICAL SCIENCES  
AARHUS UNIVERSITET

JULY 31, 2011



# Preface

Initially, stereology relied on manual observations in two-dimensional sections. The stereological techniques have continuously evolved over time. The very recent improvements in computerised image analysis allow for further development of these methods and the progress to new methods that incorporate readily available information obtained from e.g. a grayscale digital image of the two-dimensional section under consideration.

This PhD thesis is a contribution to the field of stereology and stochastic geometry. The focus is on the improvement of local stereological methods by incorporating auxiliary information within a sound statistical framework. Several issues relating to the precision of stereological estimators are dealt with. Furthermore, a flexible modelling framework for three-dimensional objects is developed under which distributional properties of local stereological estimators are derived. The thesis, submitted to the Faculty of Science, Aarhus university, consists of an introduction accompanied by four independently written papers. The introduction provides a short presentation to selected relevant parts of stereology and stochastic geometry.

My studies have been conducted under the careful supervision of Eva B. Vedel Jensen at the Department of Mathematical Sciences, Aarhus University, in collaboration with Jens R. Nyengaard at the Stereology and EM Research Laboratory, Aarhus University.

First, I would like to acknowledge Eva B. Vedel Jensen for her excellent and optimistic supervision and invaluable generous support on both the professional and personal level. Her deep dedication and love for research has been a great inspiration to me.

My co-supervisor Jens R. Nyengaard is acknowledged for his deep insight in biology and his commitment to combine the worlds of theorists and practitioners of statistics which has served as a great inspiration to me. At the Stereology and EM Research Laboratory, I also wish to thank Johnnie B. Andersen for introducing me to some sampling procedures concerning a joint project that provided me with a deeper insight into the problems that might occur when implementing theory in practice.

My PhD studies were carried out during Aug. 2007 – July 2011, with the period Apr. 2010 – July 2010 spent at Heidelberg University, where I visited Tilmann Gneiting, to whom I also owe many thanks – both for professional collaboration and great hospitality.

Furthermore, I would like to thank all my colleagues at the Department of Mathematical Sciences for providing an excellent working environment. In that regard, I would like to give a few people special acknowledgements. Concerning the proof of Theorem 1.1 in the Introduction of this thesis, a special thank you goes to Jens Ledet Jensen. Moreover, I especially wish thank Anders Rønn-Nielsen and Emil Hedevang with whom I have had many fruitful discussions during our time as office colleagues. I would also like to send a special acknowledgement to Thordis L. Thorarinsdottir who has been both a good colleague, collaborator and friend to me both during her time in Aarhus and my stay abroad in Heidelberg where she offered her great hospitality. Also I wish to extend my gratitude to Lars Madsen, who has been incredibly helpful with the final layout of this thesis.

Finally, I extend my deepest gratitude to my family, my parents Edith & Linding and my sister Anne. They have been very supportive and have always believed in me both during my PhD and the period leading up till this point.

Aarhus, Juli 31, 2011.

Linda Vadgård Hansen



# Summary – English

This thesis deals with several issues relating to the precision of stereological estimators. Furthermore, a flexible modelling framework for three-dimensional objects is developed under which distributional properties of local stereological estimators can be derived.

Image-based empirical importance sampling is considered for estimation of the intensity of point processes. Here, computerised image analysis is used to redirect the sampling. Statistical tools for assessing the efficiency of image-based empirical importance sampling are obtained. It is shown that this type of importance sampling may provide a significant increase in the precision of intensities estimators for inhomogeneous point processes. Furthermore, optimal model-based estimators of intensities are constructed.

Mixing of two well-known local stereological volume estimators are studied as a means to obtain a new local stereological volume estimator that increases the efficiency of the stereological procedure. The starting point is the nucleator, a well-established manual stereological method for estimating mean cell volume from observations on random cell transects through reference points of the cells. We consider an automated version of the nucleator using automatic segmentation of the boundaries of the cell transects. If the segmentation is judged satisfactory according to some criteria, an estimate of the cell volume is calculated automatically on the basis of the whole cell transect; otherwise the classical nucleator is used. The statistical properties of this new estimator are considered. Formulae for the bias and mean square error are derived and procedures for estimating bias and mean square error from a pilot study are provided.

A flexible framework for modelling and simulating three-dimensional star-shaped random sets is considered. The idea is to identify a star-shaped random sets with its radial function and then model the radial function as a random field over the unit sphere. The random field is given by a kernel smoothing of a Lévy basis. Under such a model we demonstrate that one can obtain sets with irregular surfaces, i.e. we can obtain sets where the surfaces have Hausdorff (fractal) dimension greater than two. We present a simulation algorithm that generates this type of star-shaped random sets.

If the model is generalised slightly, it becomes tractable for making statements on the distribution of the volume of the random sets and for local stereological volume estimators. What is needed is to make a power transformation of the

original model for the radial function. Then, the volume distribution of the random set is known and statements on the distribution of the local stereological volume estimators can be obtained.

The present thesis contains four scientific papers. The framework and main results of each paper are presented in the introduction.

# Summary – Danish

Denne afhandling behandler spørgsmål vedrørende præcisionen af stereologiske estimatorer. Endvidere betragtes en fleksibel ramme for modellering af tredimensionelle objekter under hvilken fordelingssegenskaber af lokale stereologiske estimatorer kan udledes.

Billedbaseret empirisk importance sampling studeres med henblik på estimering af intensiteten af punkt processer. Her udnyttes billedanalyse til at om dirigere samplingen. Endvidere præsenteres statistiske redskaber til at vurdere effektiviteten af billedbaseret empiriske importance sampling. Det vises, at denne form for importance sampling kan øge præcisionen af intensitets-estimatorer for inhomogene punkt processer signifikant. Desuden konstrueres optimale modelbaserede estimatorer for intensiteten af punktprocesser.

Kombinationen af to velkendte lokale stereologiske volumen estimatorer betragtes som en metode til at opnå en ny lokal stereologisk volumen estimator, der øger effektiviteten af den stereologiske procedure. Udgangspunktet er nucleator estimatoren, en veletableret manuel stereologisk metode til estimation af cellevolumen ud fra observationer på stokastiske plane snit gennem referencepunkter knyttet til cellerne. Vi betragter en automatiseret version af nucleatoren, der udnytter automatisk segmentering af randen for hvert af cellesnittene. Hvis segmentering vurderes tilfredsstillende i henhold til nogle kriterier, beregnes estimatet af cellevolumenet automatisk på grundlag af hele cellesnittet; ellers benyttes den klassiske nucleator. Statistiske egenskaber for denne nye estimator betragtes. Formler for bias og mean square error udledes og procedurer til vurdering af bias og mean square error fra en pilotstudie præsenteres.

En fleksibel ramme for modellering og simulering af tredimensionelle stjerneformede stokastiske mængder betragtes. Idéen er at identificere en stjerneformet stokastisk mængde med dens radialfunktion og derefter modellere radialfunktionen som et stokastisk felt over enhedskuglen. Det stokastiske felt er givet ved en kerne udglatning af en Lévy base. Under en sådan model kan opnås mængder med irregulære overflader, dvs. vi kan opnå mængder, hvor overfladerne har Hausdorff (fraktal) dimension større end to. Endvidere, præsenterer vi en simulationsalgoritme, der genererer denne type stjerneformede stokastiske mængder.

Generaliseres modellen en smule, bliver det muligt at fremsætte udsagn vedrørende fordelingen af den stokastiske mængde samt for lokale stereologiske volumen estimatorer. Det, der kræves, er en potenstransformation af den

oprindelige model for radialfunktionen. Herefter er volumenfordelingen af den stokastiske mængde kendt og der kan gives udsagn omkring fordelingen af de lokale stereologiske volumen estimatorer.

Denne afhandling indeholder fire videnskabelige artikler. Grundlaget og de vigtigste resultater for hver af artiklerne præsenteres i introduktionen.

# Contents

Preface . . . . .	i
Summary – English . . . . .	iii
Summary – Danish . . . . .	v
Accompanying Papers . . . . .	xi

<b>Introduction . . . . .</b>	<b>1</b>
1.1 Stereology . . . . .	1
1.2 Non-uniform sampling (Paper A) . . . . .	2
1.2.1 The nucleator and non-uniform sampling . . . . .	4
1.3 Semi-automatic procedures (Paper B) . . . . .	6
1.4 Particle models (Paper C and Paper D) . . . . .	8
1.4.1 Lévy particles (Paper C and Paper D) . . . . .	9
1.4.2 Stereology and Lévy particles (Paper D) . . . . .	11
1.4.3 Surface behaviour of Lévy particles (Paper C) . . . . .	13
1.5 Surface behaviour of a Lévy field on $\mathbb{R}^2$ . . . . .	14
Bibliography . . . . .	16

## *Paper A*

<b>Image-based empirical importance sampling: an efficient way of estimating intensities . . . . .</b>	<b>21</b>
A.1 Introduction . . . . .	23
A.2 The set-up . . . . .	24
A.3 A simple condition for gain in efficiency . . . . .	26
A.4 Efficiency for homogeneous point processes . . . . .	28
A.5 Design-based inference . . . . .	30
A.6 Model-based inference . . . . .	32
A.7 Concluding remarks . . . . .	35
A.8 Acknowledgements . . . . .	35
Bibliography . . . . .	36
Appendix 1 . . . . .	36
Appendix 2 . . . . .	38
Appendix 3 . . . . .	38

***Paper B***

<b>The semi-automatic nucleator . . . . .</b>	<b>41</b>
B.1 Introduction . . . . .	43
B.2 Theoretical background . . . . .	44
B.2.1 The estimators . . . . .	45
B.2.1.1 The classical nucleator . . . . .	45
B.2.1.2 The integrated nucleator . . . . .	46
B.2.1.3 The automatic nucleator . . . . .	46
B.2.1.4 The semi-automatic nucleator . . . . .	47
B.2.2 Variance and mean square error relations . . . . .	47
B.2.3 Estimation of bias and MSE . . . . .	48
B.2.4 Discrimination between estimators . . . . .	49
B.3 A comparative study . . . . .	50
B.3.1 Materials and preparation methods . . . . .	50
B.3.2 Sampling and segmentation . . . . .	50
B.3.3 Constructing the true cell transect . . . . .	50
B.3.4 Comparison of estimators . . . . .	53
B.4 Discussion . . . . .	55
B.5 Acknowledgements . . . . .	57
Bibliography . . . . .	57
Appendix . . . . .	58

***Paper C***

<b>Lévy particles: Modelling and simulating star-shaped random sets . . . . .</b>	<b>61</b>
C.1 Introduction . . . . .	63
C.2 Preliminaries . . . . .	65
C.3 Hausdorff dimension . . . . .	66
C.4 Isotropic kernels . . . . .	68
C.4.1 Von Mises–Fisher kernel . . . . .	68
C.4.2 Uniform kernel . . . . .	70
C.4.3 Power kernel . . . . .	71
C.5 Examples . . . . .	73
C.5.1 Simulation algorithm . . . . .	73
C.5.2 Celestial bodies . . . . .	74
C.5.3 Planar Lévy particles . . . . .	76
C.6 Discussion . . . . .	78
Bibliography . . . . .	80

***Paper D***

<b>Local stereology for Lévy particles . . . . .</b>	<b>83</b>
D.1 Introduction . . . . .	85
D.2 The model and theoretical background . . . . .	86

D.3	Isotropic kernels . . . . .	87
D.3.1	The antipodal uniform kernel . . . . .	88
D.3.2	The antipodal exponential kernel . . . . .	89
D.4	Volume prediction . . . . .	91
D.5	Volume prediction under the power 3 Lévy model . . . . .	93
D.6	Discussion . . . . .	96
	Bibliography . . . . .	98
	Appendix . . . . .	99





# Accompanying Papers

<b>Paper A</b> . . . . .	21
Hansen, L. V., Kiderlen, M. and Jensen, E. B. V. (2011): Image-based empirical importance sampling: an efficient way of estimating intensities. <i>Scandinavian Journal of Statistics</i> , in press.	
<b>Paper B</b> . . . . .	41
Hansen, L. V., Nyengaard, J. R., Andersen, J. B. and Jensen, E. B. V. (2011): The semi-automatic nucleator. <i>Journal of Microscopy</i> , 242: 206–215.	
<b>Paper C</b> . . . . .	61
Hansen, L. V., Thorarinsdottir, T. L. and Gneiting, T. (2011): Lévy particles: Modelling and simulating star-shaped compact sets, Manuscript.	
<b>Paper D</b> . . . . .	83
Hansen, L. V. and Jensen, E. B. V. (2011): Local stereology for Lévy particles, Manuscript.	



# Introduction

## 1.1 Stereology

Originally, stereology was defined as *the spatial interpretation of sections* (Weibel, 1979). This early description captures the essence of this interdisciplinary field very well as stereological methods are designed to provide statistically sound information on three-dimensional objects from observations made in two dimensional sections. The main practical purpose of stereology is to extract quantitative information of three-dimensional objects from microscopic images. A main advantage is that three-dimensional reconstruction of the object is not necessary in order to do so which makes stereology a relatively cheap technique and hence its practical appeal. Stereological methods can also be developed for observations on a one-dimensional transect, a two-dimensional projection or a three-dimensional volume image. The procedures are appreciated within many fields such as life sciences, geology, soil science, material science and engineering.

Though, the introduction of the term stereology and the creation of The International Society of Stereology (ISS) in 1963 are both rather recent events the history of stereology dates back much further. Historical references often mentioned are *Cavalieri's principle* dating back to F. B. Cavalieri (1598–1647), cf. e.g. Baddeley and Jensen (2005), and *Buffon's needle problem* stated by the count of Buffon, G. L. Leclerc (1701–1788), in 1777<sup>1</sup>. One of the first reports of stereological techniques used in combination with microscopes is given by the young French mineralogist A.E. Delesse (1817–1881) in 1847 (Delesse, 1847, 1848). Modern stereology can be considered as spatial sampling theory (Baddeley, 1993) and it has strong analogies within survey sampling (Cochran, 1977; Thompson, 1992).

A solid mathematical foundation for stereology was laid down only some decades ago in a series of papers from the nineteen seventies by Roger Miles and Pamela Davy (Davy and Miles, 1977; Miles and Davy, 1976, 1977; Miles, 1978a,b) where they also emphasised the connection to survey sampling. The exposition of stereology made by Weibel (1979, 1980) celebrated the sampling approach of Miles and Davy and presented sampling principles for many of the standard stereological techniques. Yet new stereological principles keep occurring

---

<sup>1</sup>in *Histoire Naturelle Générale et Particulière, Servant de Suite à l'Histoire Naturelle de l'Homme, Supplements*, Forth volume, XXIII, 1777, p 95–100

and the field is in constant development. For review papers on stereology see e.g. Baddeley (1993); Jensen et al. (1985); Stoyan (1990).

A fundamental property of the stereological procedures is that no model assumptions on the object of interest are necessary (here disregarding assumptions of compactness and so on) which makes techniques relevant for practitioners in e.g. biomedicine where properties such as spatial homogeneity may not be satisfied.

With a design based approach where the probe (e.g. a line, a line system, a point grid, a plane etc.) providing the observations is randomised appropriately one can obtain unbiased estimators of e.g. volume, surface area and length. In the eyes of a practitioner unbiasedness is an important property of an estimator. However, a statistician does not settle with unbiasedness, the precision of the estimator is an equally important issue.

The present thesis contains four scientific papers, all dealing with issues relating to the precision of the stereological estimators. In *Paper A*, image-based empirical importance sampling is introduced as an efficient tool for improving the precision of estimators of intensities. *Paper B* describes a new local stereological estimator of volume, the semi-automatic nucleator, with higher precision than the traditional nucleator estimator. The new estimator uses computerised image analysis. In *Paper C*, a flexible model for star-shaped random sets is developed where the radial function is modelled by a kernel smoothing of a Lévy basis. For the resulting random sets, called Lévy particles, it is possible to derive formulae for the precision of a number of local stereological estimators, see *Paper D*.

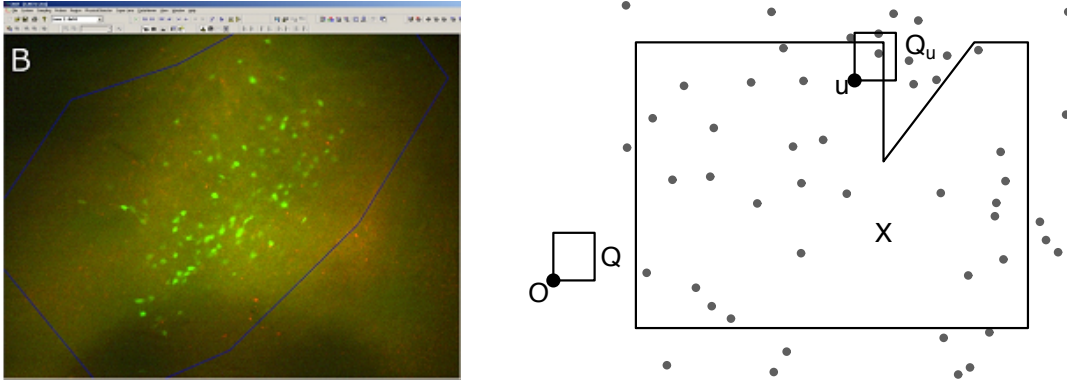
The main results of the four papers are described below.

## 1.2 Non-uniform sampling (Paper A)

In *Paper A* we assume that a realisation of a point process  $\Phi$  is available for observation in a bounded subset  $X$  of  $\mathbb{R}^2$  of area  $A(X)$ . The aim is to predict the total number  $N(\Phi \cap X)$  of points in  $X$  or equivalently the area weighted total number  $N_A = N(\Phi \cap X)/A(X)$ . In the applications we have in mind, the points of  $\Phi$  represent the positions of the objects in a digital or analog image. In the example of Figure 1.1 (left),  $X$  corresponds to the region delineated by blue at low magnification and  $\Phi$  represents the green cells which are of interest. Counting of the cells has to be performed at high magnification and it is impracticable to count every single cell falling inside  $X$ .

Hence, the aim is to obtain  $N(\Phi \cap X)$  from observation in a randomly placed window  $Q_U = Q + U$ , hitting  $X$ . Here,  $Q \subset \mathbb{R}^2$  is bounded and is assumed to contain the origin  $O$ . The position of the window is determined by the random vector  $U \in \mathbb{R}^2$ . For an illustration of the framework see Figure 1.1 (right).

Say, we are allowed to base our prediction of  $N_A$  on the observations from  $N$  windows  $Q_{U_1}, Q_{U_2}, \dots, Q_{U_N}$  then, dependent on the distribution of the points  $\Phi$ , different strategies might preferred. First, the variables  $U_1, U_2, \dots, U_N$  can be chosen independent and uniformly within the set  $\bar{X} := \{u \in \mathbb{R}^2 : X \cap Q_u \neq \emptyset\}$ .



**Figure 1.1:** Left. In this microscopy image the green cells constitute a realisation of  $\Phi$  and the region  $X$  is delineated in blue. Right. The red dots show a realisation of the point process  $\Phi$  which is available for observation within the set  $X$  delineated by the blue line. The window  $Q$  and the by  $u$  translated version thereof  $Q_u$  are shown in black.

This type is known as simple random sampling. Alternatively,  $U_1, U_2, \dots, U_N$  can be chosen in an equidistant systematic manor with uniform placement. This approach is known as systematic uniform random sampling.

In the applications to microscopy, the points of  $\Phi$  typically represent the positions of the objects in a digital or analog image. As mentioned earlier, examination of a window and counting the points can only be performed at high magnification of the microscopic image and is therefore time consuming. Hence, it is crucial to obtain as much information as possible from the time spent. Even though systematic uniform random sampling introduce a significant increase in efficiency compared to simple random sampling many windows might contain very little information if the points are inhomogeneously distributed.

Recently non-uniform sampling has been introduced to stereology in the biomedical literature under the name of *the proportionator*. This technique utilise computerised analysis of microscopy images to redirect the sampling of an inhomogeneous cell population (Gardi et al., 2008, 2007). The proportionator proves to be of utmost relevance when dealing with inhomogeneous cell population and it is found to outperform systematic uniform random sampling.

In *Paper A*, a solid mathematical foundation for the proportionator is provided by relating the proportionator to the well understood statistical concept of importance sampling. The idea underlying importance sampling is to sample from an alternative distribution than the one from which we want to perform inference. A good alternative distribution is one that redirect the sampling to the part of the state space which contributes the most to the parameter of interest (Asmussen and Glynn, 2007). By redirecting the sampling according to computerised analysis of microscopy images this is exactly what the proportionator technique set out to exploit. We notice that, also the proportionator technique has an analogy within survey sampling known as the Horvitz–Thompson estimator (Horvitz and Thompson, 1952).

In applications, it might not always be straightforward how to choose the alternative sampling distribution. If  $p$  is the traditional sampling distribution over the region  $X$ , we study in *Paper A* the performance of a sampling distribution of the form

$$\frac{Z(\cdot)}{\int Z} p(\cdot | Z)$$

Here,  $Z$  denotes a random field (available on  $\bar{X}$ ) and  $\int Z$  denotes integration of  $Z$  over its entire domain with respect to Lebesgue measure.

In applications,  $Z$  may be readily available for instance via colour proportions in a grayscale image (Gardi et al., 2007, Fig. 3, Fig. 4 and Fig. 5). The gain in efficiency obtained depend of course on the information contained in the grayscale image. If the redirection of the sampling is done appropriately a gain in efficiency without increasing the workload can be expected by using the proportionator. Based on three biological examples Gardi et al. (2007) report an 8–25 fold increase in efficiency and these findings are supported by substantial simulation studies (Gardi et al., 2008).

In addition to laying out a sound mathematical foundation for this type of empirical importance sampling *Paper A* provide statistical tools for assessing its efficiency. Moreover, optimal model based estimators of intensities are constructed under the assumption of a proportional regression. The proposed optimal estimator and a corresponding estimator obtained from weighted random sampling is studied in the light of the data example of Gardi et al. (2007, Fig. 3). In that concrete example it is found that given the random field  $Z$  and the sampling design, the variance of the traditional proportionator estimator is about three times the variance of the optimal estimator.

For the proportionator method some additional readily available information is required in order to establish the non-uniform sampling distribution. This information can be available as a grayscale image of the object under study. Due to the immense progress of computerised image analysis such auxiliary information can be easily accessible without increasing the workload. The positive results on non-uniform sampling reported above encouraged the implementation to other stereological methods.

### 1.2.1 The nucleator and non-uniform sampling

A number of stereological volume estimators are available in the literature, cf. e.g. Jensen (1998, Table 7.1). All of these estimators are unbiased and do not require any specific shape assumptions. One, *the nucleator* (Gundersen, 1988) has enjoyed repeated use in application. For recent applications to neuroscience and biology, see Abrahão et al. (2009) and Melo et al. (2009). Below and in *Paper B*, we set out to improve the efficiency of the nucleator by incorporating auxiliary information.

The nucleator belongs to a branch of stereology, called local stereology, where the random section plane is restricted to pass through a fixed reference point

(Jensen, 1998). In biological applications, such a reference point could typically be the nucleus or nucleolus of a cell. If no obvious reference point is visible, then the centre of mass might be used as the reference point.

Let  $Y$  denote a three-dimensional particle containing a fixed reference point  $o$ . Following the notation of *Paper B* we let  $L_2$  denote an isotropic plane through  $o$ . The normal vector of such a random plane is uniformly distributed on a unit half-sphere. Furthermore,  $L_1$  is a line in  $L_2$  through  $o$ . Notice that,  $L_1$  can be uniquely determined by the angle  $\theta \in [0, \pi)$  it generates with a fixed axis within  $L_2$ . If needed, this specific line is denoted  $L_1(\theta)$ .

Then the nucleator estimator of  $V(Y)$  is given by

$$m_{\text{cl}_1}(Y \cap L_1) = 2\pi \int_{Y \cap L_1} d(y, o)^2 dy,$$

where  $L_1$  is uniformly oriented, that is  $\theta$  is uniformly distributed on  $[0, \pi)$ . Here,  $d(y, x)$  denotes the Euclidean distance between the points  $x, y \in \mathbb{R}^3$ . In the case where  $Y \cap L_1$  forms a single line segment  $[y_-, y_+]$  the estimator takes the simple form

$$m_{\text{cl}_1}(Y \cap L_1) = \frac{2\pi}{3} (d(y_-, o)^3 + d(y_+, o)^3).$$

The nucleator is an unbiased estimator of  $V(Y)$ , i.e.  $E(m_{\text{cl}_1}(Y \cap L_1)) = V(Y)$ .

To increase the precision of the nucleator estimator it is found useful to base the estimator is based on a systematic set of lines within  $L_2$ . In applications, this systematic version of the nucleator is frequently used and systems of two or four lines are often sufficient. We indicate the number of lines on which the nucleator estimator is based by the lower index, e.g.  $m_{\text{cl}_2}$  is the notation used for a nucleator based on two systematic lines.

Setting out to combine non-uniform sampling and the nucleator principle we need some auxiliary information to redirect the sampling. In the application to microscopy, this additional information can be available from segmentation of the microscopy image. Let  $\tilde{Y}_2$  denote the segmented cell transect  $Y \cap L_2$ . If the segmentation procedure is without error,  $\tilde{Y}_2 = Y \cap L_2$ ; otherwise  $\tilde{Y}_2$  may be regarded as an estimate of  $Y \cap L_2$ . The idea is then to reduce the within section variance of the classical nucleator by changing the orientation distribution of  $\theta$  from uniform to a weighted distribution by using the information from  $\tilde{Y}_2$ . More precisely, we considered lines  $L_1 = L_1(\theta)$  with the following orientation density distribution  $q$ , depending on the shape of  $\tilde{Y}_2$

$$q(\theta) \propto \int_{\tilde{Y}_2 \cap L_1} d(y, o)^2 dy, \quad \theta \in [0, \pi).$$

The new estimator is then given by

$$m_{\text{prop}}(Y \cap L_1(\theta)) = m_{\text{cl}_1}(Y \cap L_1(\theta)) / (\pi q(\theta)),$$

where  $\theta$  is distributed according to  $q$ . Noticing that

$$E(m_{\text{prop}}(Y \cap L_1) \mid L_2) = \frac{1}{\pi} \int_0^\pi m_{\text{cl}_1}(Y \cap L_1(\theta)) \, d\theta$$

and utilising that the nucleator in an unbiased estimator of  $V(Y)$  one finds that also  $m_{\text{prop}}(Y \cap L_1)$  is an unbiased estimator of  $V(Y)$ .

Statements concerning the variance of  $m_{\text{prop}}$  can be accessed from the variance decomposition

$$\text{Var}(m_{\text{prop}}(Y \cap L_1)) = \text{Var}(E[m_{\text{prop}}(Y \cap L_1) \mid L_2]) + E[\text{Var}(m_{\text{prop}}(Y \cap L_1) \mid L_2)].$$

It is seen that the minimal within section variance is obtained when, for almost all  $L_2$ ,  $\text{Var}(m_{\text{prop}}(Y \cap L_1) \mid L_2) = 0$ . This is achieved if  $m_{\text{cl}_1}(\tilde{Y}_2 \cap L_1(\theta))$  is proportional to  $m_{\text{cl}_1}(Y \cap L_1(\theta))$ . In applications, this implies that the gain in efficiency from using  $m_{\text{prop}}$  instead of  $m_{\text{cl}_1}$  depends on the quality of  $\tilde{Y}_2$ .

This proportionator nucleator approach is an attempt to use more of the observed information from the section plane  $L_2$  and incorporate this in the estimation procedure in order to increase the efficiency in terms of within section variance reduction and at the same time keep the workload at a minimum. A simulation study in Jensen (2000) indicates that it can indeed be expected for such a proportionator nucleator volume estimator to have a reduced within section variability if there is an appropriate resemblance between the actual transect of the object of interest and the rough estimate thereof.

In the neurone data set analysed in this thesis (see *Paper B* below) it was found that the estimator variance of  $m_{\text{prop}}$  is smaller than the estimator variance of  $m_{\text{cl}_1}$  but larger than the estimator variance of  $m_{\text{cl}_2}$ . As stated in the discussion of *Paper B* this result might be explained by the following: (i) the estimator  $m_{\text{cl}_2}$  has already a small within section variability due to the antithetic effect of the two perpendicular lines and (ii) the true cell transects of this particular data set are not very irregular in shape and hence the variance reducing effect of weighted sampling cannot compensate for the effect that some extra within section variability is introduced when  $m_{\text{prop}}$  is used in case of non-perfect segmentation.

In *Paper B* to be presented next, we use the information available in the segmented cell transects  $\tilde{Y}_2$  in a semi-automatic estimation procedure. Thereby, an attractive new method for local stereological volume estimation, providing an increase in efficiency compared to the classical nucleator, is achieved.

### 1.3 Semi-automatic procedures (Paper B)

The basic idea of *Paper B* is to combine the nucleator technique with the so-called *integrated nucleator*. The integrated nucleator is a local estimator of volume based on information from the entire transect  $Y \cap L_2$ . The estimator is



given by (Jensen, 1998)

$$m_{\text{int}}(Y \cap L_2) = 2 \int_{Y \cap L_2} d(y, o) \, dy^2.$$

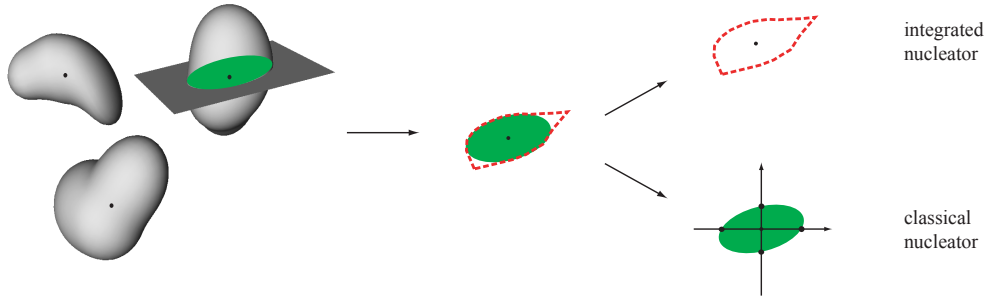
Here,  $dy^2$  denotes the element of the area measure on  $L_2$ .

Even though, the integrated nucleator has zero within section variance,  $m_{\text{int}}$  has not been used much in practice. The explaining factor might be that  $m_{\text{int}}$  requires well-behaved automatic image segmentation.

In *Paper B* we propose a semi-automatic procedure that requires an expert to judge whether the segmentation  $\tilde{Y}_2$  is satisfactory (according to some criteria) and then a semi-automatic estimation procedure is performed,

$$m_{\text{semi}}(Y \cap L_2) = \begin{cases} m_{\text{int}}(\tilde{Y}_2), & \text{if } \tilde{Y}_2 \text{ is judged satisfactory,} \\ m_{\text{cl}_2}(Y \cap L_1), & \text{otherwise} \end{cases}.$$

For an illustration see Figure 1.2. The resulting estimator is called *the semi-automatic nucleator*.



**Figure 1.2:** Illustration of the semi-automatic nucleator. The cell transect  $Y \cap L_2$  is shown as the green set while the segmented cell transect  $\tilde{Y}_2$  is delineated by the dashed red curve. The integrated nucleator is used if the segmentation is judged satisfactory, otherwise the classical nucleator is used.

When  $\tilde{Y}_2$  is judged satisfactory  $m_{\text{semi}}$  benefits from the low variation property of  $m_{\text{int}}$ . Otherwise, the expert conducts a traditional nucleator procedure, using  $m_{\text{cl}_2}$  and disregarding  $\tilde{Y}_2$ . Hence, the workload of the semi-automatic procedure is closely related to the segmentation quality and it will never be larger than that associated with  $m_{\text{cl}_2}$ .

It is not guaranteed that  $m_{\text{semi}}$  is unbiased but usually the bias will be small. In *Paper B* the mean square error as well as variance relations for all the above mentioned local stereological volume estimators are derived. It is shown that  $m_{\text{semi}}$  is likely to have smaller mean square error than  $m_{\text{cl}_2}$ . The magnitude of the gain in efficiency by using  $m_{\text{semi}}$  instead of  $m_{\text{cl}_2}$  will depend on the shapes of the cells and, for a given cell population, on how large the fraction of cells with satisfactory segmentation is. The semi-automatic estimator is exemplified and compared to the various other estimators in a study of

somatostatin positive inhibitory interneurons from mice hippocampi, observed by optical fluorescent microscopy. In this example it is found that, in order to obtain a relative error of the estimate of mean cell volume of 0.05, it is needed to sample and analyze 134, 150 and 189 cells using the integrated nucleator, semi-automatic nucleator and classical nucleator, respectively. Therefore, the semi-automatic procedure introduce a significant reduction in workload in this example.

For the general application to microscopy, it is required that the computerised image segmentation is able to detect the transect boundary reasonably well in order to obtain a significant gain in efficiency. Thus, the new semi-automatic method presented in *Paper B* is mainly applicable in microscopes without over-projection, such as confocal laser microscopes.

## 1.4 Particle models (Paper C and Paper D)

As pointed earlier out one of the major advances of stereology is that no specific model assumptions are required for the construction of unbiased stereological estimation of a number of quantitative properties of spatial structures. But for the assessment of the precision of the estimators, a modelling approach is required. This is the topic of *Paper C* and *Paper D* of the present thesis. Below, we give a short summary of the results obtained. We refer to the objects under study as particles but the concept is to be understood in a much broader context as such a model can be used for modelling a wide range of objects such as grains (Sato et al., 1996), cells (Wicksell, 1925, 1926) and asteroids and comets (Muinonen, 1998).

In the beginning of the twentieth century, Wicksell proposed the corpuscle problem in two consecutive papers Wicksell (1925, 1926) where the main task is to recover the size distribution of spherical particles. This is now known as the Wicksell problem and has been revisited numerous times (Mecke and Stoyan, 1980; Hoogendoorn, 1992) starting a line of papers in the literature. For reviews on the subject see Ripley (1981); Stoyan et al. (1995); Weibel (1980). The Wicksell problem still enjoys attention in the mathematical literature but for many applications Wicksells spherical model is often too restrictive. In the nineteen seventies, it was suggested to model a particle population by isotropically oriented (prolate or oblate) ellipsoids positioned in space according to a Poisson point process (Cruz-Orive, 1976, 1978). Later the isotropic orientation assumption was relaxed (Girling, 1993). One pursued method for randomising the spherical particle is to randomise its radius, e.g. using a lognormal distribution (Muinonen et al., 1996).

As pointed out by Miller et al. (1994) most real-world shapes and images are strongly structured but at the same time they are indeed highly variable and not rigid, like crystals. Miller et al. (1994) consider modelling a spatial particle with no obvious landmarks. They represent the particle by a global translation and a normal deformation of a sphere where the normal deformation

is defined using the orthonormal spherical harmonic basis, see also Grenander (1993). The spherical deformation model is revisited in Hobolth (2003) where it is employed to summarise the shape of star-shaped three-dimensional objects using few parameters. See also the early paper Hobolth and Jensen (2000) where they consider a two-dimensional particle modelling of similar type.

Common for the above models are that they assume a predetermined rigid shape (spherical or ellipsoidal) or use a deformation of such. Over the last decades, a new approach for particle modelling has been suggested. The idea is to identify the particle  $Y$  with its radial function  $R_Y$ . This approach requires that  $Y$  is compact and star-shaped with respect to some fixed point  $o$  inside  $Y$ . Let  $\mathbb{S}^2 = \{x \in \mathbb{R}^3 : \|x\| = 1\}$  denote the unit sphere. Then, the radial function in direction  $u \in \mathbb{S}^2$  is given by

$$R_Y(u) = \max\{r \geq 0 : o + ru \in Y\},$$

(Gardner, 2006). For an illustration of a two-dimensional particle  $Y$  and its radial function in direction  $u$ , see Figure 1.4 (left).

The approach is now to assume a random field model for the radial function. In Hobolth et al. (2003) they propose a parametric model for the normalised radial function using polar Fourier expansion where the expected phase amplitudes are modelled by a simple regression and the parameters of the regression have nice geometric interpretations.

### 1.4.1 Lévy particles (Paper C and Paper D)

Both *Paper C* and *Paper D* consider three-dimensional particle models using a random field model for the associated radial function. In *Paper C* the radial function is modelled as a kernel smoothing of a Lévy basis of the form

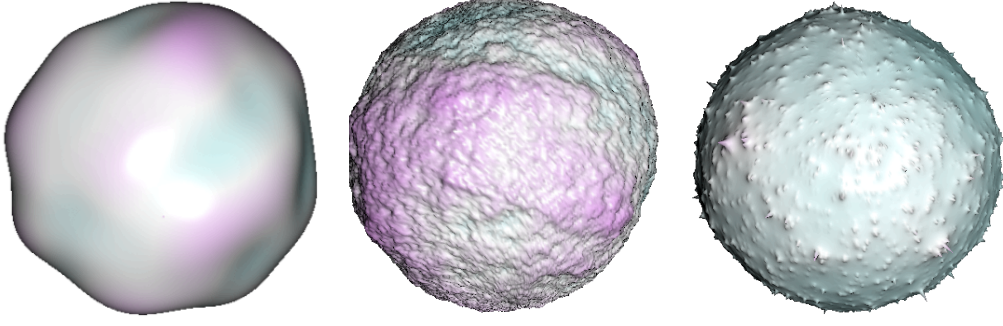
$$R_Y(u) = \int_{\mathbb{S}^2} K(v, u) \Gamma(dv), \quad u \in \mathbb{S}^2. \quad (1.1)$$

Here,  $K$  denotes a deterministic kernel function and  $\Gamma$  is a Lévy basis on the Borel subsets of  $\mathbb{S}^2$ . Figure 1.3 shows examples of simulated Lévy particles. The examples are taken from *Paper C* which contains further examples of simulated Lévy particles along with the simulation algorithm used.

A more general version of the model (1.1) is considered in Jónsdóttir et al. (2008). In *Paper D* a power transformation of the model is considered.

In the following a short introduction to Lévy bases as considered in connection with the model (1.1) is provided along with some nice properties thereof. For further expositions of Lévy bases, cf. Hellmund et al. (2008); Jónsdóttir et al. (2008); Barndorff-Nielsen and Schmiegel (2004).

A random measure  $\Gamma$  on the Borel subsets  $\mathcal{B}(\mathbb{S}^2)$  of  $\mathbb{S}^2$  is said to be a Lévy basis (Barndorff-Nielsen and Schmiegel, 2004) if  $\Gamma$  is (i) an infinitely divisible and (ii) independently scattered random measure, i.e. if  $\Gamma$  fulfils



**Figure 1.3:** Examples of simulated Lévy particles under the model 1.1.

- (i) for every  $A \in \mathcal{B}(\mathbb{S}^2)$  and every  $n \in \mathbb{N}$  there exist a sequence of independent and identically distributed random variables  $\Gamma_1, \Gamma_2, \dots, \Gamma_n$  such that  $\sum_{i=1}^n \Gamma_i \sim \Gamma(A)$ .
- (ii) for  $\{A_n\}_n \subseteq \mathcal{B}(\mathbb{S}^2)$  a sequence of disjoint Borel sets, then  $\{\Gamma(A_n)\}_n$  are independent random variables and  $\sum_n \Gamma(A_n) = \Gamma(\cup_n A_n)$  almost surely.

Furthermore,  $\Gamma$  is called a non-negative Lévy basis if  $\Gamma(A)$  is positive for all  $A \in \mathcal{B}(\mathbb{S}^2)$ . In *Paper C* and *Paper D*, three different Lévy bases are considered, a Gaussian basis, a gamma basis and an inverse Gaussian basis, respectively, see Table 1.1.

**Table 1.1:** The three Lévy bases considered in the thesis.

Lévy basis	$\Gamma(\mathrm{d}v)$	parameters
Gaussian	$\mathcal{N}(\mu \mathrm{d}v, \sigma^2 \mathrm{d}v)$	$\mu \in \mathbb{R}, \sigma^2 > 0$
gamma	$\mathrm{Ga}(\kappa \mathrm{d}v, \tau)$	$\kappa > 0, \tau > 0$
inverse Gaussian	$\mathrm{IG}(\delta \mathrm{d}v, \gamma)$	$\delta > 0, \gamma > 0$

For the gamma distribution  $\mathrm{Ga}(\kappa, \tau)$  and the inverse Gaussian distribution  $\mathrm{IG}(\delta, \gamma)$  the following parametrisation of the respective densities are used

$$\frac{\tau^\kappa}{\Gamma(\kappa)} y^{\kappa-1} e^{-\tau y}, y > 0. \quad \text{and} \quad \frac{\delta e^{\delta \gamma}}{\sqrt{2\pi}} y^{3/2} \exp\left\{-\frac{1}{2}\left(\delta^2 \frac{1}{y} + \gamma^2 y\right)\right\}, y > 0$$

An attractive property of a Lévy integral such as (1.1) is the tractability of its cumulant function. Here, following (Barndorff-Nielsen and Schmiegel, 2004), the cumulant function  $C(\cdot, W)$  of a random variable  $W$  denotes the logarithm of its characteristic function, i.e.  $C(t, W) = \log \mathbb{E}(\exp\{itW\})$ , for  $t \in \mathbb{R}$ . The tractability of the cumulant function of the integral (1.1) is a consequence of the Lévy–Khintchine representation of  $C(t, \Gamma(A))$ , cf. e.g. Hellmund et al. (2008),

$$C(t, \Gamma(A)) = ita(A) + \frac{1}{2}t^2b(A) + \int_{\mathbb{R}} [e^{itr} - 1 - itr\mathbb{1}_{[-1,1]}(r)] U(\mathrm{d}r, A).$$

Here,  $a$  is a signed measure,  $b$  is a positive measure and, for fixed  $r$ ,  $U(dr, \cdot)$  is a measure on  $\mathcal{B}(\mathbb{S}^2)$ . Moreover, for fixed  $A$  then  $U(\cdot, A)$  is a Lévy measure on the Borel sets  $\mathcal{B}(\mathbb{R})$  of  $\mathbb{R}$ . Indeed,  $U$  if a measure on the product space  $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{S}^2)$  and the triplet  $(a, b, U)$  is known as the characteristic triplet of  $\Gamma$ .

In this thesis, only Lévy bases where the measures  $a$  and  $b$  are absolutely continuous with respect to the element of Lebesgue measure on  $\mathbb{S}^2$  are considered, i.e.  $a(dv) = \tilde{a}(v) dv$  and  $b(dv) = \tilde{b}(v) dv$ . Moreover, the considered Lévy bases have the property that  $U(\cdot, dv) = V(\cdot, v) dv$ , where  $V(\cdot, v)$  is a Lévy measure for each fixed  $v \in \mathbb{S}^2$ . The cumulant function of  $\Gamma(A)$  has the representation

$$C(t, \Gamma(A)) = \int_A C(t, \Gamma'(v)) dv,$$

where  $\Gamma'(v)$  is the random variable with cumulant function

$$C(t, \Gamma'(v)) = it\tilde{a}(v) + \frac{1}{2}t^2\tilde{b}(v) + \int_{\mathbb{R}} [e^{itr} - 1 - itr\mathbb{1}_{[-1,1]}(r)] V(dr, v).$$

The so-called *spot variable*  $\Gamma'(v)$  characterises  $\Gamma$  at position  $v$ .

Integration of a measurable function  $f$  on  $(\mathbb{S}^2, \mathcal{B}(\mathbb{S}^2))$  with respect to a Lévy basis  $\int_{\mathbb{S}^2} f(v) \Gamma(dv)$  obey an appealing cumulant function property (Hellmund et al., 2008; Rajput and Rosinski, 1989),

$$C\left(t, \int_{\mathbb{S}^2} f(v) \Gamma(dv)\right) = \int_{\mathbb{S}^2} C(tf(v), \Gamma'(v)) dv.$$

Using this result, it is possible to obtain moment results for  $\int_{\mathbb{S}^2} f(v) \Gamma(dv)$ . For the first and second order moments one finds,

$$\mathbb{E}\left(\int_{\mathbb{S}^2} f(v) \Gamma(dv)\right) = \int_{\mathbb{S}^2} f(v) \Gamma'(v) dv$$

and

$$\text{Var}\left(\int_{\mathbb{S}^2} f(v) \Gamma(dv)\right) = \int_{\mathbb{S}^2} f(v)^2 \Gamma'(v) dv.$$

For two measurable functions  $f_1, f_2$ , the mutual covariance of the corresponding Lévy integrals is given by

$$\text{Cov}\left(\int_{\mathbb{S}^2} f_1(v) \Gamma(dv), \int_{\mathbb{S}^2} f_2(v) \Gamma(dv)\right) = \int_{\mathbb{S}^2} f_1(v)f_2(v) \Gamma'(v) dv. \quad (1.2)$$

Notice that, for each  $u$ , the right hand side of (1.1) is a Lévy integral and hence the above moment relations apply for the model. For the Lévy bases of Table 1.1 the function  $\tilde{a}$  and  $\tilde{b}$  and  $V(r, v)$ , for every fixed  $r$ , are independent of  $v$  and hence the Lévy basis  $\Gamma$  is homogeneous (Hellmund et al., 2008) and the distribution of the  $\Gamma'(v)$  does not depend on  $v$ .

Lévy bases have also been used in point process modelling for driving the random intensity of shot-noise Cox point processes (Wolpert and Ickstadt, 1998) and of the more general Cox point processes (Hellmund et al., 2008).

### 1.4.2 Stereology and Lévy particles (Paper D)

It is generally not possible to provide explicit expressions on the precision of a design based estimator. Often, the precision is discussed in terms of asymptotic variance results. The task is often simplified (or made possible) by assuming a model for the object under study. Naturally, any model assumptions need to be justifiable for the applications in mind and can therefore not be too restrictive.

With the earlier presented nucleator estimators in mind a model for a suitable power transformation of the radial function seems attractable. This is exactly what is considered in *Paper D* and this modelling approach goes well together with local stereological methods. The considered model for the radial function is

$$R_Y(u) = \left\{ \int_{\mathbb{S}^2} K(v, u) \Gamma(\mathrm{d}v) \right\}^{1/p}, \quad u \in \mathbb{S}^2. \quad (1.3)$$

Here,  $\Gamma$  is a nonnegative Lévy basis and  $p > 0$ . In *Paper D* the focus is on the case where  $K$  is isotropic, in that, for points  $v, u \in \mathbb{S}^2$ ,  $K(v, u) = k(\vartheta(v, u))$  depends only on the great-circle distance between  $v$  and  $u$  only. It is shown that under the model (1.3) with  $p = 3$  and an arbitrary isotropic kernel, the distribution of the volume of the random particle  $Y$  belongs to the same distributional family as the Lévy basis itself. Moreover, the distribution of local stereological volume estimators such as the nucleator and the integrated nucleator are considered and their respective distributions are derived.

Local stereology goes particularly well together with a modelling approach based on the radial function. The reason is that an observation of the radial function in a planar section  $Y \cap L_2$  in direction  $u$  coincides with the radial function of  $Y$  for each  $u \in \mathbb{S}^2 \cap L_2$ , i.e.  $R_{Y \cap L_2}(u) = R_Y(u)$ , for  $u \in Y \cap L_2$ . Here,  $L_2$  denotes a plane through  $o$ .

If instead observations are available from projections of the particle onto a projection plane, then a different course of action might be desirable. In geometric tomography where information is often available from particle projections the support function plays an important role. The reason is that the support function is an important tool when analysing particle projections, see Gardner (2006). The support function is defined for convex sets and hence we now restrict  $Y$  to be convex. The support function  $H_Y$  of  $Y$  is given by (Gardner, 2006),

$$H_Y(u) = \max\{u \cdot v : v \in Y\}, \quad u \in \mathbb{S}^2.$$

The support function in direction  $u$  gives the distance from  $o$  to the supporting hyperplane with normal vector  $u$ . For an illustration in a two-dimensional setting, see Figure 1.4 (right).

Following (1.1) the support function  $H_Y(u)$  is modelled by

$$H_Y(u) = \int_{\mathbb{S}^2} h(v, u) \Gamma(\mathrm{d}v), \quad u \in \mathbb{S}^2. \quad (1.4)$$

For the right hand side of (1.4) to be a support function it is required that  $\Gamma$  is nonnegative and that for each  $v \in \mathbb{S}^2$ , the kernel function  $h(v, \cdot)$  is a support



**Figure 1.4:** For the two-dimensional particles  $Y$  (grey), the radial function  $R_Y(u)$  in direction  $u$  (left) and the support function  $H_Y(\tilde{u})$  in direction  $\tilde{u}$  along with the supporting hyperplane (dashed line) with normal vector  $\tilde{u}$  (right) are shown.

function. Then the linearity of integration with respect to  $\Gamma(dv)$  assure that the right side of (1.4) becomes a support function (Schneider, 1993, Theorem 1.7.1). Recall that a function is sub-linear if and only if it is both convex and homogeneous. Thus, as  $x \mapsto x^{1/p}$  is convex for  $p \leq 1$ , we also have the more general model

$$H_Y(u) = \left\{ \int_{\mathbb{S}^2} h(v, u) \Gamma(dv) \right\}^{1/p}, \quad p \leq 1.$$

An explicit relation between the radial function and the support function might allow for a comparison of local and pivotal volume estimators under the same Lévy particle model. Such a relation is at present unknown to us. Though, for a strict convex set  $Y$  a more implicit relation do exist, as  $R_Y(\nabla_{H_Y}(u)) = 1$ , where  $\nabla_{H_Y}(u)$  denotes the gradient of  $H_Y$  at  $u$  (Schneider, 1993, Corollary 1.7.3).

We notice that the idea of leaving a strictly design based set-up and instead use a combined design based and model based set-up was also used in *Paper A* where it allowed us to obtain an optimal predictor for the total number of cells in a cell population.

### 1.4.3 Surface behaviour of Lévy particles (Paper C)

Due to the freedom in the choice of both kernel and Lévy basis the Lévy particle model (1.1) is very flexible. Thus, it is possible to model a rich class of particles. *Paper C* studies some aspects of this model with focus on the surface behaviour of  $Y$  considered in terms of the Hausdorff dimension.

The Hausdorff (fractal) dimension  $D(Z)$  of a set  $Z$  is a number quantifying the roughness of the set  $Z$  when observed at infinitesimal scale. The Hausdorff dimension of a random field is closely connected to the second order properties of the random field (Adler, 1981). For stationary random fields,  $D$  is determined by the behaviour of the covariance function near the origin.

In *Paper C* the model (1.1) is studied for isotropic kernels. For such a kernel, using (1.2), the correlation of  $R_Y(u), R_Y(u')$  of points  $u, u'$  at great-circle distance  $\theta \in [0, \pi]$ , can be expressed as

$$C(\theta) = \frac{2}{c_2} \int_0^\pi \int_0^\pi k(\lambda) k(\arccos(\sin \lambda \sin \theta \cos \phi + \cos \lambda \cos \theta)) \sin \lambda \, d\lambda \, d\phi, \quad (1.5)$$

where  $c_2 = 2\pi \int_0^\pi k(\lambda)^2 \sin \lambda \, d\lambda$ .

Let  $Z = \{(u, R_Y(u)) : u \in \mathbb{S}^2\}$  denote the surface (shell) of  $Y$  and let  $Z_c = \max\{Z, c\}$ ,  $c > 0$ , denote a cut-off version thereof. If  $\Gamma$  is a Gaussian Lévy basis then *Paper C* establish that the Hausdorff dimension of  $Z_c$ , is determined by the asymptotic behaviour of  $C(\theta) - C(0)$  near zero. For different parametric kernel families the asymptotic behaviour of  $C(\theta) - C(0)$  near zero is determined and the corresponding Hausdorff dimension of  $Z_c$  is provided. For *the von Mises–Fisher kernel*,

$$k(\theta) = \exp\{a \cos \theta\}, \quad 0 \leq \theta \leq \pi, a \in \mathbb{R},$$

and *the uniform kernel*

$$k(\theta) = \mathbb{1}_{\{\theta < r\}}, \quad 0 \leq \theta \leq \pi, a \in \mathbb{R},$$

the particle surface  $Z_c$  is shown to have Hausdorff dimension 2 and 2.5, respectively. For the *power kernel*,

$$k(\theta) = (\theta/\pi)^{-q} - 1, \quad 0 < \theta \leq \pi, 0 < q < 1,$$

the asymptotic behaviour of  $C(\theta) - C(0)$  near zero is conjectured to be proportional to  $\theta^{2(1-q)}$  which leads to the corresponding particle surface having Hausdorff dimension  $2 + q$ .

In the process of determining the asymptotic behaviour of  $C(\theta) - C(0)$  near zero for the truncated kernel family, the scheme for a random field over  $\mathbb{R}^2$  was also considered. In the following the results found in this planar setting are reported. These results do not appear in any of the four papers attached to the present thesis.

## 1.5 Surface behaviour of a Lévy field on $\mathbb{R}^2$

Let  $X = \{X(t) : t \in \mathbb{R}^2\}$  be the centred, real valued random field, given by

$$X(t) = \int_{\mathbb{R}^2} K(s, t) \Gamma(ds).$$

Here,  $K$  is a deterministic kernel function and  $\Gamma$  is the centred, homogeneous Gaussian basis on the Borel subsets of  $\mathbb{R}^2$  given by  $\Gamma(dv) \sim \mathcal{N}(0, \sigma^2 dv)$ , for  $\sigma^2 > 0$ . This model is a special case of the linear spatio-temporal Lévy model



proposed in Jónsdóttir et al. (2008). We assume that  $K$  is isotropic, that is, for  $t, s \in \mathbb{R}^2$ ,  $K(v, u) = k(\|t - u\|)$  is a , i.e.  $k(s, t)$  is a function of the Euclidean distance  $\|t - s\|$  between  $t$  and  $s$  only.

Let  $(t_1, t_2) \in \mathbb{R}^2$  such that  $\|t_1 - t_2\| = h \in [0, \infty)$ . Then, using polar decomposition and Jónsdóttir et al. (2008, Eq. (11)), we find

$$\text{Cov}(X(t_1), X(t_2)) = 2\sigma^2 \int_0^\infty r k(r) \int_0^\pi k(\sqrt{r^2 + h^2 - 2rh \cos \phi}) d\phi dr. \quad (1.6)$$

This shows that  $\text{Cov}(X(t_1), X(t_2))$  only depends on  $\|t_1 - t_2\|$ , i.e.  $X$  is isotropic. That is, the isotropic nature of the kernel function is carried over to the random field. Let  $c_2 = \int_0^\infty r k(r)^2 dr$ . Then, the correlation function of  $X$  becomes

$$C(h) = \frac{1}{\pi c_2} \int_0^\infty r k(r) \int_0^\pi k(\sqrt{r^2 + h^2 - 2rh \cos \phi}) d\phi dr.$$

The Gaussian distribution of the Lévy basis is carried over to the random field such that  $X$  is again a centred Gaussian random field with covariance determined by (1.6). The Hausdorff dimension of the graph  $\text{Gr } X = \{(t, X(t)) : t \in \mathbb{R}^2\}$  of  $X$  is  $D(\text{Gr } X) = 3 - \alpha/2$ , where  $\alpha$  is given by

$$\begin{aligned} \alpha &= \sup \{ \beta > 0 : [C(0) - C(h)] = o(h^\beta), h \downarrow 0 \} \\ &= \inf \{ \beta > 0 : h^\beta = o([C(0) - C(h)]), h \downarrow 0 \}, \end{aligned}$$

(Adler, 1981). The number  $\alpha$  is called the (fractal) index of  $X$ .

In the following, the behaviour of  $[C(0) - C(h)]$  is studied in this planar setting under the assumption of the truncated power kernel. The truncated power kernel is in this setting given by

$$k(r) = \begin{cases} r^{-q} - 1, & 0 < r \leq 1 \\ 0, & \text{otherwise} \end{cases}, \quad 0 < q < 1. \quad (1.7)$$

Provided that the integral (1.8) is strictly positive the asymptotic behaviour of  $C(0) - C(h)$  near 0 is given by Theorem 1.1 below. The Hausdorff dimension  $D(\text{Gr } X) = 2 + q$  is then a direct consequence thereof.

**Theorem 1.1.** *Let  $0 < q < 1$  and*

$$A_q = \int_0^\infty x^{1-q} \int_0^\pi [x^{-q} - (x^2 + 1 - 2x \cos \phi)^{-q/2}] d\phi dx \in [0, \infty). \quad (1.8)$$

*Let  $C$  be the covariance function (1.6) with  $k$  the power kernel (1.7). Provided that  $A_q > 0$ , then*

$$[C(0) - C(h)] \sim 2\pi c_2 A_q h^{2(1-q)}, \quad h \rightarrow 0.$$

If needed, the integral  $A_q$  can be found using numerical integration.

*Proof.* Let  $h \geq 0$ . For  $r \in (0, 1)$ , let  $g(h, r, \phi) = \sqrt{r^2 + h^2 - 2rh \cos \phi}$  and  $A(r) = \{ \phi \in [0, \pi] : g(h, r, \phi) \in (0, 1] \}$ . Then,

$$\frac{C(0) - C(h)}{\pi c_2} = \int_0^1 r[r^{-q} - 1] \left\{ \int_0^\pi [r^{-q} - 1] d\phi - \int_{A(r)} [g(h, r, \phi)^{-q} - 1] d\phi \right\} dr.$$

For  $0 < r < 1 - h$ ,  $A(r) = [0, \pi]$ , whereas  $A(r) \subset [0, \pi]$ , for  $1 - h < r < 1$ . Hence, we conduct the integration with respect to  $r$  in a two-step procedure. First, notice that a Taylor expansion around zero implies  $((1 - h)^{-q} - 1) = O(h)$ . Furthermore, for  $h$  small enough, say  $h < h_0$ , the integral difference within the curly brackets is smaller than one. Now, for  $h < h_0$ , an application of the first mean value theorem for integration yield that there exists  $t \in (1 - h, 1)$  such that

$$\begin{aligned} & \int_{1-h}^1 r[r^{-q} - 1] \left\{ \int_0^\pi [r^{-q} - 1] d\phi - \int_{A(r)} [g(h, r, \phi)^{-q} - 1] d\phi \right\} dr \\ &= ht[t^{-q} - 1] \left\{ \int_0^\pi [t^{-q} - 1] d\phi - \int_{A(t)} [g(h, t, \phi)^{-q} - 1] d\phi \right\} \\ &\leq h[(1 - h)^{-q} - 1] \\ &= O(h^2). \end{aligned}$$

Applying the substitution  $r = hx$  yields for the remaining integral term

$$\begin{aligned} & \int_0^{1-h} r[r^{-q} - 1] \int_0^\pi [r^{-q} - g(h, r, \phi)^{-q}] d\phi dr \\ &= h^{2(1-q)} \int_0^{\frac{1-h}{h}} x[x^{-q} - h^q] \int_0^\pi [x^{-q} - g(1, x, \phi)^{-q}] d\phi dx. \end{aligned}$$

Notice that, for  $x \rightarrow 0$ ,

$$\int_0^\pi [x^{-q} - g(1, x, \phi)^{-q}] d\phi = \pi x^{-q} - \int_0^\pi g(1, x, \phi)^{-q} d\phi = \pi x^{-q} + O(x^0) \sim \pi x^{-q}.$$

Furthermore, for  $x \rightarrow \infty$ , a Taylor expansion of  $(1 + y)^{-q/2}$  around  $y = 0$  yields,

$$\int_0^\pi [x^{-q} - g(1, x, \phi)^{-q}] d\phi = x^{-q} \int_0^\pi [1 - g(1, x^{-1}, \phi)^{-q}] d\phi \sim -\frac{\pi q^2}{4} x^{-q-2}.$$

Thus, both integrals

$$\int_0^\infty x^{1-q} \int_0^\pi [x^{-q} - g(1, x, \phi)^{-q}] d\phi dx$$

and

$$\int_0^\infty x \int_0^\pi [x^{-q} - g(1, x, \phi)^{-q}] d\phi dx$$

exist both at 0 and at  $\infty$ . Hence, we get the following limit result

$$\begin{aligned} & \int_0^{\frac{1-h}{h}} x [x^{-q} - h^q] \int_0^\pi [x^{-q} - g(1, x, \phi)^{-q}] \, d\phi \, dx \\ & \xrightarrow{h \rightarrow 0} \int_0^\infty x^{1-q} \int_0^\pi [x^{-q} - g(1, x, \phi)^{-q}] \, d\phi \, dx < \infty. \end{aligned}$$

which concludes the proof.  $\square$

The literature contains a rich discussion of Hausdorff dimension and the estimation thereof for Gaussian random fields over general Euclidean spaces, cf. Adler (1981); Xue and Xiao (2009); Gneiting et al. (2011) and references therein.

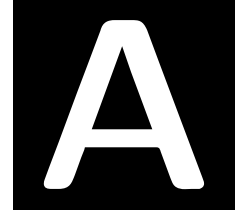
## Bibliography

- Abrahão, L. M., Nyengaard, J. R., Sasahara, T. H., Gomes, S. P., Oliveira Fda, R., Ladd, F. V., Ladd, A. A., Melo, M. P., Machado, M. R., Melo, S. R., and Ribeiro, A. A. (2009). Asymmetric post-natal development of superior cervical ganglion of paca (Agouti paca). *International Journal of Developmental Neuroscience*, 27:37–45.
- Adler, R. J. (1981). *The Geometry of Random Fields*. John Wiley and Sons, Chichester.
- Asmussen, S. and Glynn, P. W. (2007). *Stochastic Simulation*. Springer Science+Business Media, LLC.
- Baddeley, A. and Jensen, E. B. V. (2005). *Stereology for statisticians*. Chapman & Hall, Boca Raton.
- Baddeley, A. J. (1993). Stereology and survey sampling theory. In *Proceedings of the 49th Session of the International Statistical Institute*, volume 2, pages 435–449.
- Barndorff-Nielsen, O. E. and Schmiegel, J. (2004). Lévy-based tempo-spatial modelling; with applications to turbulence. *Uspekhi Matematicheskikh Nauk*, 159:63–90.
- Cochran, W. G. (1977). *Sampling Techniques*. John Wiley & Sons, New York.
- Cruz-Orive, L. M. (1976). Particle size-shape distributions: the general spheroid problem. I. Mathematical model. *Journal of Microscopy*, 107:235–253.
- Cruz-Orive, L. M. (1978). Particle size-shape distributions: the general spheroid problem. II. Stochastic model and particle guide. *Journal of Microscopy*, 112:153–167.
- Davy, P. J. and Miles, R. E. (1977). Sampling theory for opaque spatial specimens. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):56–65.
- Delesse, A. E. O. J. (1847). Procédé mécanique pour déterminer la composition des roches. *Comptes Rendues de l'Académie des Sciences*, 25:544–545.
- Delesse, A. E. O. J. (1848). Procédé mécanique pour déterminer la composition des roches. *Anales des Mines*, 13:379–388. Quatrième série.

- Gardi, J., Nyengaard, J., and Gundersen, H. (2008). The proportionator: Unbiased stereological estimation using biased automatic image analysis and non-uniform probability proportional to size sampling. *Computers in Biology and Medicine*, 38(3):313–328.
- Gardi, J. E., Nyengaard, J. R., and Gundersen, H. J. G. (2007). Automatic sampling for unbiased and efficient stereological estimation using the proportionator in biological studies. *Journal of Microscopy*, 230:108–120.
- Gardner, R. J. (2006). *Geometric tomography*. Cambridge University Press, New York.
- Girling, A. J. (1993). Shape analysis for the anisotropic corpuscle problem. *Journal of the Royal Statistical Society. Series B (Methodological)*, 55(3):675–686.
- Gneiting, T., Sevcikova, H., and Percival, D. B. (2011). Estimators of Fractal Dimension: Assessing the Roughness of Time Series and Spatial Data. <http://adsabs.harvard.edu/abs/2011arXiv1101.1444G>.
- Grenander, U. (1993). *General pattern theory*. Oxford University Press, Oxford.
- Gundersen, H. J. G. (1988). The nucleator. *Journal of Microscopy*, 151:3–21.
- Hellmund, G., Prokešová, M., and Jensen, E. B. V. (2008). Lévy based cox point processes. *Advances in Applied Probability*, 40:603–629.
- Hobolth, A. (2003). The spherical deformation model. *Biostatistics*, 4(4):583–595.
- Hobolth, A. and Jensen, E. B. V. (2000). Modelling stochastic changes in curve shape, with an application to cancer diagnostics. *Advances in Applied Probability*, 32(2):344–362.
- Hobolth, A., Pedersen, J., and Jensen, E. (2003). A continuous parametric shape model. *Annals of the Institute of Statistical Mathematics*, 55:227–242.
- Hoogendoorn, A. (1992). Estimating the weight undersize distribution for the wicksell problem. *Statistica Neerlandica*, 46(4):259–282.
- Horvitz, D. G. and Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47:663–685.
- Jensen, E. B., Baddeley, A. J., Gundersen, H. J. G., and Sundberg, R. (1985). Recent trends in stereology. *International Statistical Review*, 53(1):99–108.
- Jensen, E. B. V. (1998). *Local Stereology*. World Scientific Publishing, Singapore.
- Jensen, E. B. V. (2000). On the variance of local stereological volume estimators. *Image Analysis and Stereology*, 19:15–18.
- Jónsdóttir, K. Y., Schmiegel, J., and Jensen, E. B. V. (2008). Lévy based growth models. *Bernoulli*, 14:62–90.
- Mecke, J. and Stoyan, D. (1980). Stereological problems for spherical particles. *Mathematische Nachrichten*, 96(1):311–317.
- Melo, S. R., Nyengaard, J. R., da Roza Oliveira, F., Ladd, F. V. L., Abrahão, L. M. B., Machado, M. R., Sasahara, T. H., de Melo, M. P., and Ribeiro, A. A. C. (2009). The developing left superior cervical ganglion of pacas (agouti paca). *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology*, 292(7):966–975.

- Miles, R. E. (1978a). *The importance of proper model specification in stereology*, pages 115–136. Springer.
- Miles, R. E. (1978b). The sampling, by quadrats, of planar aggregates. *Journal of Microscopy*, 113:257–267.
- Miles, R. E. and Davy, P. J. (1976). Precise and general conditions for the validity of a comprehensive set of stereological fundamental formulae. *Journal of Microscopy*, 107:211–226.
- Miles, R. E. and Davy, P. J. (1977). On the choice of quadrats in stereology. *Journal of Microscopy*, 110:27–44.
- Miller, M. I., Joshi, S., Maffitt, D. R., McNally, J. G., and Grenander, U. (1994). Membranes, mitochondria and amoebae: shape models. *Journal of Applied Statistics*, 21(1):141–163.
- Muinsonen, K. (1998). Introducing the Gaussian shape hypothesis for asteroids and comets. *Astronomy and Astrophysics*, 332:1087–1098.
- Muinsonen, K., Nousiainen, T., Fast, P., Lumme, K., and Peltoniemi, J. I. (1996). Light scattering by Gaussian random particles: ray optics approximation. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 55(5):577–601.
- Rajput, B. S. and Rosinski, J. (1989). Spectral representations of infinitely divisible processes. *Probability Theory and Related Fields*, 82:451–487.
- Ripley, B. D. (1981). *Spatial statistics*. John Wiley & Sons, Inc., Hoboken, New Jersey.
- Sato, E., Kondo, N., and Wakai, F. (1996). Particle size, shape and orientation distributions: A general spheroid problem and application to deformed  $\text{Si}_3\text{N}_4$  microstructures. *Philosophical Magazine A*, 74(1):215–228.
- Schneider, R. (1993). *Convex bodies: the Brunn-Minkowski theory*. University Press.
- Stoyan, D. (1990). Stereology and stochastic geometry. *International Statistical Review*, 58(3):227–242.
- Stoyan, D., Kendall, W. S., and Mecke, J. (1995). *Stochastic Geometry and Its Applications*. Akademie-Verlag, 2nd edition.
- Thompson, S. (1992). *Sampling*. John Wiley & Sons.
- Weibel, E. R. (1979). *Stereological methods, 1. Practical methods for biological morphometry*. Academic Press, London.
- Weibel, E. R. (1980). *Stereological methods, 2. Theoretical foundations*. Academic Press, London.
- Wicksell, S. D. (1925). The corpuscle problem. A mathematical study of a biometric problem. *Biometrika*, 17:84–89.
- Wicksell, S. D. (1926). The corpuscle problem. Second memoir. Case of ellipsoidal corpuscles. *Biometrika*, 18:152–172.
- Wolpert, R. L. and Ickstadt, K. (1998). Poisson/gamma random field models for spatial statistics. *Biometrika*, 85(2):251–267.
- Xue, Y. and Xiao, Y. (2009). Fractal and Smoothness Properties of Space-Time Gaussian Models. <http://adsabs.harvard.edu/abs/2009arXiv0912.0285X>.





# Image-based empirical importance sampling: an efficient way of estimating intensities

Linda V. Hansen, Markus Kiderlen and Eva B. Vedel Jensen

*Centre for Stochastic Geometry and Advanced Bioimaging, Department of Mathematical Sciences, Aarhus University*

**Abstract:** Very recently, it has been suggested in the biomedical literature to combine computerised image analysis with non-uniform sampling in order to increase the efficiency of estimators of intensities of biological cell populations. We give this ingenious idea of empirical importance sampling a stochastic formulation, using point process theory and modern sampling theory. We develop statistical tools for assessing its efficiency and construct optimal model-based estimators of intensities. Examples of applications of empirical importance sampling in microscopy are provided.

**Key words:** image analysis, importance sampling, probability proportional to size, proportionator, stereology.





## A.1 Introduction

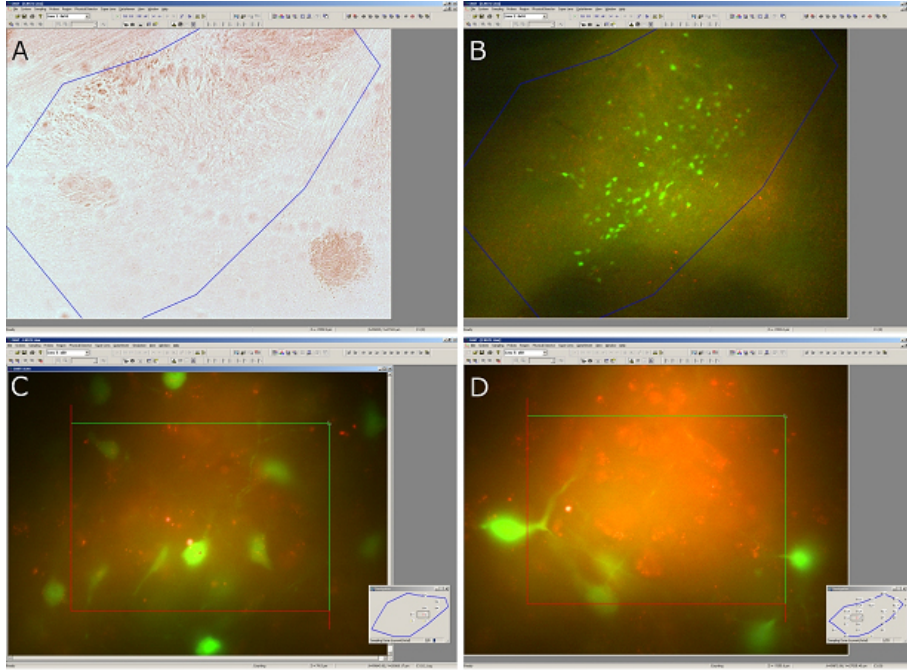
Importance sampling is a general statistical technique for estimating properties of a particular distribution, based on samples from a different distribution than the one of interest, cf. Asmussen and Glynn (2007) and references therein. Depending on the application, the term may refer to the process of sampling from this alternative distribution, the process of inference, or both. By choosing the alternative distribution appropriately, importance sampling may result in a marked increase in estimator efficiency. The basic idea is to choose an alternative sampling distribution in such a way that most of the sampling is done in the part of the state space that contributes the most to the parameter of interest.

This type of sampling technique has recently been introduced in computerised analysis of microscopy images under the name of *the proportionator*, Gardi et al. (2008, 2007a), see also the early paper Dorph-Petersen et al. (2000). This technique addresses the essential problem in the biomedical sciences that cell populations often show pronounced inhomogeneity, being present only in structured layers or showing marked gradients. Observing such a cell population, using a systematic set of fields of view, will be highly inefficient because most fields will contain no or very few cells.

An example of application of the technique suggested in Gardi et al. (2008, 2007a) is shown in Figure A.1. The aim is here to estimate the total number of green GFP-expressing neurones, see panel B. Under low magnification, the complete region (panel A) of interest is delineated, and by automatic image analysis every field of view inside this region is given a weight proportional to the amount of green colour observed under fluorescence illumination. A field of view is automatically sampled with a probability proportional to its weight and the number of neurones seen in the sampled field is determined under high magnification by an expert, see panels C and D. Further examples of computer-assisted spatial sampling may be found in Gardi et al. (2006, 2007b); Gundersen (2002).

If the weight assigned to a field of view is positively correlated with the number of cells seen in the field, the sampling is directed towards fields of view with high number of cells and an increase in efficiency is expected. In Gardi et al. (2007a), increase in efficiency ranging from  $8\times$  to  $25\times$  was indeed observed in three biological examples without increasing the workload. This finding was supported by extensive simulation studies Gardi et al. (2008), demonstrating the beneficial effect of this type of empirical importance sampling.

A number of important aspects of the statistical inference for this type of data has, however, not been addressed in Gardi et al. (2008, 2007a). The present paper fills this gap. We provide statistical methods of judging from a pilot experiment whether the estimator based on weighted sampling suggested in Gardi et al. (2007a) will indeed have an increased efficiency. Such knowledge is important for the practitioner before conducting a large-scale time-consuming microscopy study with a particular choice of weight. Furthermore, after the weighted sampling has been performed, an important issue is to construct



**Figure A.1:** Estimating total number of GFP-expressing neurones in transgenic mice brain. Panels A and B show the same region of interest at  $10\times$  magnification in bright field (panel A) and during colour identification, using fluorescence light (panel B). Counting is performed by an expert using a  $60\times$  magnification, as shown in panels C and D. The small inserts indicate the positions of the sampled windows. For more details, see Gardi et al. (2007a).

optimal estimators of intensities based on the available data. Such methods will also be developed in the present paper. Compared to the standard estimator based on weighted sampling given in Gardi et al. (2007a), the optimal estimator may have a considerably lower variance, as will be shown in the present paper by a concrete example.

The paper is organised as follows. In Section A.2, we summarise the concepts needed from point process theory and sampling theory. For more detailed accounts on these fields, cf. Møller and Waagepetersen (2004) and Thompson (1992). Section A.3 gives a condition under which the standard estimator based on weighted sampling will have an increased efficiency. In Section A.4, the gain in efficiency is assessed for homogeneous point processes, including processes from the class of Lévy driven Cox processes. Design-based inference based on systematic weighted sampling is discussed in Section A.5. Construction of optimal model-based estimators of intensities based on a specific choice of weighting is developed and exemplified by the estimation of granule cell number in Section A.6. Section A.7 contains concluding remarks. Technical details are deferred to three appendices.

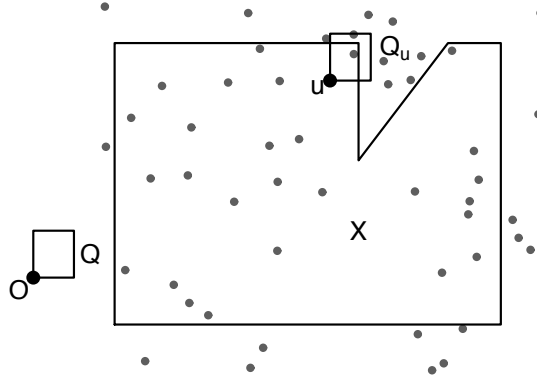
## A.2 The set-up

Throughout the paper, we will assume that a realisation of a point process  $\Phi$  is available for observation in a bounded subset  $X$  of  $\mathbb{R}^2$  of area  $A(X)$ . In addition, a non-negative random field

$$Z = \{ Z(u) : u \in X \},$$

associated with  $\Phi$ , is observed in  $X$ . In the applications we have in mind, the points of  $\Phi$  represent the positions of the objects in a digital or analog image. In the example of Figure A.1,  $X$  corresponds to the region delineated at low magnification. Determination of the total number  $N(\Phi \cap X)$  of points in  $X$  can only be performed at high magnification of the image and is impracticable. In contrast, the random field  $Z$  is readily available, e.g. from observation of colour proportions at low magnification by automatic image analysis.

Our aim is to predict  $N(\Phi \cap X)$  or equivalently  $N = N(\Phi \cap X)/A(X)$  from observation in a randomly placed window  $Q_U = Q + U$ , hitting  $X$ . Here,  $Q \subset \mathbb{R}^2$  is bounded and is assumed to contain the origin  $o$ , cf. Figure A.2. The position of the window is determined by the random vector  $U \in \mathbb{R}^2$ .



**Figure A.2:** Illustration of the stochastic set-up.

In Gardi et al. (2008, 2007a),  $\Phi$  and  $Z$  are regarded as non-random and focus is on design-unbiased predictors  $\hat{N}$  of  $N$ . A predictor  $\hat{N}$  of  $N$  is said to be design-unbiased if

$$E(\hat{N}|\Phi, Z) = N. \quad (\text{A.1})$$

The conditional mean value in (A.1) is calculated with respect to the conditional distribution of  $U$  given  $\Phi, Z$ . When  $\hat{N}$  is design-unbiased,

$$E(\hat{N}) = E(E(\hat{N}|\Phi, Z)) = E(N), \quad (\text{A.2})$$

and  $\hat{N}$  is therefore an (unconditionally) unbiased estimator of  $E(N)$ .

In the present paper, we will make statements about the quality of a predictor across different realisations of  $\Phi$  and  $Z$ . For this purpose, we will consider the prediction error  $E(\hat{N} - N)^2$ . A predictor  $\hat{N}^{(1)}$  is said to be more efficient than  $\hat{N}^{(2)}$

if  $\hat{N}^{(1)}$  has smaller prediction error than  $\hat{N}^{(2)}$ . For a design-unbiased predictor, we have  $E(\hat{N}) = E(N)$  and the prediction error can be expressed as

$$\begin{aligned} E(\hat{N} - N)^2 &= E([\hat{N} - E(\hat{N})] - [N - E(N)])^2 \\ &= \text{Var}(\hat{N}) + \text{Var}(N) - 2 \text{Cov}(\hat{N}, N) \\ &= \text{Var}(\hat{N}) - \text{Var}(N). \end{aligned} \tag{A.3}$$

### A.3 A simple condition for gain in efficiency

The standard predictor based on weighted sampling suggested in Gardi et al. (2007a) will not always result in a gain in efficiency. In this section, we will give a simple condition on  $\Phi$  and  $Z$  under which the predictor based on weighted sampling is more efficient than the one based on standard uniform random sampling. We first present the two different types of sampling considered in Gardi et al. (2008, 2007a).

In traditional sampling, the position  $u$  of the window is selected uniform randomly without any reference to the random field  $Z$ . In empirical importance sampling, the position  $u$  of the window is selected with a probability proportional to  $Z(u)$ . The two types of random windows are denoted uniform random (UR) and proportional random (PR) windows, respectively.

A UR window is distributed as  $Q_U = Q + U$  where  $U$  is independent of  $\Phi, Z$  and uniform in the set  $\bar{X}$  for which the corresponding window hits  $X$ ,

$$\bar{X} = \{u \in \mathbb{R}^2 : X \cap Q_u \neq \emptyset\}.$$

Let

$$C = A(\bar{X}) / (A(X) A(Q)). \tag{A.4}$$

The predictor

$$\hat{N}_A^{\text{UR}} = C N(\Phi \cap X \cap Q_U) \tag{A.5}$$

is a design-unbiased predictor of  $N$ . To see this, note that for any  $x \in \Phi \cap X$  we have

$$\int_{\bar{X}} \mathbb{1}_{Q+u}(x) \, du = \int_{\mathbb{R}^2} \mathbb{1}_{Q+u}(x) \, du = A(Q).$$

Therefore,

$$\begin{aligned} E[\hat{N}_A^{\text{UR}} | \Phi, Z] &= E[\hat{N}_A^{\text{UR}} | \Phi] \\ &= C \int_{\bar{X}} N(\Phi \cap X \cap (Q + u)) \frac{du}{A(\bar{X})} \\ &= \frac{1}{A(X)A(Q)} \sum_{x \in \Phi \cap X} \int_{\bar{X}} \mathbb{1}_{Q+u}(x) \, du \\ &= \frac{N(\Phi \cap X)}{A(X)} = N. \end{aligned}$$

We now consider PR windows  $Q_U$ . In order to avoid problems with edge-effects, we assume that  $Z(u)$  is defined for all  $u \in \bar{X}$ . A PR window is distributed as  $Q_U$  where  $U$  is a stochastic vector with density  $p(u|Z)$  proportional to  $Z(u)$

$$p(u|Z) = \begin{cases} \frac{Z(u)}{\int_{\bar{X}} Z(u) du}, & \text{if } u \in \bar{X}, \\ 0, & \text{otherwise.} \end{cases} \quad (\text{A.6})$$

Here and in the following, we assume that  $\int_{\bar{X}} Z(u) du > 0$  and that

$$Z(u) = 0 \quad \Rightarrow \quad N(\Phi \cap X \cap Q_u) = 0. \quad (\text{A.7})$$

The predictor suggested in Gardi et al. (2008, 2007a) takes the form

$$\hat{N}_A^{\text{PR}} = \frac{\int_{\bar{X}} Z(u) du}{A(X)A(Q)} \frac{N(\Phi \cap X \cap Q_U)}{Z(U)} \quad (\text{A.8})$$

The predictor  $\hat{N}_A^{\text{PR}}$  is of Horvitz–Thompson type, cf. Horvitz and Thompson (1952); Thompson (1992), since  $N(\Phi \cap X \cap Q_U)$  is weighted with the inverse sampling probability  $p(U|Z)^{-1}$ . The predictor  $\hat{N}_A^{\text{PR}}$  is design-unbiased for  $N$ , as can easily be seen, using the same type of reasoning as for  $\hat{N}_A^{\text{UR}}$ . The assumption (A.7) is appropriate for the applications we have in mind, cf. Gardi et al. (2007a). In the case where (A.7) does not quite hold, one may simply replace  $Z$  by  $Z + c$  where  $c$  is a small positive constant. The relative bias of  $\hat{N}_A^{\text{PR}}$

$$[\text{E}(\hat{N}_A^{\text{PR}}|\Phi, Z) - N]/N$$

is bounded by the (small) fraction of points  $x \in \Phi \cap X$  for which there exists  $u$  such that  $Z(u) = 0$  and  $x \in Q_u$ .

The predictor  $\hat{N}_A^{\text{PR}}$  is not in general more efficient than  $\hat{N}_A^{\text{UR}}$ . Using that  $\hat{N}_A^{\text{UR}}$  and  $\hat{N}_A^{\text{PR}}$  are both design-unbiased and that (A.2) and (A.3) hold for any design-unbiased predictor, it is seen that  $\hat{N}_A^{\text{PR}}$  is more efficient than  $\hat{N}_A^{\text{UR}}$  if and only if

$$\text{E}((\hat{N}_A^{\text{PR}})^2) \leq \text{E}((\hat{N}_A^{\text{UR}})^2).$$

Using that  $\text{E}((\hat{N}_A^{\text{PR}})^2) = \text{E}(\text{E}((\hat{N}_A^{\text{PR}})^2|\Phi, Z))$  and likewise for  $\hat{N}_A^{\text{UR}}$ , we find that  $\hat{N}_A^{\text{PR}}$  is more efficient than  $\hat{N}_A^{\text{UR}}$  if and only if

$$\text{E Cov}(N(\Phi \cap X \cap Q_U)^2/Z(U), Z(U)|\Phi, Z) \geq 0 \quad (\text{A.9})$$

for  $U$  uniform in  $\bar{X}$ .

Before conducting a large-scale time-consuming microscopy study or other type of study, involving weighted sampling of windows, it will be a good idea to perform a pilot study to investigate whether (A.9) is satisfied. Such a study will involve replicated observation of  $(Z(U), N(\Phi \cap X \cap Q_U))$  in windows  $Q_U$  with uniform random position  $U$  in  $\bar{X}$ . A statistical judgement of whether (A.9) is

satisfied can be performed from a plot of  $(Z(U), N(\Phi \cap X \cap Q_U)^2/Z(U))$  and an empirical estimate of the covariance in (A.9).

It is also possible to develop a model-based criterion for gain in efficiency. If the counts satisfy the following proportional regression model

$$E(N(\Phi \cap X \cap Q_U)|U, Z) = aZ(U), \quad (\text{A.10})$$

$$\text{Var}(N(\Phi \cap X \cap Q_U)|U, Z) = bZ(U)^p, \quad p \geq 1, \quad (\text{A.11})$$

then  $\hat{N}_A^{\text{PR}}$  is more efficient than  $\hat{N}_A^{\text{UR}}$ . The proof of this result is deferred to Appendix 1.

## A.4 Efficiency for homogeneous point processes

Since PR windows use the inhomogeneity of the *realised* point pattern, a gain in efficiency may be obtained even for homogeneous point processes. In this section, we provide theoretical examples of how much the efficiency can be improved. We calculate the actual gain in efficiency

$$E(\hat{N}_A^{\text{UR}} - N)^2 / E(\hat{N}_A^{\text{PR}} - N)^2$$

for various homogeneous but clustered point processes and various window sizes. The aim is to give an impression of how the magnitude of the gain in efficiency depends on the degree of inhomogeneity in the realised point pattern.

Throughout this section,  $\Phi$  will be a homogeneous point process with intensity  $\lambda$ . Then,  $E N(\Phi \cap X) = \lambda A(X)$  and its pair correlation function is translation invariant. For homogeneous  $\Phi$ , we have, cf. (Stoyan et al., 1995, (4.5.3)),

$$\text{Var}(N(\Phi \cap X)) = \lambda^2 \int_{\mathbb{R}^2} \gamma_X(y) g(y) dy + \lambda A(X) - \lambda^2 A^2(X), \quad (\text{A.12})$$

where  $g$  is the pair correlation function of  $\Phi$  and

$$\gamma_X(y) = A(X \cap (X + y))$$

is the set covariogram of  $X$ . In the following, we will, in particular, consider Cox processes. Recall that  $\Phi$  is a Cox process with driving field

$$\Lambda = \{ \Lambda(u) : u \in \mathbb{R}^2 \}$$

if, conditionally on  $\Lambda$ ,  $\Phi$  is a Poisson point process with intensity function  $\Lambda$ .

In order to assess the relative efficiency of  $\hat{N}_A^{\text{PR}}$  to  $\hat{N}_A^{\text{UR}}$  we need to express the variances in terms of second-order properties of the point process. Using (A.12) and the identity

$$\int_{\mathbb{R}^2} A(X \cap Q_u) du = A(X)A(Q), \quad (\text{A.13})$$

we get

$$\begin{aligned}
 \text{Var}(\widehat{N}_A^{\text{UR}}) &= C^2 \left( \text{Var}(\mathbb{E}[N(\Phi \cap X \cap Q_U)|U]) + \mathbb{E}(\text{Var}[N(\Phi \cap X \cap Q_U)|U]) \right) \\
 &= C^2 \lambda^2 \int_{\mathbb{R}^2} \mathbb{E}[\gamma_{X \cap Q_U}(y)] g(y) \, dy + \lambda C - \lambda^2 \\
 &= \frac{A(\bar{X})}{A(Q)^2} \frac{\lambda^2}{A(X)^2} \int_{\mathbb{R}^2} \gamma_X(y) \gamma_Q(y) g(y) \, dy + \lambda C - \lambda^2.
 \end{aligned} \tag{A.14}$$

The last equality sign follows from (A.13) with  $X$  and  $Q$  replaced by  $X \cap (X + y)$  and  $Q \cap (Q + y)$ , respectively, and the definition of  $C$ , cf. (A.4). Since, cf. (A.12),

$$\text{Var}(N) = \frac{\lambda^2}{A(X)^2} \int_{\mathbb{R}^2} \gamma_X(y) g(y) \, dy + \frac{\lambda}{A(X)} - \lambda^2, \tag{A.15}$$

the prediction error of  $\widehat{N}_A^{\text{UR}}$  is easily obtained, using (A.3),

$$\begin{aligned}
 \mathbb{E}(\widehat{N}_A^{\text{UR}} - N)^2 &= \frac{\lambda^2}{A(X)^2} \int_{\mathbb{R}^2} \gamma_X(y) g(y) \left[ \frac{A(\bar{X})}{A(Q)^2} \gamma_Q(y) - 1 \right] dy + \frac{\lambda}{A(X)} \left[ \frac{A(\bar{X})}{A(Q)} - 1 \right].
 \end{aligned} \tag{A.16}$$

The variance of  $\widehat{N}_A^{\text{PR}}$  depends on the random field  $Z$  and its interplay with  $\Phi$ . When  $\Phi$  is a Cox process with driving field  $\Lambda$  such that  $Z$  is the cumulated intensity in  $Q_u$ ,

$$Z(u) = \int_{X \cap Q_u} \Lambda(v) \, dv,$$

then, cf. Appendix 2,

$$\text{Var}(\widehat{N}_A^{\text{PR}}) = \frac{\lambda^2}{A(X)^2} \int_{\mathbb{R}^2} \gamma_X(y) g(y) \, dy + \lambda C - \lambda^2. \tag{A.17}$$

Combining with (A.15), we find, using (A.3),

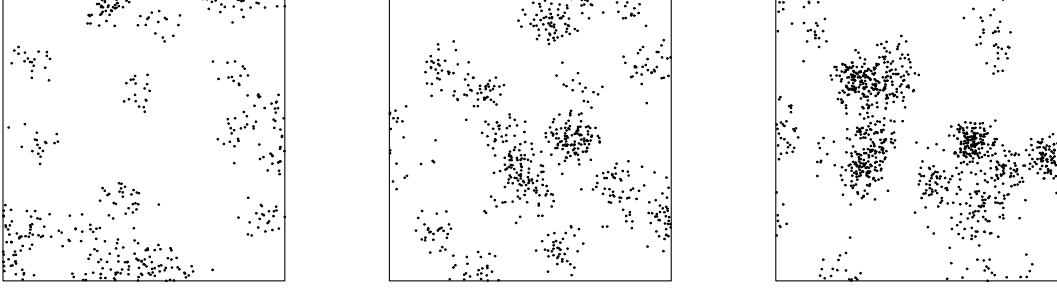
$$\mathbb{E}(\widehat{N}_A^{\text{PR}} - N)^2 = \frac{\lambda}{A(X)} \left[ \frac{A(\bar{X})}{A(Q)} - 1 \right]. \tag{A.18}$$

Note that in this case the prediction error of  $\widehat{N}_A^{\text{PR}}$  only depends on  $\lambda$ .

Using (A.16) and (A.18), we have calculated the ratio

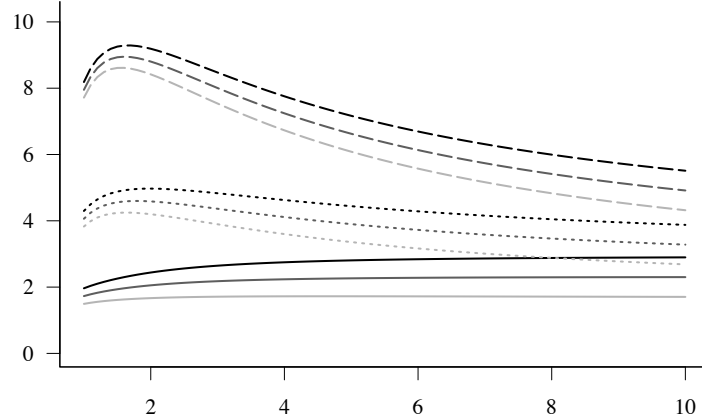
$$\mathbb{E}(\widehat{N}_A^{\text{UR}} - N)^2 \Big/ \mathbb{E}(\widehat{N}_A^{\text{PR}} - N)^2$$

for three types of Lévy driven Cox processes (Hellmund et al. (2008)). For these processes, the intensity and the pair correlation function can be determined explicitly, see (Hellmund et al., 2008, Corollary 2). Figure A.3 shows simulations in a  $[0, 10] \times [0, 10]$  window of the processes considered. Note the increasing heterogeneity in the realised point patterns. (The  $[0, 10] \times [0, 10]$  windows are just used to show realisations of the point processes and should not be confused with  $X$ .)



**Figure A.3:** Examples of realisations in a  $[0, 10] \times [0, 10]$  window of homogeneous Lévy driven Cox processes with Poisson (left), compound gamma (middle) and compound inverse Gaussian (right) Lévy bases. The driving field of the Cox processes is a kernel smoothing of the Lévy bases with a Gaussian kernel with  $\sigma = 0.4$ . Furthermore,  $c = 3$  and  $E L' = 2$ . For the definition of these parameters and further details, see Appendix 3 and (Hellmund et al., 2008, Section 4.2). The intensity  $\lambda$  is 6 in all three cases.

We have calculated the prediction error of  $\hat{N}_A^{\text{UR}}$  and  $\hat{N}_A^{\text{PR}}$  for the case where  $Q$  is a square with unit side length while  $X$  is a square with side length  $D \geq 1$ . For some technical details, see Appendix 3. In Figure A.4, we have plotted the ratio  $E(\hat{N}_A^{\text{UR}} - N)^2 / E(\hat{N}_A^{\text{PR}} - N)^2$  as a function of  $D \in [1, 10]$  for Lévy driven Cox processes with Poisson, compound gamma and compound inverse Gaussian Lévy bases and  $\lambda = 1, 10, 25$ . The gain in efficiency is largest for the most pronounced clustered point patterns. The efficiency also increases with increasing intensity of the point process. The gain in efficiency is at least a factor 2 in the majority of cases studied. Note that, for a specific point process model and intensity, the gain is largest when the size of  $X$  is small because in this case the windows  $Q_u$  more often only partly hits  $X$  and, accordingly, the parts of  $\Phi \cap X$  seen in the windows appear more inhomogeneous.



**Figure A.4:** The ratio  $E(\hat{N}_A^{\text{UR}} - N)^2 / E(\hat{N}_A^{\text{PR}} - N)^2$  as a function of the side length  $D$  of  $X$  for the Poisson (red), compound gamma (green) and compound inverse Gaussian (blue) cases, respectively. The window  $Q$  is a square of unit length. The intensities are  $\lambda = 1$  (full-drawn),  $\lambda = 10$  ( $\cdot \cdot \cdot$ ) and  $\lambda = 25$  ( $- - -$ ), respectively.



## A.5 Design-based inference

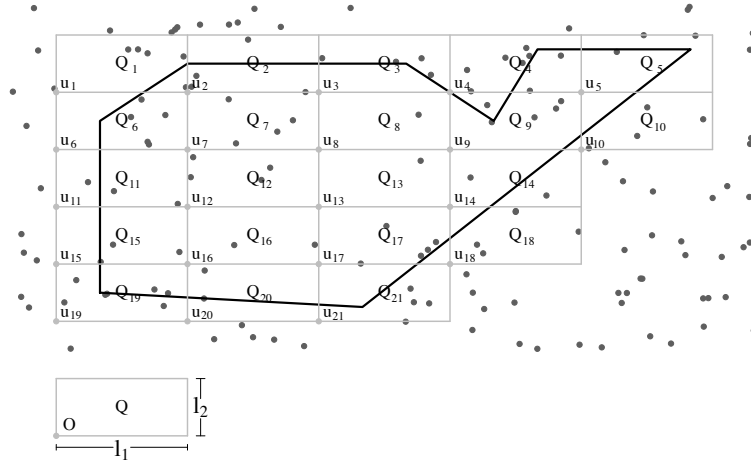
Replicated generation of PR windows in a systematic set-up has in Gardi et al. (2008, 2007a) been implemented for the analysis of microscopy sections as the one shown in Figure A.1. The sampling is denoted systematic proportional random sampling (SPRS). Below, we give SPRS sampling a probabilistic description and present the standard design-unbiased predictor suggested in Gardi et al. (2007a) based on SPRS sampling. Taking a model-based approach, predictors with considerably lower variance may be constructed, as we shall see in the section to follow.

The window  $Q = [0, l_1) \times [0, l_2)$  is a rectangle. The region  $X$  is covered by rectangles

$$G = \bigcup_{(s_1, s_2) \in \mathcal{S}} \{Q + (s_1 l_1, s_2 l_2)\},$$

where  $\mathcal{S} \subset \mathbb{Z}^2$ . The windows in  $G$  are ordered lexicographically  $Q_1, \dots, Q_N$ , where  $N$  is the number of windows in  $G$ . The windows  $Q_1, \dots, Q_N$  are translations of  $Q$  by  $u_1, \dots, u_N$ , say, see Figure A.5. The weight  $Z(u_i)$  is assigned to  $Q_i$ , for  $i = 1, \dots, N$ .

In the applications, we have in mind,  $Z(u_i)$  is a total quantity for the window  $Q_i (= Q_{u_i})$ . For instance, for the example illustrated in Figure A.1 where the number of green GFP-expressing neurones is to be estimated,  $Z(u_i)$  is the total amount of green colour observed in  $Q_i$  under fluorescence illumination.



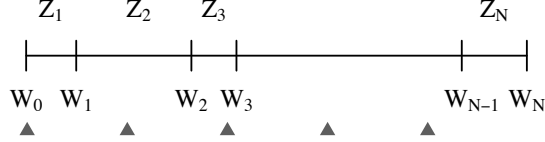
**Figure A.5:** The set-up for systematic proportional random sampling (SPRS). The point pattern  $\Phi$  (red dots) is available for observation in  $X$  (delineated by blue lines). The set  $X$  is covered by a family of lexicographically ordered non-overlapping rectangles. These rectangles are translations of the given rectangle  $Q$ .

Sampling in a systematic  $Z$ -weighted fashion is performed as follows. Let  $W_j = \sum_{i=1}^j Z(u_i)$  denote the cumulated weight, with the convention that  $W_0 = 0$ , and let  $S = \sum_{i=1}^N Z(u_i)$ . A sample of  $n \in \{1, \dots, N\}$  windows is

selected by choosing  $V_1$  uniformly in  $[0, S/n]$ , independently of  $\Phi$  and  $Z$ , and let  $V_j = V_1 + (j-1)S/n$  for  $j = 2, 3, \dots, n$ . The sampled windows are those with index in

$$J = \bigcup_{i: V_i \in [W_{j_i-1}, W_{j_i}]} \{j_i\},$$

see also Figure A.6. A window may be sampled more than once. Notice that ordinary systematic uniform random sampling (SURS) is a special case of SPRS where  $Z$  is a constant field.



**Figure A.6:** Illustration of SPRS sampling. The 1st and 3rd window are not sampled, the  $N$ th window is sampled exactly once while the 2nd window is sampled twice. We use the abbreviation  $Z_j = Z(u_j)$ .

In Gardi et al. (2007a), it is suggested to use the following predictor of  $N$

$$\hat{N}_A^{\text{SPRS}} = \frac{S}{n A(X)} \sum_{j=1}^N \#\{i \mid V_i \in [W_{j-1}, W_j]\} \frac{N(\Phi \cap X \cap Q_j)}{Z(u_j)}.$$

This predictor is design-unbiased. In fact, since  $V_i \sim U([(i-1)S/n, iS/n])$ ,

$$\begin{aligned} \mathbb{E}[\#\{i \mid V_i \in [W_{j-1}, W_j]\} \mid \Phi, Z] &= \sum_{i=1}^n \mathbb{P}(V_i \in [W_{j-1}, W_j] \mid \Phi, Z) \\ &= \frac{n}{S} \sum_{i=1}^n \int_{(i-1)\frac{S}{n}}^{i\frac{S}{n}} \mathbb{1}_{[W_{j-1}, W_j]}(v) \, dv \\ &= \frac{n}{S} (W_j - W_{j-1}) \\ &= \frac{n}{S} Z(u_j), \end{aligned}$$

which gives us

$$\mathbb{E}[\hat{N}_A^{\text{SPRS}} \mid \Phi, Z] = A(X)^{-1} \sum_{j=1}^N N(\Phi \cap X \cap Q_j) = N(\Phi \cap X) / A(X).$$

## A.6 Model-based inference

After having performed SPRS sampling with a particular choice of weight, an important issue is to construct optimal model-based predictors of  $N$  based on the available data

$$(Z(u_j), N(\Phi \cap X \cap Q_j)), \quad j \in J.$$

For an introduction to model-based sampling theory, cf. e.g. (Thompson, 1992, Section 2.7 and Chapter 8).

Under a proportional regression model

$$\mathbb{E}(N(\Phi \cap X \cap Q_j) | J, Z) = aZ(u_j), \quad (\text{A.19})$$

$$\text{Var}(N(\Phi \cap X \cap Q_j) | J, Z) = bZ(u_j)^p, \quad (\text{A.20})$$

$j \in J$ , we will construct a predictor with minimal model-variance among predictors of the form

$$\hat{N} = \frac{S}{A(X)} \sum_{j \in J} \alpha_j(J, Z) \frac{N(\Phi \cap X \cap Q_j)}{Z(u_j)}, \quad (\text{A.21})$$

where  $\alpha_j(J, Z) \geq 0$  and  $\sum_{j \in J} \alpha_j(J, Z) = 1$ , provided that the counts

$$N(\Phi \cap X \cap Q_j), \quad j \in J,$$

can be regarded as uncorrelated, given  $J$  and  $Z$ . This assumption may be checked empirically by analysing the spatial correlation of the residuals.

First, notice that for a predictor of the form (A.21)

$$\mathbb{E}(\hat{N} | J) = \frac{a}{A(X)} \mathbb{E}(S).$$

Furthermore, because of (A.19),

$$\mathbb{E}(N) = A(X)^{-1} \mathbb{E} N(\Phi \cap X) = A(X)^{-1} \sum_{j=1}^N \mathbb{E} N(\Phi \cap X \cap Q_j) = \frac{a}{A(X)} \mathbb{E}(S).$$

It follows that  $\mathbb{E}(\hat{N} | J) = \mathbb{E}(N)$ , so any predictor of the form (A.21) is indeed model-unbiased. Using the assumption of uncorrelatedness, the model-variance becomes

$$\begin{aligned} \text{Var}(\hat{N} | J) &= \mathbb{E}(\text{Var}(\hat{N} | J, Z)) + \text{Var}(\mathbb{E}(\hat{N} | J, Z)) \\ &= \frac{b}{A(X)^2} \mathbb{E}(S^2 \sum_{j \in J} \alpha_j(J, Z)^2 Z(u_j)^{p-2}) + \frac{a^2}{A(X)^2} \text{Var}(S). \end{aligned}$$

Using Cauchy–Schwartz’ inequality it follows that  $\text{Var}(\hat{N} | J)$  is minimised for

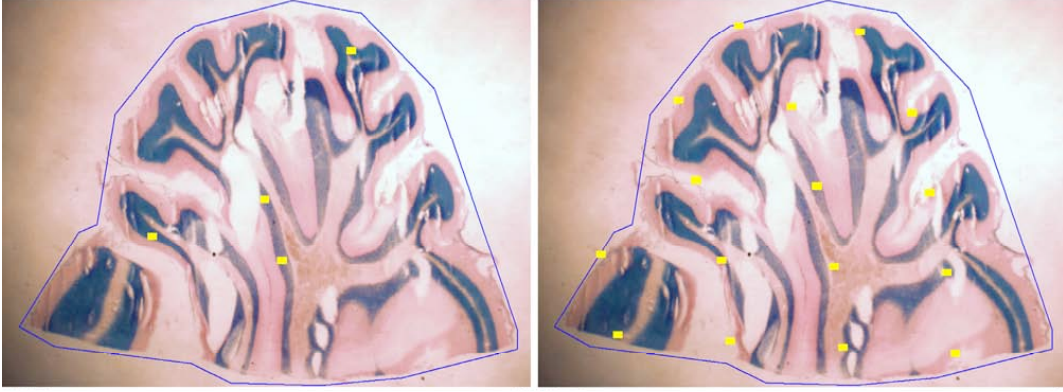
$$\alpha_j(J, Z) = Z(u_j)^{2-p} / \sum_{j \in J} Z(u_j)^{2-p}.$$

The minimal variance becomes

$$\text{Var}(\hat{N} | J) = \frac{b}{A(X)^2} \mathbb{E}\left(\frac{S^2}{\sum_{j \in J} Z(u_j)^{2-p}}\right) + \frac{a^2}{A(X)^2} \text{Var}(S).$$

If  $p = 2$  and the sampling is performed such that a window is only sampled once, then the minimal model-variance predictor of the form (A.21) coincides with  $\hat{N}_A^{\text{SPRS}}$ . In other cases, the minimal model-variance predictor will be superior to  $\hat{N}_A^{\text{SPRS}}$  in terms of increased efficiency.

We have applied this type of model-based inference on data collected by SPRS sampling for the estimation of the intensity of granule cells in rat cerebellum. SPRS sampling of the granule cell layer (blue) is shown in Figure A.7 (left) where almost all selected sampling windows (small yellow rectangles) lie in the granule layer. For comparison, Figure A.7 (right) shows traditional systematic uniform random sampling where many of the sampling windows fall outside the blue region of interest. A total of 124 windows was selected by SPRS sampling on 14 sections, as the one shown in Figure A.7. For more details, see Gardi et al. (2007a).



**Figure A.7:** Left: A systematic proportional random sample (SPRS) of the granule cell layer (blue). The selected sampling windows are shown as small yellow rectangles. Right: A systematic uniform random sample (SURS).

A plot of the observed counts versus weights shows that a proportional regression model is not appropriate. Mean counts are shown in Figure A.8 as a function of grouped weights. A relationship of the type

$$E(N(\Phi \cap X \cap Q_j) | J, Z) = a_1 Z(u_j)^{a_2}$$

with  $a_2 \neq 1$  is more appropriate. The fitted curve in Figure A.8 (left) has parameters  $a_1 = 1$  and  $a_2 = 0.595$ . We therefore transform the weights

$$\tilde{Z}(u_j) = Z(u_j)^{a_2},$$

such that the proportional regression is fulfilled for the transformed weights. Figure A.8 (right) shows the empirical variances of the counts as a function of the transformed weights. A relationship of the type (A.20) with  $p = 0$  seems to be appropriate. Under this type of model, the optimal predictor takes the form

$$\hat{N}_A^{\text{opt}} = \frac{\sum_{i=1}^N \tilde{Z}(u_i)}{A(X)} \frac{\sum_{j \in J} \tilde{Z}(u_j) N(\Phi \cap X \cap Q_j)}{\sum_{j \in J} \tilde{Z}(u_j)^2},$$

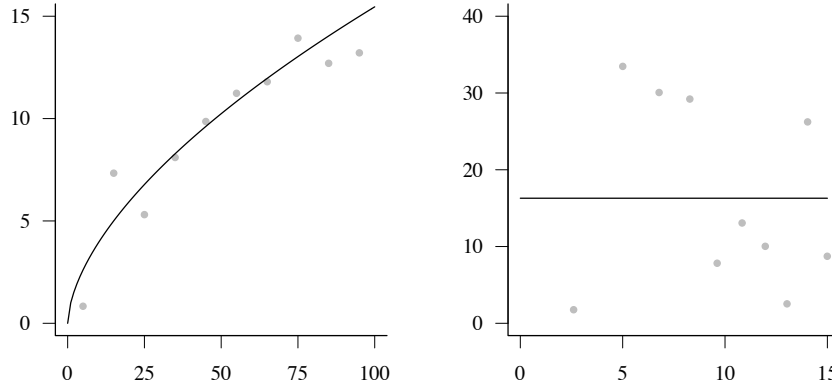
while the predictor suggested in Gardi et al. (2007a) based on the transformed weights becomes (assuming no window is sampled more than once)

$$\hat{N}_A^{\text{SPRS}} = \frac{\sum_{i=1}^N \tilde{Z}(u_i)}{nA(X)} \sum_{j \in J} N(\Phi \cap X \cap Q_j) / \tilde{Z}(u_j).$$

The ratio between the conditional variances is for  $p = 0$

$$\frac{\text{Var}(\hat{N}_A^{\text{SPRS}} | J, Z)}{\text{Var}(\hat{N}_A^{\text{opt}} | J, Z)} = \left( \frac{1}{n} \sum_{j \in J} \tilde{Z}(u_j)^2 \right) \left( \frac{1}{n} \sum_{j \in J} \tilde{Z}(u_j)^{-2} \right).$$

In the concrete example, this ratio takes the value 2.90 and the optimal model-based predictor  $\hat{N}_A^{\text{opt}}$  represents an increase in efficiency of a factor 3 compared to the standard predictor  $\hat{N}_A^{\text{SPRS}}$  suggested in Gardi et al. (2007a).



**Figure A.8:** Left: Mean counts as a function of grouped weights together with the fitted relationship. Right: Empirical variances of counts as a function of grouped transformed weights  $\tilde{Z}$ . For details, see text.

In the example discussed in this section, the proportional regression model (A.19)–(A.20) was satisfied after a transformation of the weights. Likewise, a fulfilment of the criterion (A.9) presented in Section A.3 may require a transformation of the weights.

## A.7 Concluding remarks

In this paper, we have studied how computerised image analysis can be combined with non-uniform sampling in order to increase the efficiency of estimators of intensities of biological cell populations. We have provided conditions under which the proposed non-uniform sampling results in a gain in efficiency and constructed optimal model-based estimators of intensities.

We believe that the principle of empirical importance sampling has a much wider range of applications than in microscopy. It is likely to be useful in other areas of spatial sampling where point patterns show realised inhomogeneity, e.g.

in precision farming and satellite image analysis. It might also be the solution to the problem of low resolution in modern MR and PET scanners. An initial complete scan at low resolution may be used to direct the sampling towards the region of interest which is subsequently scanned at a high resolution.

## A.8 Acknowledgements

This work has been supported by the Danish Council for Strategic Research and Centre for Stochastic Geometry and Advanced Bioimaging, funded by a grant from the Villum Foundation.

## Bibliography

- Asmussen, S. and Glynn, P. W. (2007). *Stochastic Simulation*. Springer Science+Business Media, LLC.
- Dorph-Petersen, K.-A., Gundersen, H. J. G., and Jensen, E. B. V. (2000). Non-uniform systematic sampling in stereology. *Journal of Microscopy*, 200(2):148–157.
- Gardi, J., Nyengaard, J., and Gundersen, H. (2008). The proportionator: Unbiased stereological estimation using biased automatic image analysis and non-uniform probability proportional to size sampling. *Computers in Biology and Medicine*, 38(3):313–328.
- Gardi, J. E., Nyengaard, J. R., and Gundersen, H. J. G. (2006). Using biased image analysis for improving unbiased stereological number estimation - a pilot simulation study of the smooth fractionator. *Journal of Microscopy*, 222(3):242–250.
- Gardi, J. E., Nyengaard, J. R., and Gundersen, H. J. G. (2007a). Automatic sampling for unbiased and efficient stereological estimation using the proportionator in biological studies. *Journal of Microscopy*, 230:108–120.
- Gardi, J. E., Wulfsohn, D., and Nyengaard, J. R. (2007b). A handheld support to facilitate stereological measurements and mapping of branching structures. *Journal of Microscopy*, 227:127–139.
- Gundersen, H. J. G. (2002). The smooth fractionator. *Journal of Microscopy*, 207(3):191–210.
- Hellmund, G., Prokešová, M., and Jensen, E. B. V. (2008). Lévy based cox point processes. *Advances in Applied Probability*, 40:603–629.
- Horvitz, D. G. and Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47:663–685.
- Møller, J. and Waagepetersen, R. P. (2004). *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC, Boca Raton.
- Stoyan, D., Kendall, W. S., and Mecke, J. (1995). *Stochastic Geometry and Its Applications*. Akademie-Verlag, 2nd edition.
- Thompson, S. (1992). *Sampling*. John Wiley & Sons.

## Appendix 1

In this appendix, we show that  $\hat{N}_A^{\text{PR}}$  is more efficient than  $\hat{N}_A^{\text{UR}}$  if the proportional regression model given by (A.10) and (A.11) is satisfied. Since  $\hat{N}_A^{\text{UR}}$  and  $\hat{N}_A^{\text{PR}}$  are both design-unbiased, it suffices to show that

$$\text{Var}(\hat{N}_A^{\text{PR}}) \leq \text{Var}(\hat{N}_A^{\text{UR}}). \quad (\text{A.22})$$

Throughout this appendix,  $U$  will denote a uniform random vector on  $\bar{X}$  and  $W$  a random vector with density (A.6). We show (A.22) by showing the following two relations

$$\text{Var}(\mathbb{E}(\hat{N}_A^{\text{PR}} \mid W, Z)) \leq \text{Var}(\mathbb{E}(\hat{N}_A^{\text{UR}} \mid U, Z)) \quad (\text{A.23})$$

$$\mathbb{E}(\text{Var}(\hat{N}_A^{\text{PR}} \mid W, Z)) \leq \mathbb{E}(\text{Var}(\hat{N}_A^{\text{UR}} \mid U, Z)). \quad (\text{A.24})$$

First we show (A.23). Since  $\hat{N}_A^{\text{UR}}$  and  $\hat{N}_A^{\text{PR}}$  are both design-unbiased, it suffices to show that

$$\mathbb{E}(\mathbb{E}^2(\hat{N}_A^{\text{PR}} \mid W, Z)) \leq \mathbb{E}(\mathbb{E}^2(\hat{N}_A^{\text{UR}} \mid U, Z)).$$

We find

$$\mathbb{E}(\hat{N}_A^{\text{PR}} \mid W, Z) = Ca \int_{\bar{X}} Z(u) \frac{du}{A(\bar{X})} = Ca \mathbb{E}(Z(U) \mid Z),$$

and

$$\mathbb{E}(\hat{N}_A^{\text{UR}} \mid U, Z) = Ca Z(U).$$

Since

$$\mathbb{E}(\mathbb{E}^2(Z(U) \mid Z)) \leq \mathbb{E}(\mathbb{E}(Z(U)^2 \mid Z))$$

inequality (A.23) follows.

In order to show (A.24), we use (A.11) and find

$$\begin{aligned} \text{Var}(\hat{N}_A^{\text{UR}} \mid U, Z) &= C^2 b Z(U)^p, \\ \text{Var}(\hat{N}_A^{\text{PR}} \mid W, Z) &= \left( \frac{\int_{\bar{X}} Z(u) du}{A(X)A(Q)} \right)^2 b Z(W)^{p-2}. \end{aligned}$$

Using that

$$\mathbb{E}(Y^p) \geq \mathbb{E}(Y) \mathbb{E}(Y^{p-1}), \quad p \geq 1, \quad (\text{A.25})$$

we finally get

$$\begin{aligned} \mathbb{E}(\text{Var}(\hat{N}_A^{\text{PR}} \mid W, Z)) &= \mathbb{E} \left[ \left( \int_{\bar{X}} Z(u) du / (A(X)A(Q)) \right)^2 b \mathbb{E}(Z(W)^{p-2} \mid Z) \right] \\ &= C^2 b \mathbb{E} \left[ \left( \int_{\bar{X}} Z(u) \frac{du}{A(\bar{X})} \right)^2 \int_{\bar{X}} [Z(u)^{p-1} / \int_{\bar{X}} Z(u) du] du \right] \\ &= C^2 b \mathbb{E}[\mathbb{E}(Z(U) \mid Z) \mathbb{E}(Z(U)^{p-1} \mid Z)] \\ &\leq C^2 b \mathbb{E}[\mathbb{E}(Z(U)^p \mid Z)] \\ &= \mathbb{E}(\text{Var}(\hat{N}_A^{\text{UR}} \mid U, Z)). \end{aligned}$$

The relation (A.25) can be shown as follows. For  $p, q \geq 0$ , we have

$$\begin{aligned} \mathbb{E}(Y^q) \mathbb{E}(Y^p) &= ((\mathbb{E}(Y^q))^{1/q})^q ((\mathbb{E}(Y^p))^{1/p})^p \\ &\leq (\mathbb{E}(Y^{p+q}))^{q/(p+q)} (\mathbb{E}(Y^{p+q}))^{p/(p+q)} \\ &= \mathbb{E}(Y^{p+q}), \end{aligned}$$

using that  $p \mapsto (\mathbb{E}(Y^p))^{1/p}$  is increasing for  $p \geq 0$ , due to Jensen's inequality. Replacing  $q$  and  $p$  by 1 and  $p - 1$ , respectively, we have shown (A.25).

## Appendix 2

In this appendix, we show (A.17). Note that, given  $U$  and  $\Lambda$ ,  $N(\Phi \cap X \cap Q_U)$  is Poisson distributed with mean

$$Z(U) = \int_{X \cap Q_U} \Lambda(v) \, dv.$$

It follows that

$$\mathbb{E}(\hat{N}_A^{\text{PR}} \mid U, \Lambda) = \int_{\bar{X}} Z(u) \, du \Big/ (A(X)A(Q)) = \int_X \Lambda(v) \frac{dv}{A(X)}$$

and

$$\text{Var}(\hat{N}_A^{\text{PR}} \mid U, \Lambda) = \left( \int_X \Lambda(v) \frac{dv}{A(X)} \right)^2 \frac{1}{Z(U)}.$$

Therefore, since  $\mathbb{E} \Lambda(v) = \lambda$ ,

$$\mathbb{E}(\text{Var}(\hat{N}_A^{\text{PR}} \mid U, \Lambda)) = \frac{A(\bar{X})}{A(X)^2 A(Q)} \mathbb{E} \left( \int_X \Lambda(v) \, dv \right) = C\lambda. \quad (\text{A.26})$$

Furthermore, since  $\mathbb{E}(\Lambda(v_1)\Lambda(v_2)) = g(v_1 - v_2)\lambda^2$ ,

$$\begin{aligned} \text{Var}(\mathbb{E}(\hat{N}_A^{\text{PR}} \mid U, \Lambda)) &= \mathbb{E} \left( \frac{1}{A(X)^2} \int_X \int_X \Lambda(v_1)\Lambda(v_2) \, dv_1 \, dv_2 \right) - \lambda^2 \\ &= \frac{\lambda^2}{A(X)^2} \int_X \int_{X-v} g(u) \, du \, dv - \lambda^2 \\ &= \frac{\lambda^2}{A(X)^2} \int_{\mathbb{R}^2} \gamma_X(u) g(u) \, du - \lambda^2. \end{aligned} \quad (\text{A.27})$$

Combining (A.26) and (A.27), we finally get (A.17).

## Appendix 3

In this appendix, we will derive an explicit expression for  $\mathbb{E}(\hat{N}_A^{\text{UR}} - N)^2$  for the Lévy driven Cox processes in Figure A.3.



By (Hellmund et al., 2008, Corollary 2), one can associate a random variable  $L'$ , the so-called spot-variable, to a Lévy driven Cox process such that

$$\lambda = c \mathbb{E} L',$$

$$g(y) = 1 + \frac{\text{Var}(L')}{(\mathbb{E} L')^2} \frac{I_k(y)}{c},$$

where  $I_k(y) = \int_{\mathbb{R}^2} k(y+n)k(n) \, dn$  and  $c > 0$  is an explicitly known constant. For a Gaussian kernel

$$k(y) = \frac{1}{2\pi\sigma^2} \exp(-\|y\|^2/(2\sigma^2)), \quad y \in \mathbb{R}^2, \quad (\text{A.28})$$

we get

$$g(y) = 1 + \frac{\text{Var}(L')}{\mathbb{E} L'} \frac{\exp(-\|y\|^2/(4\sigma^2))}{4\pi\sigma^2\lambda}, \quad y \in \mathbb{R}^2. \quad (\text{A.29})$$

Throughout this appendix,  $D \geq 1$ . For  $Q = [0, 1]^2$  and  $X = [0, D]^2$ , we have

$$A(Q) = 1, \quad A(X) = D^2, \quad A(\bar{X}) = (D+1)^2.$$

Furthermore,

$$\gamma_X(y_1, y_2) = \begin{cases} (D - |y_1|)(D - |y_2|) & \text{if } (y_1, y_2) \in [-D, D]^2 \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\gamma_Q(y_1, y_2) = \begin{cases} (1 - |y_1|)(1 - |y_2|) & \text{if } (y_1, y_2) \in [-1, 1]^2 \\ 0 & \text{otherwise.} \end{cases}$$

For the calculation of the prediction error of  $\hat{N}_A^{\text{UR}}$  we use that

$$\int_{\mathbb{R}^2} \gamma_X(y) \gamma_Q(y) \, dy = \left( \int_{-1}^1 (D - |y_1|)(1 - |y_1|) \, dy_1 \right)^2 = \left( D - \frac{1}{3} \right)^2,$$

$$\begin{aligned} & \int_{\mathbb{R}^2} \gamma_X(y) \gamma_Q(y) \exp\{-\|y\|^2/(4\sigma^2)\} \, dy \\ &= \left( \int_{-1}^1 (D - |y_1|)(1 - |y_1|) \exp\{-y_1^2/(4\sigma^2)\} \, dy_1 \right)^2 \\ &= 4 \left( D\sqrt{\pi\sigma^2} [2\Phi(\frac{1}{\sqrt{2\sigma^2}}) - 1] + 2\sigma^2(D+1)[e^{(4\sigma^2)^{-1}} - 1] + \int_0^1 y^2 e^{\frac{-y^2}{4\sigma^2}} \, dy \right)^2. \end{aligned}$$

Furthermore,

$$\int_{\mathbb{R}^2} \gamma_X(y) \, dy = \left( \int_{-D}^D (D - |y_1|) \, dy_1 \right)^2 = D^4,$$

$$\begin{aligned}
\int_{\mathbb{R}^2} \gamma_X(y) \exp\left\{\frac{-\|y\|^2}{4\sigma^2}\right\} dy &= \left( \int_{-D}^D (D - |y_1|) \exp\left\{\frac{-y_1^2}{4\sigma^2}\right\} dy_1 \right)^2 \\
&= 4 \left( D\sqrt{\pi\sigma^2} \left[ 2\Phi\left(\frac{D}{\sqrt{2\sigma^2}}\right) - 1 \right] + 2\sigma^2 \left[ \exp\left\{\frac{-D^2}{4\sigma^2}\right\} - 1 \right] \right)^2.
\end{aligned}$$

Using (A.16) and (A.29), we finally get

$$\begin{aligned}
&\mathbb{E}(\widehat{N}_A^{\text{UR}} - \widehat{N})^2 \\
&= \lambda(1 + 2/D) - \lambda^2 + \lambda^2((D+1)^2(D-1/3)^2) / D^4 \\
&\quad - \lambda \frac{\text{Var}(L')}{D^4 \pi \sigma^2 \mathbb{E} L'} \left( D\sqrt{\pi\sigma^2} \left[ 2\Phi\left(\frac{D}{\sqrt{2\sigma^2}}\right) - 1 \right] + 2\sigma^2 \left[ \exp\left\{\frac{-D^2}{4\sigma^2}\right\} - 1 \right] \right)^2 \\
&\quad + \lambda \frac{(D+1)^2 \text{Var}(L')}{D^4 \pi \sigma^2 \mathbb{E} L'} \left( D\sqrt{\pi\sigma^2} \left[ 2\Phi\left(\frac{1}{\sqrt{2\sigma^2}}\right) - 1 \right] + 2\sigma^2(D+1) \left[ e^{\frac{-1}{4\sigma^2}} - 1 \right] \right. \\
&\quad \left. + \int_0^1 y^2 \exp\left\{\frac{-y^2}{4\sigma^2}\right\} dy \right)^2.
\end{aligned}$$

# The semi-automatic nucleator

Linda Vadgård Hansen<sup>1</sup>, Jens R. Nyengaard<sup>2,3</sup>, Johnnie Bremholm Andersen<sup>3</sup>  
and Eva B. Vedel Jensen<sup>1,2</sup>

*<sup>1</sup>Department of Mathematical Sciences, Aarhus University; <sup>2</sup>Centre for Stochastic Geometry and Advanced Bioimaging, Aarhus University; <sup>3</sup>Stereology and EM Laboratory, Aarhus University*

**Abstract:** The nucleator is a well-established manual stereological method of estimating mean cell volume from observations on random cell transects through reference points of the cells. In this paper, we present an automated version of the nucleator that uses automatic segmentation of the boundaries of the cell transects. An expert supervises the process. If the segmentation is judged to be satisfactory, an estimate of the cell volume is calculated automatically on the basis of the whole cell transect. In the remaining cases, the expert intervenes and uses the classical nucleator. The resulting estimator is called the semi-automatic nucleator. In the present paper, we study the statistical properties of the semi-automatic nucleator. Formulae for the bias and mean square error are derived. The semi-automatic nucleator may have a small bias but will still in most cases be more efficient than the classical nucleator. Procedures for estimating bias and mean square error from a pilot study are provided. The application of the semi-automatic nucleator is illustrated in a study of somatostatin positive inhibitory interneurons which were genetically labelled with green fluorescent protein (GFP). The cells were sampled with an optical disector. The centre of mass in a central cell transect was used as reference point. It is found in this study that the number of cells needed for obtaining, for instance, a 5 % precision of the estimate of mean cell volume is 150 and 189 for the semi-automatic and the classical nucleator, respectively. Taking into account that the time spent analysing one cell is shorter for the semi-automatic nucleator than for the classical nucleator, the semi-automatic nucleator is superior to the classical nucleator.

**Key words:** computerised image analysis, local stereology, nucleator, volume.



## B.1 Introduction

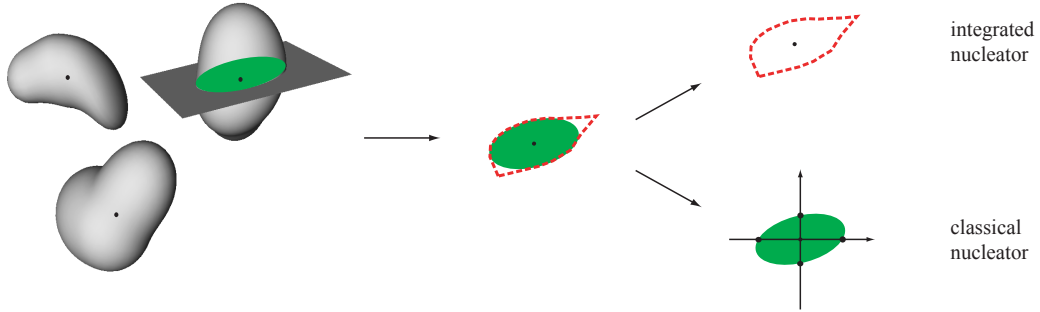
By means of local stereology, it is possible to determine the size of an object from random sections through a reference point (Jensen, 1998). Local stereological techniques can be applied without specific assumptions about the shape of the object, which is an important advantage compared to earlier methods, depending on shape assumptions such as spherical or ellipsoidal shape (Wicksell, 1925, 1926; Cruz-Orive, 1976, 1978). The local stereological methods do not have these severe shape restrictions, but local stereological estimators may have a high variability if the object is far from being spherical and/or the reference point is not centrally positioned within the object. In case the reference point is very far from being centrally positioned, the recent pivotal estimators based on the invariator principle are to be preferred, see Cruz-Orive (2005) and the accompanying papers Cruz-Orive (2008) and Cruz-Orive et al. (2010).

The nucleator in its original form, the 'classical' nucleator (Gundersen, 1988), is used for estimating cell volumes from observations in thick transparent sections. Cells are sampled when their reference point comes into focus. In the case of isotropic sections, two perpendicular lines are normally used on a sampled cell transect and the expert indicates by the computer mouse the four (or perhaps more) intersection points between the lines and the cell transect boundary, see the lower right illustration of Figure B.1. It is the expert that decides the position of the intersection points.

If it is possible by automatic image analysis to identify the boundaries of the cell transects, it is more powerful to use the information available in the whole cell transect than the information from two lines. There exists such an estimator of cell volume that uses the whole cell transect. The theory of this estimator has been known for quite long, see Jensen (2000) and references therein, but the estimator has not been in common use because automatic identification of the transect boundaries has not earlier appeared to be a realistic possibility. It can be shown that if the classical nucleator estimate is calculated on the basis of an increasing number of lines, then the estimate will come closer and closer to the value obtained when the whole cell transect is used directly. A natural name for the estimator that uses the whole cell transect is, therefore, *the integrated nucleator*.

It is important to know how much more precise the integrated nucleator is compared to the classical nucleator based on measurements along two lines. In fact, the classical nucleator is already quite precise if the reference point is centrally positioned because of the antithetic effect of the two perpendicular lines. The gain in precision by using the integrated nucleator depends on the shape of the cells. If the cells are of perfect spherical shape with the centres as reference points, there is no gain. A simulation study in Jensen (2000) showed that if the cells are prolate ellipsoids with centres as reference points and ratio between major and minor axis equal to 2, then the CE of the integrated nucleator will be 83 % of that of the classical nucleator. If the cells are prolate ellipsoids with ratio equal to 4, then the percent will be 64. Furthermore, using the

integrated nucleator, it is not needed to spend time indicating the position of the intersection points between the lines and the cell boundary. All this is under the assumption that the automatic identification of the cell transect boundaries is correct.



**Figure B.1:** Illustration of the semi-automatic nucleator. The true cell transect is shown as the green set while the segmented cell transect is delineated by the red broken curve. The integrated nucleator is used if the segmentation is judged to be satisfactory, otherwise the classical nucleator is used.

In the present paper, we study the performance of an intermediate option between the classical and the integrated nucleator: *the semi-automatic nucleator* where the expert supervises the process, see Figure B.1. The first step of the semi-automatic nucleator is an automatic identification of the cell transect boundary but it is not a requirement that the identification is correct. If the expert judges that the identification of the cell transect boundary is satisfactory, then the integrated nucleator based on the automatically segmented cell transect is used. If instead the identification of the cell transect boundary is judged unsatisfactory, the expert intervenes and indicates by the mouse the four (or more) intersection points between the lines and the real cell transect boundary. The semi-automatic nucleator may have a small bias but will, as we shall see, still in most cases be more efficient than the classical nucleator.

In Section B.2 we introduce the various estimators of volume mentioned above. As not all estimators are guaranteed unbiased we discuss mean square error as well as variance relations. In Section B.3 we compare the various estimators in a study of somatostatin positive inhibitory interneurons from mice hippocampi, observed by optical fluorescent microscopy. Further aspects are discussed in Section B.4. Some mathematical derivations are deferred to an Appendix.

## B.2 Theoretical background

In the following we describe various procedures for estimating mean particle volume. They all consist of a sampling step followed by a measurement step and use a reference point associated with each particle in the sample.

Throughout this paper,  $Y$  will denote a random particle with the origin  $o$  as reference point, i.e.  $Y$  is a compact subset of  $\mathbb{R}^3$  containing  $o$ . The aim is to estimate the mean particle volume  $\mu = \mathbb{E} V(Y)$ . For example, if the particle population under study consists of  $N$  particles,  $Y_1, Y_2, \dots, Y_N$ , and  $Y$  is selected uniform randomly among the  $N$  particles, then

$$\mu = \mathbb{E} V(Y) = \frac{1}{N} \sum_{n=1}^N V(Y_n).$$

The estimators to be considered use an isotropic plane  $L_2$  through  $o$ . The normal vector of such a random plane is uniformly distributed on a unit half-sphere. Furthermore, we will let  $L_1$  be a line in  $L_2$  through  $o$ . Every such line can be uniquely determined by the angle  $\theta \in [0, \pi)$  it generates with a fixed axis within  $L_2$ . When we need to be specific this line is denoted  $L_1(\theta)$ .

### B.2.1 The estimators

Several estimators of volume will be introduced below. For easy reference they are all listed in Table B.1.

**Table B.1:** List of volume estimators (notation and name).

$m_{\text{int}}$	integrated nucleator
$m_{\text{aut}}$	automatic nucleator
$m_{\text{cl}_1}$	classical nucleator using one line
$m_{\text{cl}_2}$	classical nucleator using two lines
$m_{\text{semi}}$	semi-automatic nucleator

#### B.2.1.1 The classical nucleator

The classical nucleator  $m_{\text{cl}_1}$  uses measurements along one isotropic line  $L_1$  in  $L_2$  and the unbiased estimator is given by, see e.g. Jensen (1998),

$$m_{\text{cl}_1} = m_{\text{cl}_1}(Y \cap L_1) = 2\pi \int_{Y \cap L_1} d(y, o)^2 \, dy.$$

Here,  $d(\cdot, \cdot)$  is the Euclidean distance. If  $Y \cap L_1$  consists of a single line segment  $[y_-, y_+]$ , then  $m_{\text{cl}_1}$  takes the simple form

$$m_{\text{cl}_1} = \frac{2\pi}{3} (d(y_-, o)^3 + d(y_+, o)^3).$$

If  $Y \cap L_1$  is a union of such line segments,  $m_{\text{cl}_1}$  involves the measurements of distances from  $o$  to the endpoints of all line segments. In practice, usually an expert decides where the intersection points between the line and the boundary of  $Y$  are positioned.

In Gundersen (1988), it was suggested to use measurements along more than one line to increase the efficiency of the estimator. The resulting estimator  $m_{\text{cl}_2}$  is based on measurements along two perpendicular lines. This unbiased estimator is given by

$$m_{\text{cl}_2} = \frac{1}{2}(m_{\text{cl}_1}(Y \cap L_1(\Theta)) + m_{\text{cl}_1}(Y \cap L_1(\Theta + \pi/2))),$$

where  $\Theta$  is uniform on  $[0, \pi/2)$ . Introducing a perpendicular line implies a significant reduction in estimator variance if a large intersection with the cell at one of the lines typically is accompanied by a relatively small intersection at the perpendicular line (antithetic effect) and  $m_{\text{cl}_2}$  is much used in practice today. Recent references are Abrahão et al. (2009) and Melo et al. (2009).

### B.2.1.2 The integrated nucleator

The integrated nucleator is an unbiased estimator of particle volume based on the whole transect  $Y \cap L_2$ , see e.g. Jensen (1998) and references therein. The estimator takes the following form

$$m_{\text{int}} = m_{\text{int}}(Y \cap L_2) = 2 \int_{Y \cap L_2} d(y, o) dy^2.$$

Here,  $dy^2$  denotes the element of the area measure on  $L_2$ . This estimator has not been in common use in the study of biological cell populations because accurate recognition of  $Y \cap L_2$  has not appeared to be a realistic possibility. Today, due to improved labelling techniques, automatic identification of  $Y \cap L_2$  is no longer unrealistic.

Using polar decomposition in  $L_2$ , we find that

$$m_{\text{int}} = 2 \int_0^\pi \int_{Y \cap L_1(\theta)} d(y, o)^2 dy^1 d\theta = \int_0^\pi m_{\text{cl}_1}(Y \cap L_1(\theta)) \frac{d\theta}{\pi}. \quad (\text{B.1})$$

It follows that

$$\text{E}(m_{\text{cl}_1}|Y, L_2) = \text{E}(m_{\text{cl}_2}|Y, L_2) = m_{\text{int}}(Y \cap L_2). \quad (\text{B.2})$$

Equation (B.1) also shows that the integrated nucleator can be regarded as a classical nucleator based on an infinite number of lines.

### B.2.1.3 The automatic nucleator

If it is possible by automatic image analysis to recognise the cell transect  $Y \cap L_2$ , it appears natural to use this information in connection with the integrated nucleator. Let  $\tilde{Y}_2$  be the readily available estimate of the cell transect  $Y \cap L_2$  obtained using computerised image analysis and assume that  $\tilde{Y}_2$  contain  $o$ . Furthermore, let

$$m_{\text{aut}} = m_{\text{int}}(\tilde{Y}_2).$$

If the segmentation is perfect, i.e.  $\tilde{Y}_2 = Y \cap L_2$ , then the automatic nucleator  $m_{\text{aut}}$  provides an unbiased estimator of the volume of  $Y$ . If, on the other hand, the segmentation is of poor quality, then  $m_{\text{aut}}$  may be heavily biased.



#### B.2.1.4 The semi-automatic nucleator

As mentioned above, the advantages of the automatic nucleator depend on a satisfactory automatic identification of the cell transect boundaries. Below we present an intermediate possibility between the classical nucleator and the automatic nucleator where the expert supervises the measurement process performed using the automatic nucleator and only interferes if the recognition of the cell transect boundaries is not satisfactory. In such cases, the expert will perform the measurements needed for the classical nucleator. We will refer to this estimator as the semi-automatic nucleator  $m_{\text{semi}}$ .

If the segmentation is judged satisfactory, then  $m_{\text{semi}} = m_{\text{aut}}$ , otherwise  $m_{\text{semi}} = m_{\text{cl}_2}$ . In contrast to the classical and integrated nucleator, the semi-automatic nucleator may have a small bias. Let  $A$  be the event that the segmentation is accepted, i.e. judged satisfactory. It can be shown, see the Appendix, that

$$\text{bias}(m_{\text{semi}}) := E(m_{\text{semi}}) - \mu = p E(m_{\text{aut}} - m_{\text{int}}|A),$$

where  $p = \mathbb{P}(A)$  is the probability that the segmentation is accepted and  $E(m_{\text{aut}} - m_{\text{int}}|A)$  is the mean difference between the automatic and integrated nucleator among cells for which the segmentation is accepted. The bias will be small because  $m_{\text{aut}} - m_{\text{int}}$  is small for a cell with accepted segmentation.

### B.2.2 Variance and mean square error relations

If  $m$  denotes either  $m_{\text{cl}_1}$  or  $m_{\text{cl}_2}$ , then we have the following equation and lower bound for the variance of the estimator

$$\begin{aligned} \text{Var}(m) &= \text{Var}(E(m|Y, L_2)) + E \text{Var}(m|Y, L_2) \\ &= \text{Var}(m_{\text{int}}) + E \text{Var}(m|Y, L_2) \\ &\geq \text{Var}(m_{\text{int}}). \end{aligned} \tag{B.3}$$

At the second equality sign, we have used (B.2).

To compare estimators that are not necessarily unbiased the mean square error (MSE) is a more appropriate measure of variability. The MSE of an estimator  $m$  of  $\mu$  is given by, see e.g. Cochran (1977),

$$\text{MSE}(m) = E(m - \mu)^2 = \text{Var}(m) + \text{bias}(m)^2.$$

As  $m_{\text{int}}$  is unbiased we have that

$$\text{bias}(m) = E m - E m_{\text{int}}.$$

If  $m$  is unbiased then  $\text{MSE}(m) = \text{Var}(m)$ . Therefore the variance relations in (B.3) also holds for the corresponding MSEs, when  $m$  is unbiased.

For  $m_{\text{semi}}$ , we find, see the Appendix,

$$\text{MSE}(m_{\text{semi}}) = p \text{MSE}(m_{\text{aut}}|A) + (1 - p) \text{MSE}(m_{\text{cl}_2}|A^c), \tag{B.4}$$

where  $\text{MSE}(m_{\text{aut}}|A)$  is the mean square error of the automatic nucleator among cells for which a satisfactory segmentation is obtained. Likewise,  $\text{MSE}(m_{\text{cl}_2}|A^c)$  is the mean square error of  $m_{\text{cl}_2}$  among cells for which the segmentation is not satisfactory. The semi-automatic nucleator  $m_{\text{semi}}$  will be more precise than the classical nucleator  $m_{\text{cl}_2}$  if

$$\text{MSE}(m_{\text{semi}}) \leq \text{Var}(m_{\text{cl}_2}) = \text{MSE}(m_{\text{cl}_2}). \quad (\text{B.5})$$

Using (B.4), it follows that (B.5) is equivalent to

$$\text{MSE}(m_{\text{aut}}|A) \leq \text{MSE}(m_{\text{cl}_2}|A). \quad (\text{B.6})$$

The inequality (B.6) is satisfied if  $m_{\text{aut}}$  is replaced by  $m_{\text{int}}$ , see the Appendix, and is therefore likely to hold since  $\text{MSE}(m_{\text{aut}}|A)$  is calculated for cells with satisfactory segmentation. The magnitude of the gain in efficiency by using  $m_{\text{semi}}$  instead of  $m_{\text{cl}_2}$  will depend on the shapes of the cells and, for a given cell population, on how large the fraction  $p$  of cells with satisfactory segmentation is. Note also that the workload associated with determining  $m_{\text{semi}}$  will not be larger than that associated with  $m_{\text{cl}_2}$ .

### B.2.3 Estimation of bias and MSE

In order to determine which one of the estimators will be the most efficient one in an actual study, it is advisable to perform a pilot study.

Let  $Y_1, \dots, Y_M$  denote the sample of cells in such a study. For all estimators  $m$  except  $m_{\text{semi}}$ , we imagine that we perform  $N$  replicated determinations  $m_{i1}, \dots, m_{iN}$  of the estimator for each cell  $Y_i$ ,  $i = 1, \dots, M$ . The MSE of  $m$  can then be estimated by

$$\widehat{\text{MSE}}(m) = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (m_{ij} - \hat{\mu})^2. \quad (\text{B.7})$$

In (B.7) and the following, we use  $\hat{\mu} = \bar{m}_{\text{int}}$ , but an estimate of  $\mu$  based on one of the other unbiased estimators might be used as well.

Since  $m_{\text{int}}$  and  $m_{\text{aut}}$  depend on the cell transect as a whole, we will typically have  $N = 1$  for these estimators. For  $m_{\text{cl}_1}$  and  $m_{\text{cl}_2}$ , we may be interested in estimating the within section variance due to the random positioning of the lines inside the cell transects and in this case  $N > 1$ . The within section variance can be estimated by

$$\frac{1}{M(N-1)} \sum_{i=1}^M \sum_{j=1}^N (m_{ij} - \bar{m}_i)^2.$$

For  $m_{\text{semi}}$ , we let  $I(|I|)$  denote the (number of cells in the) sub-population of the sampled cells that are judged to have a satisfactory segmentation. The

probability of a satisfactory segmentation is estimated by  $\hat{p} = |I|/M$ . Then, an estimate of  $\text{MSE}(m_{\text{semi}})$  is, cf. (B.4),

$$\begin{aligned}\widehat{\text{MSE}}(m_{\text{semi}}) &= \hat{p} \widehat{\text{MSE}}(m_{\text{aut}}|A) + (1 - \hat{p}) \widehat{\text{MSE}}(m_{\text{cl}_2}|A^c) \\ &= \frac{|I|}{M} \frac{1}{|I|} \sum_{i \in I} (m_{\text{aut},i} - \hat{\mu})^2 + \left(1 - \frac{|I|}{M}\right) \frac{1}{(M-|I|)N} \sum_{i \notin I} \sum_{j=1}^N (m_{\text{cl}_2,ij} - \hat{\mu})^2 \\ &= \frac{1}{M} \left( \sum_{i \in I} (m_{\text{aut},i} - \hat{\mu})^2 + \frac{1}{N} \sum_{i \notin I} \sum_{j=1}^N (m_{\text{cl}_2,ij} - \hat{\mu})^2 \right).\end{aligned}$$

Among the estimators considered, only  $m_{\text{aut}}$  and  $m_{\text{semi}}$  may be biased. For  $m = m_{\text{aut}}$ , the bias is estimated by

$$\widehat{\text{bias}}(m_{\text{aut}}) = \bar{m}_{\text{aut}\cdot} - \hat{\mu},$$

while the bias of  $m_{\text{semi}}$  can be estimated by

$$\widehat{\text{bias}}(m_{\text{semi}}) = \frac{1}{M} \sum_{i \in I} (m_{\text{aut},i} - m_{\text{int},i}).$$

#### B.2.4 Discrimination between estimators

A yardstick for the precision of an estimator of the mean cell volume  $\mu$  is the number  $n$  of cells needed to obtain a given precision. The estimator

$$\hat{m} = \frac{1}{n} \sum_{i=1}^n m(Y_i)$$

of mean cell volume has MSE of the form

$$\text{MSE}(\hat{m}) = \frac{1}{n} \text{Var}(m) + \text{bias}(m)^2.$$

It is seen that in order to obtain a relative error  $\rho = \sqrt{\text{MSE}(\hat{m})}/\mu$  of the estimate  $\hat{m}$  of mean cell volume, we need to sample

$$n = \frac{\text{MSE}(m) - \text{bias}(m)^2}{\rho^2 \mu^2 - \text{bias}(m)^2} \quad (\text{B.8})$$

cells. If we let

$$\text{Relative bias}(m) = \text{bias}(m) / \mu$$

and

$$\text{Relative error}(m) = \sqrt{\text{MSE}(m)} / \mu,$$

then (B.8) reduces to

$$n = \frac{\text{Relative error}^2(m) - \text{Relative bias}^2(m)}{\rho^2 - \text{Relative bias}^2(m)}. \quad (\text{B.9})$$

Estimates of  $\mu$ ,  $\text{bias}(m)$  and  $\text{MSE}(m)$  are available from the pilot study, cf. Section B.2.3.

### B.3 A comparative study

In this section we will compare the described estimators in a study of somatostatin positive inhibitory interneurons from transgenic GFP-GAD mice hippocampi, observed by optical fluorescent microscopy.

#### B.3.1 Materials and preparation methods

The animal study was approved by the Danish Animal Experiments Inspectorate. Two GFP-GAD (FVB-TgN (GadGFP) 45704Swn) mice were anaesthetised by sodium pentobarbital (50 mg/kg i.p.) and transcardially perfused with phosphate-buffered 4 % paraformaldehyde. Following post-fixation in 4 % paraformaldehyde overnight at 4 °C, the brains were removed, and the hippocampi were cut out and embedded in 5 % agar in the isector (Nyengaard and Gundersen, 1992) to generate isotropic sections. Using a Vibratome (Vibratome, St. Louis, USA), the brains were sectioned exhaustively at 65  $\mu\text{m}$ . Every sixth section was selected for sampling. For counterstaining, these sections were transferred into DAPI (Sigma, St. Louis, USA) solution. Each section was wet mounted on a super frost slide and was dried at room temperature for only 10 min. An aqueous mounting media was used to adhere the cover-glass and care was taken to remove the excess mounting media. *Z*-stacks were recorded at a confocal microscope (Zeiss LSM 510 META system), using a 40 $\times$  NA 1.2 C-Apochromat objective and systematic sampling. The laser line used was 488 nm, image size 225  $\mu\text{m} \times 225 \mu\text{m}$  and voxel size 0.44  $\mu\text{m} \times 0.44 \mu\text{m} \times 0.44 \mu\text{m}$  (undersampling in the XY plane with respect to the optimum resolution which is 0.1  $\mu\text{m} \times 0.1 \mu\text{m}$  at this wavelength and NA).

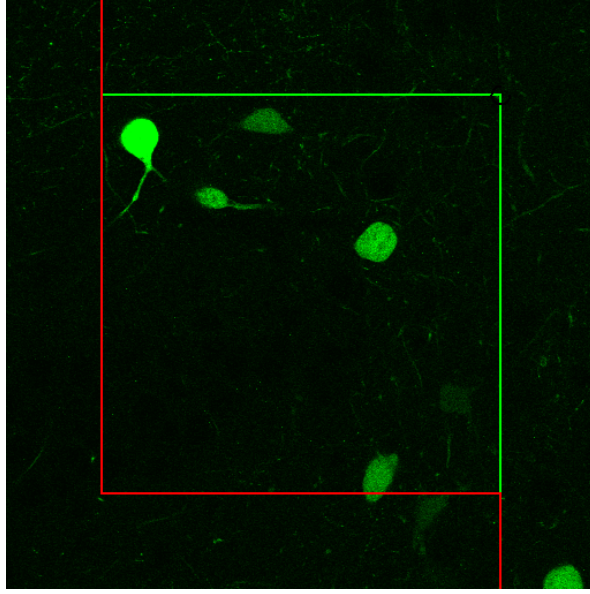
#### B.3.2 Sampling and segmentation

Figure B.2 shows examples of somatostatin positive inhibitory interneurons (green) in an optical section of tissue as seen under a confocal microscope. A characteristic of these cells is the dendrites/axons which are also visible in Figure B.2. The dendrites will not be regarded as part of the actual cell body and, accordingly, they do not contribute to the cell volume.

A total of 91 cells were sampled using an optical disector within an isotropic thick section and an unbiased counting frame. For each sampled cell a segmentation of the boundary of the central cell transect was performed using the max-vol function in Visiormorph (Visiopharm, Hørsholm, Denmark). The segmentation results in a set of *xy*-coordinates on the segmentation boundary. These 91 sets of *xy*-coordinates constitute the data to be used in the following analysis.

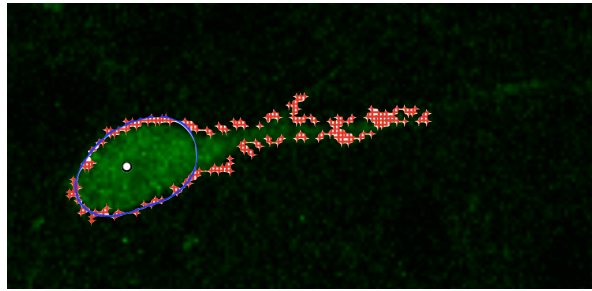
#### B.3.3 Constructing the true cell transect

In order to study the performance of the different types of estimators, we fitted for each sampled cell transect a cubic smoothing spline to those *xy*-coordinates not originating from the dendrites or from other cells visible from other layers



**Figure B.2:** Somatostatin positive inhibitory interneurons in mice hippocampi have been genetically labelled with green fluorescent protein (GFP) and observed under a confocal microscope. The optical disector with the unbiased counting frame is used to sample the interneurons.

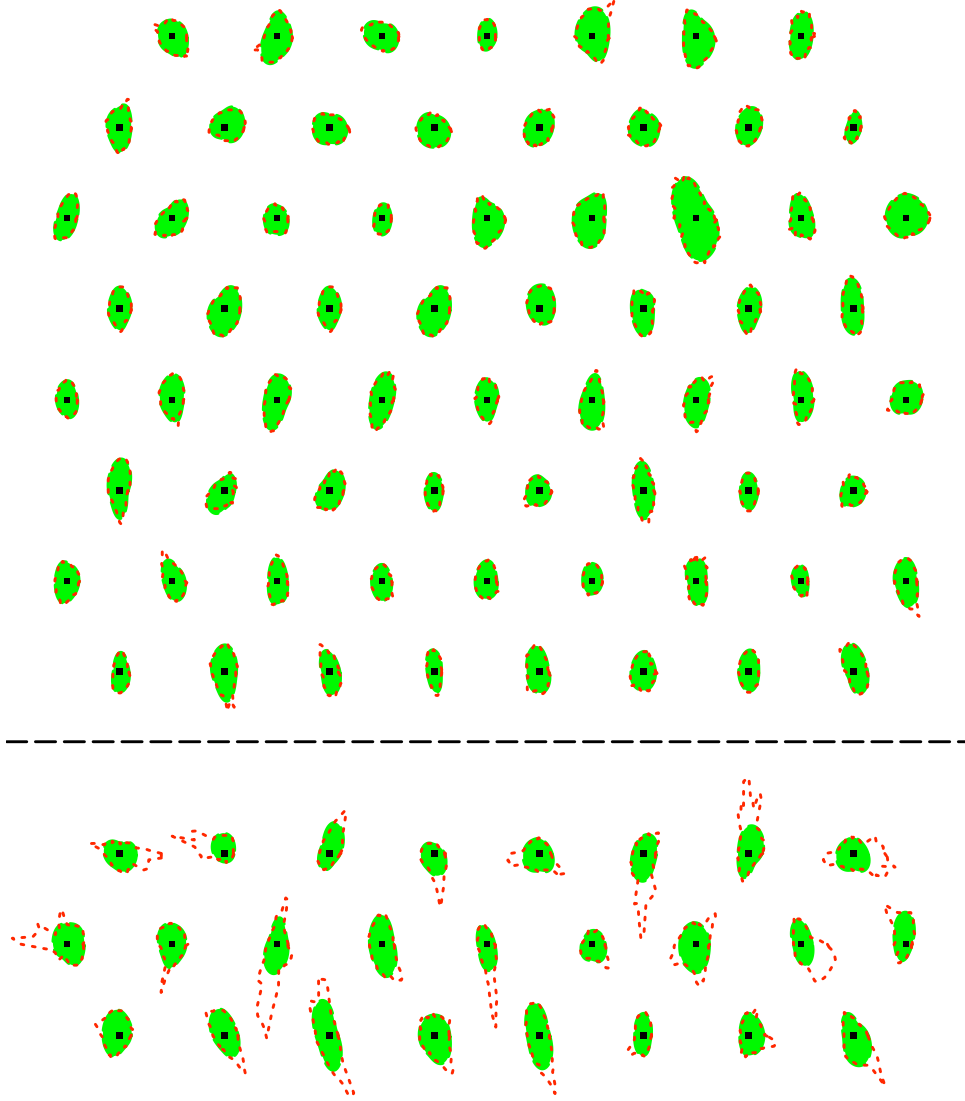
of the thick section. Subsequently, the spline curve representing  $Y \cap L_2$  was approved by an expert and used in the following as the true cell transect. In the succeeding analysis  $m_{\text{int}}$  was calculated by numerical integration using (B.1) with  $Y$  replaced by the set delineated by the spline curve.



**Figure B.3:** A somatostatin positive inhibitory interneuron as observed under the confocal laser microscope. The observed  $xy$ -coordinates on the segmentation boundary are shown (red crosses) along with the calculated spline (blue line) and the centre of mass relative to the spline (white dot).

Figure B.3 shows an example of a sampled cell (green) as seen under the confocal laser microscope. The observed  $xy$ -coordinates (red crosses) are superimposed along with the calculated spline (blue line) and the centre of mass of the spline (white dot). As these cells do not have a natural reference point we will use the centre of mass for this purpose. This centre of mass in the central

section of the cell may be regarded as an approximation to the centre of mass in 3D. This approximation may result in a small bias for all the estimators considered in the present paper.

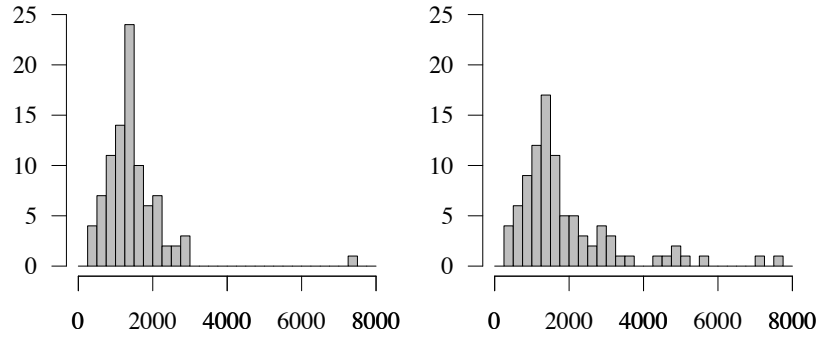


**Figure B.4:** The 'true' cell transect (green) and the boundary of the segmented cell (red broken line) for each of the 91 sampled somatostatin positive inhibitory interneurons. The dot indicates the reference point of the cell. The cell transects above the horizontal line have satisfactory segmentation.

Figure B.4 shows for each cell the true cell transect (green) and the boundary of the segmented cell (red broken line). The dot indicates the centre of mass according to the true cell transect. The cell transects above the horizontal line of Figure B.4 have satisfactory segmentation while for those below the segmented cell transects are too large, either because parts of the dendrites are included or neighbouring cells interfere.

### B.3.4 Comparison of estimators

These visual observations are reflected in the histograms in Figure B.5 displaying the estimates of  $m_{\text{int}}$  (left) and  $m_{\text{aut}}$  (right), respectively. Furthermore, the histograms indicate a right skewed distribution, especially for  $m_{\text{aut}}$ , with means  $E(m_{\text{int}})$  and  $E(m_{\text{aut}})$  estimated by  $1446 \mu\text{m}^3$  and  $1904 \mu\text{m}^3$ , respectively. Furthermore, the estimates of  $\text{Var}(m_{\text{int}})$  and  $\text{Var}(m_{\text{aut}})$  are  $706\,620 \mu\text{m}^6$  and  $1\,944\,011 \mu\text{m}^6$ , respectively. The estimate of the bias of  $m_{\text{aut}}$  is  $458 \mu\text{m}^3$  resulting in an estimated MSE of  $m_{\text{aut}}$  of  $2\,132\,314 \mu\text{m}^6$ . The estimated relative bias of  $m_{\text{aut}}$  is quite large, i.e.  $(\bar{m}_{\text{aut.}} - \hat{\mu})/\hat{\mu} = 0.32$ . It is therefore not advisable to use the automatic nucleator in this study.



**Figure B.5:** Histograms of the estimated volumes ( $\mu\text{m}^3$ ) based on the integrated nucleator (left) and the automatic nucleator (right).

For the implementation of the semi-automatic nucleator, we used a distance in the judgement of the discrepancy between the true cell transect  $Y \cap L_2$  and the automatically segmented cell transect called  $\tilde{Y}_2$ . The segmentation was judged satisfactory if

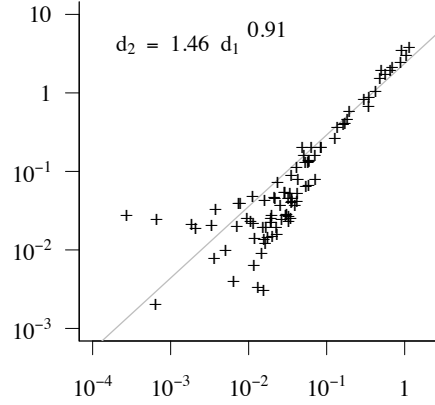
$$d(Y \cap L_2, \tilde{Y}_2) < \varepsilon$$

where  $\varepsilon \geq 0$  and  $d$  denotes a distance on the set of subsets of  $L_2$ . Two examples of distances were considered

$$d_1(B_1, B_2) = \frac{|A(B_1) - A(B_2)|}{A(B_1)} \quad \text{and} \quad d_2(B_1, B_2) = \frac{|m_{\text{int}}(B_1) - m_{\text{int}}(B_2)|}{m_{\text{int}}(B_1)}.$$

Here,  $A$  denotes area and  $B_1, B_2 \subset L_2$ . In practice, the distance  $d_1$  seems straightforward for an expert to evaluate since the distance only involves differences in area between the true cell transect and the segmentation. But as it is the difference in estimated volume that is important  $d_2$  seems more appropriate. Figure B.6 shows  $d_2(Y \cap L_2, \tilde{Y}_2)$  plotted against  $d_1(Y \cap L_2, \tilde{Y}_2)$  in a double logarithmic scale along with a fitted log-linear regression  $d_2 = 1.46 d_1^{0.91}$ . It is seen from this plot, that the two distances provide quite similar results.

In the analysis below we used  $d_2$  and  $\varepsilon = 0.15$ . A total of 66 cells had a satisfactory segmentation according to this criterion. They are shown above the horizontal line of Figure B.4. The probability  $p$  is estimated by  $66/91 = 0.73$ .



**Figure B.6:** Plot of  $d_2(Y \cap L_2, \tilde{Y}_2)$  against  $d_1(Y \cap L_2, \tilde{Y}_2)$  in a double logarithmic scale.

Using the empirical relationship between  $d_1$  and  $d_2$ ,  $d_2(Y \cap L_2, \tilde{Y}_2) = 0.15$  corresponds to  $d_1(Y \cap L_2, \tilde{Y}_2) = 0.08$ .

For each sampled cell, Figure B.7 shows the volume estimates  $m_{\text{aut}}$ ,  $m_{\text{cl}_1}$ ,  $m_{\text{cl}_2}$  and  $m_{\text{semi}}$  plotted against  $m_{\text{int}}$ , respectively. The scale is  $\mu\text{m}^3$ . The upper left plot indicates that  $m_{\text{aut}}$  is quite heavily biased while the remaining three plots are in accordance with the theoretical result that  $m_{\text{int}}$ ,  $m_{\text{cl}_1}$  and  $m_{\text{cl}_2}$  are unbiased. Also,  $m_{\text{semi}}$  appears virtually unbiased and somewhat more precise than  $m_{\text{cl}_2}$  which again is more precise than  $m_{\text{cl}_1}$ . In the upper left plot of Figure B.7, the points with large distance to the identity line come from cells with unsatisfactory segmentation.

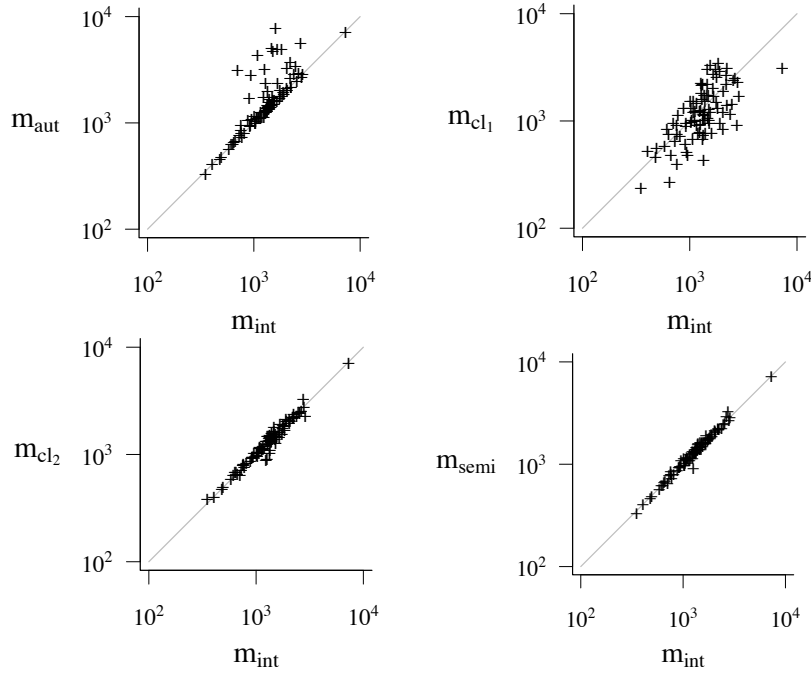
Table B.2 shows, for each estimator considered, the relative bias and the relative error of the estimated cell volume, when analysing one cell. For  $m_{\text{cl}_1}$  and  $m_{\text{cl}_2}$ , we also estimated the within section variance which was found to constitute 89 % and 29 % of the total variance, respectively. Still  $m_{\text{semi}}$  is more precise than  $m_{\text{cl}_2}$ , see Table B.2.

**Table B.2:** For each of the estimators considered, the relative bias and the relative error of the estimated cell volume, when analysing one cell, are shown. Note that  $m_{\text{semi}}$  is virtually unbiased.

	$m_{\text{int}}$	$m_{\text{aut}}$	$m_{\text{cl}_1}$	$m_{\text{cl}_2}$	$m_{\text{semi}}$
Relative bias	—	0.32	—	—	0.004
Relative error	0.58	1.01	1.79	0.69	0.61

Using (B.9) and the results shown in Table (B.2), we can estimate the number of cells needed to analyse in order to obtain a given relative error  $\rho$  of the estimated mean cell volume. Figure B.8 shows the number of cells needed to be analysed for  $m_{\text{int}}$  (full),  $m_{\text{semi}}$  (dashed) and  $m_{\text{cl}_2}$  (dotted), respectively, to obtain a given relative error of the estimated mean cell volume between 0.02 and 0.10. For a given relative error in this interval it is seen that we always





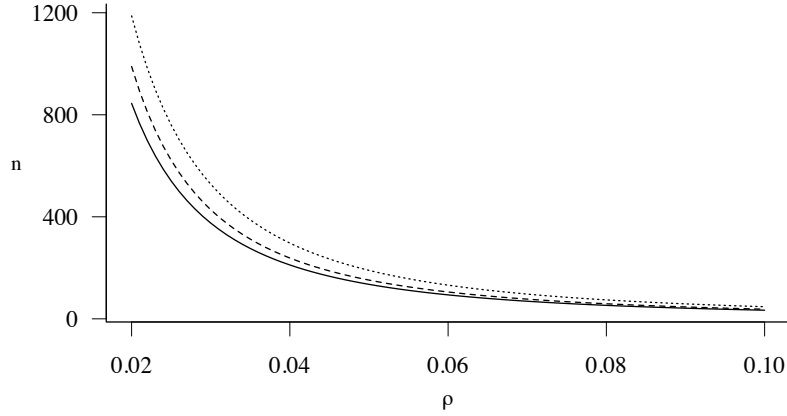
**Figure B.7:** Volume estimates  $\hat{m}_{\text{aut}}$ ,  $\hat{m}_{\text{cl}_1}$ ,  $\hat{m}_{\text{cl}_2}$  and  $\hat{m}_{\text{semi}}$  plotted against  $\hat{m}_{\text{int}}$  in a double logarithmic scale. The scale is  $\mu\text{m}^3$ .

need to include less cells for  $m_{\text{semi}}$  than for  $m_{\text{cl}_2}$ . As an example, in order to obtain a relative error of the estimate of mean cell volume of 0.05, it is needed to sample and analyse 134, 150 and 189 cells using the integrated nucleator, semi-automatic nucleator and classical nucleator, respectively. The use of  $m_{\text{int}}$  is not feasible in this study because it requires the correct segmentation of *all* cell transects. The semi-automatic nucleator is more efficient than the classical nucleator, since the same precision can be obtained by sampling fewer cells in the case of the semi-automatic nucleator and in addition the time spent analysing one cell is shorter.

## B.4 Discussion

In the present paper, we have proposed a new method of estimating mean cell volume, the semi-automatic nucleator. The method uses sections through reference points of the cells. Automatic segmentation of the cell transects is an integral part of the method. A gain in efficiency compared to the classical nucleator can generally be expected. The magnitude of the gain in efficiency increases with the fraction of analysed cell transects with satisfactory segmentation.

An expert supervises the process and interferes if the segmentation of a particular cell transect is judged unsatisfactory. It is important that the expert is trained in using an appropriate threshold for deciding whether a segmentation is satisfactory.



**Figure B.8:** The number of cells  $n$  needed to be analysed for  $m_{\text{int}}$  (full),  $m_{\text{semi}}$  (dashed) and  $m_{\text{cl}_2}$  (dotted), respectively, as a function of the relative error  $\rho$  of the estimate of the mean cell volume.

The efficiency of nucleator type estimators depends on the positioning of the reference point inside the cell. If the reference point is very far from being centrally positioned, we advice to use the pivotal estimators suggested in Cruz-Orive (2005, 2008) and Cruz-Orive et al. (2010) instead.

Recently, weighted sampling has been introduced in stereology under the name of the *proportionator*, as a successful method of reducing variance in the stereological estimation of number (Gardi et al., 2007, 2008). During the research work reported in the present paper, we also considered to reduce the within section variance of the classical nucleator by changing the distribution of the lines in the section from uniform to a weighted orientation distribution, using the information available in the segmented cell transect  $\tilde{Y}_2$ .

More specifically, we considered lines  $L_1 = L_1(\theta)$  with the following orientation density distribution  $q$ , depending on the shape of  $\tilde{Y}_2$ ,

$$q(\theta) \propto \int_{\tilde{Y}_2 \cap L_1(\theta)} d(y, o)^2 dy,$$

for  $\theta \in [0, \pi)$ . This density favours lines with a large intersection with the cell. The following estimator

$$\frac{m_{\text{cl}_1}(Y \cap L_1(\Theta))}{\pi q(\Theta)}, \quad (\text{B.10})$$

where  $\Theta$  has density  $q$ , is an unbiased estimator of  $V(Y)$ . To see this note that

$$\mathbb{E}(m_{\text{cl}_1}(Y \cap L_1(\Theta)) / (\pi q(\Theta)) | Y, L_2) = \int_0^\pi m_{\text{cl}_1}(Y \cap L_1(\theta)) \frac{d\theta}{\pi} = \mathbb{E}(m_{\text{cl}_1} | Y, L_2).$$

In the case of a perfect segmentation, i.e.  $\tilde{Y}_2 = Y \cap L_2$ , the values of this estimator will not depend on  $\Theta$  and therefore its within section variability will be zero.

When comparing the estimator in (B.10) to the classical nucleator estimators, using the data analysed in Section B.3, we found that the estimator in (B.10)

had smaller variance than  $m_{cl_1}$  but larger variance than  $m_{cl_2}$ . Two reasons that the estimator (B.10) failed to outperform  $m_{cl_2}$  might be: *i*) the estimator  $m_{cl_2}$  has already a small within section variability due to the antithetic effect of the two perpendicular lines and *ii*) the true cell transects are not very irregular in shape and the variance reducing effect of weighted sampling cannot compensate for the effect that some extra within section variability is introduced when (B.10) is used in case of non-perfect segmentation.

Current developments in computerised image analysis make it possible to produce a rough voxel image of a sampled cell in which virtual central sections can be generated (Allan Rasmusson, personal communication). This feature is not yet an integrated part of any available computer software but it might be used in future applications of stereology in the following manner. If a rough voxel image of a sampled cell is available, we may use a distribution of the orientation of the section plane  $L_2$  that depends on this rough voxel image. Implementation of weighted sampling in this connection may lead to a further reduction in estimator variance that will affect all the estimators presented in this paper.

## B.5 Acknowledgements

This project has been supported by the Danish Council for Strategic Research and by Centre for Stochastic Geometry and Advanced Bioimaging, funded by a grant from The Villum Foundation.

## Bibliography

- Abrahão, L. M., Nyengaard, J. R., Sasahara, T. H., Gomes, S. P., Oliveira Fda, R., Ladd, F. V., Ladd, A. A., Melo, M. P., Machado, M. R., Melo, S. R., and Ribeiro, A. A. (2009). Asymmetric post-natal development of superior cervical ganglion of paca (Agouti paca). *International Journal of Developmental Neuroscience*, 27:37–45.
- Cochran, W. G. (1977). *Sampling Techniques*. John Wiley & Sons, New York.
- Cruz-Orive, L. M. (1976). Particle size-shape distributions: the general spheroid problem. I. Mathematical model. *Journal of Microscopy*, 107:235–253.
- Cruz-Orive, L. M. (1978). Particle size-shape distributions: the general spheroid problem. II. Stochastic model and particle guide. *Journal of Microscopy*, 112:153–167.
- Cruz-Orive, L. M. (2005). A new stereological principle for test lines in three dimensional space. *Journal of Microscopy*, 219:18–28.
- Cruz-Orive, L. M. (2008). Comparative precision of the pivotal estimators of particle size. *Image Analysis and Stereology*, 27:17–22.
- Cruz-Orive, L. M., Ramos-Herrera, M. L., and Artacho-Pérula, E. (2010). Stereology of isolated objects with the invariator. *Journal of Microscopy*, 240(2):94–110.

- Gardi, J., Nyengaard, J., and Gundersen, H. (2008). The proportionator: Unbiased stereological estimation using biased automatic image analysis and non-uniform probability proportional to size sampling. *Computers in Biology and Medicine*, 38(3):313–328.
- Gardi, J. E., Wulfsohn, D., and Nyengaard, J. R. (2007). A handheld support to facilitate stereological measurements and mapping of branching structures. *Journal of Microscopy*, 227:127–139.
- Gundersen, H. J. G. (1988). The nucleator. *Journal of Microscopy*, 151:3–21.
- Jensen, E. B. V. (1998). *Local Stereology*. World Scientific Publishing, Singapore.
- Jensen, E. B. V. (2000). On the variance of local stereological volume estimators. *Image Analysis and Stereology*, 19:15–18.
- Melo, S. R., Nyengaard, J. R., da Roza Oliveira, F., Ladd, F. V. L., Abrahão, L. M. B., Machado, M. R., Sasahara, T. H., de Melo, M. P., and Ribeiro, A. A. C. (2009). The developing left superior cervical ganglion of pacas (agouti paca). *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology*, 292(7):966–975.
- Nyengaard, J. R. and Gundersen, H. J. G. (1992). The isector: a simple and direct method for generating isotropic, uniform random sections from small specimens. *Journal of Microscopy*, 165:427–431.
- Wicksell, S. D. (1925). The corpuscle problem. A mathematical study of a biometric problem. *Biometrika*, 17:84–89.
- Wicksell, S. D. (1926). The corpuscle problem. Second memoir. Case of ellipsoidal corpuscles. *Biometrika*, 18:152–172.

## Appendix

In this Appendix, we will derive the results on bias and MSE of  $m_{\text{semi}}$  presented in Section B.2.1.4 and Section B.2.2 of the main text.

Let  $A_\varepsilon = \{ d(Y \cap L_2, \tilde{Y}_2) < \varepsilon \}$ . Then, the semi-automatic nucleator is given by

$$m_{\text{semi}} = \mathbb{1}_{A_\varepsilon} m_{\text{aut}} + \mathbb{1}_{A_\varepsilon^c} m_{\text{cl}_2},$$

where  $\mathbb{1}_A$  is the indicator function of  $A$ . Given  $Y$  and  $L_2$ , the distribution of  $m_{\text{cl}_2}$  depends only on the lines  $L_1(\Theta)$  and  $L_1(\Theta + \pi/2)$  where  $\Theta$  is uniform on  $[0, \pi/2)$ . It follows that  $m_{\text{cl}_2}$  and  $A_\varepsilon$  are conditional independent given  $Y$  and  $L_2$  and we find, using (B.2),

$$\begin{aligned} \mathbb{E}(m_{\text{semi}} | Y, L_2) &= \mathbb{E}(\mathbb{1}_{A_\varepsilon} m_{\text{aut}} | Y, L_2) + \mathbb{E}(\mathbb{1}_{A_\varepsilon^c} | Y, L_2) \mathbb{E}(m_{\text{cl}_2} | Y, L_2) \\ &= \mathbb{E}(\mathbb{1}_{A_\varepsilon} m_{\text{aut}} | Y, L_2) + \mathbb{E}(\mathbb{1}_{A_\varepsilon^c} | Y, L_2) m_{\text{int}} \\ &= \mathbb{E}(\mathbb{1}_{A_\varepsilon} m_{\text{aut}} + \mathbb{1}_{A_\varepsilon^c} m_{\text{int}} | Y, L_2). \end{aligned}$$

Thereby,

$$\mathbb{E}(m_{\text{semi}}) = \mathbb{E}(\mathbb{1}_{A_\varepsilon} m_{\text{aut}} + \mathbb{1}_{A_\varepsilon^c} m_{\text{int}}).$$

It follows that

$$\begin{aligned}
\text{bias}(m_{\text{semi}}) &= E(m_{\text{semi}}) - E(m_{\text{int}}) \\
&= E(\mathbb{1}_{A_\varepsilon} m_{\text{aut}} + \mathbb{1}_{A_\varepsilon^c} m_{\text{int}}) - E(\mathbb{1}_{A_\varepsilon} m_{\text{int}} + \mathbb{1}_{A_\varepsilon^c} m_{\text{int}}) \\
&= E(\mathbb{1}_{A_\varepsilon} (m_{\text{aut}} - m_{\text{int}})) \\
&= p E(m_{\text{aut}} - m_{\text{int}} | A_\varepsilon).
\end{aligned}$$

Using the distance  $d_2$ , we have the following inequality between estimators

$$\mathbb{1}_{A_\varepsilon} (1 - \varepsilon) m_{\text{int}} < \mathbb{1}_{A_\varepsilon} m_{\text{aut}} < \mathbb{1}_{A_\varepsilon} (1 + \varepsilon) m_{\text{int}}.$$

Therefore,

$$\begin{aligned}
E(m_{\text{semi}}) &< (1 + \varepsilon) E(\mathbb{1}_{A_\varepsilon} m_{\text{int}}) + E(\mathbb{1}_{A_\varepsilon^c} m_{\text{int}}) \\
&= \varepsilon E(\mathbb{1}_{A_\varepsilon} m_{\text{int}}) + E(m_{\text{int}}) \\
&= \varepsilon p E(m_{\text{int}} | A_\varepsilon) + \mu.
\end{aligned}$$

Likewise,  $E(m_{\text{semi}}) > -\varepsilon p E(m_{\text{int}} | A_\varepsilon) + \mu$  and we get a bound for the relative bias

$$\frac{|E(m_{\text{semi}}) - \mu|}{\mu} \leq \varepsilon p \frac{E(m_{\text{int}} | A_\varepsilon)}{\mu}.$$

The results on the MSE of  $m_{\text{semi}}$  is derived as follows

$$\begin{aligned}
\text{MSE}(m_{\text{semi}}) &= E(\mathbb{1}_{A_\varepsilon} [m_{\text{aut}} - \mu]^2 + \mathbb{1}_{A_\varepsilon^c} [m_{\text{cl}_2} - \mu]^2) \\
&= p \text{MSE}(m_{\text{aut}} | A_\varepsilon) + (1 - p) \text{MSE}(m_{\text{cl}_2} | A_\varepsilon^c).
\end{aligned}$$

Finally, we show that (B.6) is satisfied when  $m_{\text{aut}}$  is replaced by  $m_{\text{int}}$ , i.e.

$$\text{MSE}(m_{\text{int}} | A_\varepsilon) \leq \text{MSE}(m_{\text{cl}_2} | A_\varepsilon).$$

Utilising that  $m_{\text{cl}_2}$  and  $A_\varepsilon$  are conditionally independent given  $Y$  and  $L_2$ , we find

$$E(m_{\text{cl}_2} | A_\varepsilon) = E(E(m_{\text{cl}_2} | A_\varepsilon, Y, L_2) | A_\varepsilon) = E(m_{\text{int}} | A_\varepsilon)$$

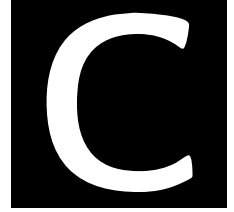
and

$$\begin{aligned}
E(m_{\text{cl}_2}^2 | A_\varepsilon) &= \text{Var}(m_{\text{cl}_2} | A_\varepsilon) + (E(m_{\text{cl}_2} | A_\varepsilon))^2 \\
&\geq \text{Var}(E(m_{\text{cl}_2} | A_\varepsilon, Y, L_2) | A_\varepsilon) + (E(m_{\text{int}} | A_\varepsilon))^2 \\
&= \text{Var}(m_{\text{int}} | A_\varepsilon) + (E(m_{\text{int}} | A_\varepsilon))^2 \\
&= E(m_{\text{int}}^2 | A_\varepsilon).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\text{MSE}(m_{\text{cl}_2} | A_\varepsilon) &= E((m_{\text{cl}_2} - \mu)^2 | A_\varepsilon) \\
&= E(m_{\text{cl}_2}^2 | A_\varepsilon) - 2\mu E(m_{\text{cl}_2} | A_\varepsilon) + \mu^2 \\
&\geq E(m_{\text{int}}^2 | A_\varepsilon) - 2\mu E(m_{\text{int}} | A_\varepsilon) + \mu^2 \\
&= \text{MSE}(m_{\text{int}} | A_\varepsilon).
\end{aligned}$$





# Lévy particles: Modelling and simulating star-shaped random sets

Linda V. Hansen<sup>1</sup>, Thordis L. Thorarinsdottir<sup>2</sup> and Tilmann Gneiting<sup>2</sup>

<sup>1</sup>*Centre for Stochastic Geometry and Advanced Bioimaging, Aarhus University;*

<sup>2</sup>*Institute of Applied Mathematics, University of Heidelberg*

**Abstract:** Lévy particles provide a flexible framework for modelling and simulating three-dimensional star-shaped random sets. The radial function of a Lévy particle arises from a kernel smoothing of a Lévy basis, and is associated with an isotropic random field on the sphere. If the kernel is proportional to a von Mises–Fisher density, or uniform on a spherical cap, the correlation function of the associated random field admits a closed form expression. Using a Gaussian basis, the fractal or Hausdorff dimension of the surface of the Lévy particle reflects the decay of the correlation function at the origin, as quantified by the fractal index. Power kernels are conjectured to yield particles with boundaries of any Hausdorff dimension between 2 and 3.

**Key words:** celestial body, correlation function, fractal index, Hausdorff dimension, Lévy basis, random field on a sphere, simulation of star-shaped random sets





## C.1 Introduction

Mathematical models for three-dimensional particles have received great interest in astronomy, botany, geology, material science and zoology, among many other disciplines. Early approaches include that of Wicksell (1925, 1926), who addressed the estimation of the number and the size distribution of three-dimensional corpuscles in biological tissue from planar sections. He proposed both a spherical model and a more flexible ellipsoidal model. Wicksell's ideas were elaborated half a century later by Cruz-Orive (1976, 1978), who studied more general particle size-shape distributions in both continuous and discrete settings.

While some particles such as crystals have a rigid shape, many real-world objects are star-shaped, highly structured and stochastically varying. As a result, flexible yet parsimonious models for star-shaped random sets have been in high demand. Grenander and Miller (1994) proposed a model for two-dimensional featureless objects with no obvious landmarks, which are represented by a deformed polygon along with a Gaussian shape model. This was investigated further in Kent et al. (2000) and Hobolth et al. (2002), and a non-Gaussian extension was suggested by Hobolth et al. (2003). Miller et al. (1994) proposed an isotropic deformation model that relies on spherical harmonics and was studied by Hobolth (2003), where it was applied to monitor tumour growth. A related Gaussian random shape model was studied by Muinonen et al. (1996) and used by Muñoz et al. (2007) to represent Saharan desert dust particles.

In this paper we propose a flexible framework for modelling three-dimensional star-shaped particles, where the radial function is a random field on the sphere that arises through a kernel smoothing of a Lévy basis. Specifically, let  $Y \subset \mathbb{R}^3$  be a three-dimensional compact set, which is star-shaped with respect to an interior point  $o$ . Then there is a one-to-one correspondence between the set  $Y$  and its radial function  $X = \{X(u) : u \in \mathbb{S}^2\}$ , where

$$X(u) = \max\{r \geq 0 : o + ru \in Y\}, \quad u \in \mathbb{S}^2,$$

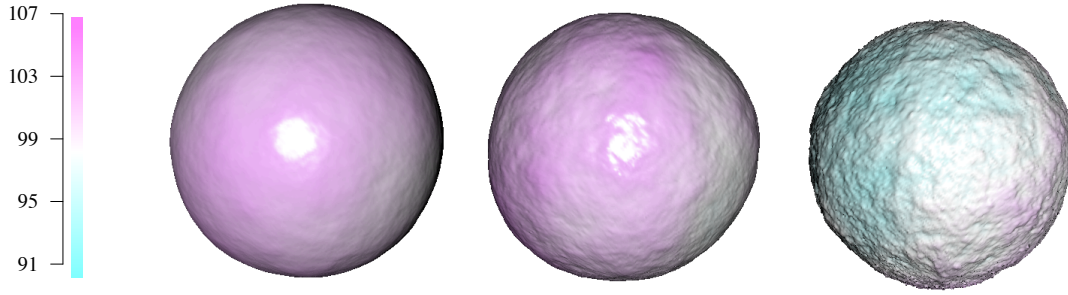
with  $\mathbb{S}^2 = \{x \in \mathbb{R}^3 : \|x\| = 1\}$  denoting the unit sphere in  $\mathbb{R}^3$ . We model  $X$  as a real-valued random field on  $\mathbb{S}^2$  via a kernel smoothing of a Lévy basis, in that

$$X(u) = \int_{\mathbb{S}^2} K(v, u) \Gamma(\mathrm{d}v), \quad u \in \mathbb{S}^2, \quad (\text{C.1})$$

where  $K : \mathbb{S}^2 \times \mathbb{S}^2 \rightarrow \bar{\mathbb{R}}$  is a suitable kernel function, and  $\Gamma$  is a Lévy basis on the Borel subsets of  $\mathbb{S}^2$ , that is, an infinitely divisible and independently scattered random measure. If  $X$  is a nonnegative process, the random particle can then be described as the set

$$Y = \bigcup_{u \in \mathbb{S}^2} \{o + ru : 0 \leq r \leq X(u)\} \subset \mathbb{R}^3, \quad (\text{C.2})$$

so that the particle contains the centre  $o$ , which without loss of generality can be assumed to be the origin, and the distance in direction  $u$  from  $o$  to the particle



**Figure C.1:** Lévy particles with mean  $\mu_X = 100$  and variance  $\sigma_X^2 = 10$ , using a Gaussian Lévy basis and the power kernel (C.14) with  $q = 0.05$  (left),  $q = 0.25$  (middle) and  $q = 0.5$  (right).

boundary is given by  $X(u)$ . A potentially modified particle  $Y_c$  arises in the case of a general, not necessarily nonnegative process, where we replace  $X(u)$  by  $X_c(u) = \max\{c, X(u)\}$  for some  $c > 0$ . We call  $Y$  or  $Y_c$  a Lévy particle, with realisations being illustrated in Figure C.1. The Lévy particle framework is a special case of the linear spatio-temporal Lévy model proposed by Jónsdóttir et al. (2008) in the context of tumour growth, and allows for flexible shape and scale. Alternatively, it can be seen as a generalisation and a three-dimensional extension of the model proposed in Hobolth et al. (2003), while also being a generalisation of the Gaussian random shape models of Miller et al. (1994) and Muinonen et al. (1996).

The realisations in Figure C.1 demonstrate that the boundary or surface of a Lévy particle allows for regular as well as irregular behaviour. The roughness or smoothness of the surface in the limit as the observational scale becomes infinitesimally fine can be quantified by the fractal or Hausdorff dimension, which for a surface in  $\mathbb{R}^3$  varies between 2 and 3, with the lower limit corresponding to a smooth, differentiable surface, and the upper limit corresponding to an excessively rough, space-filling surface (Falconer, 1990). The concept dates back to Hausdorff (1918) and has attracted much attention due to the work of Mandelbrot, who argued that fractal objects and surfaces are ubiquitous in nature (Mandelbrot, 1982). Under a Gaussian Lévy basis, the Hausdorff dimension of the surface of an isotropic Lévy particle is determined solely by the behaviour of the correlation function of the associated random field on the sphere. We investigate the properties of Lévy particles under parametric families of isotropic kernel functions, including power kernels, and kernels that are proportional to von Mises–Fisher densities, or uniform on spherical caps. Von Mises–Fisher and uniform kernels generate Gaussian particles with boundaries of Hausdorff dimension 2 and 2.5, respectively. Power kernels generate Gaussian Lévy particles whose surface we conjecture to attain any Hausdorff dimension between 2 and 3.

The remainder of the paper is organised as follows. Section C.2 recalls basic properties of Lévy bases and of the radial function in the Lévy particle model (C.1). In Section C.3 we show how to derive the Hausdorff dimension

of an isotropic Gaussian Lévy particle from the infinitesimal behaviour of the correlation function of the underlying random field at the origin. Section C.4 introduces the aforementioned families of isotropic kernels and discusses the properties of the associated correlation functions and Lévy particles. Section C.5 presents a simulation algorithm and simulation examples, including a case study on celestial bodies and a discussion of planar Lévy particles. The paper ends with a discussion in Section C.6.

## C.2 Preliminaries

The properties of the random function (C.1) that characterises a Lévy particle process depend both on the kernel function  $K$  and the Lévy basis  $\Gamma$ . For the Lévy basis, we use two of the types considered by Jónsdóttir et al. (2008). Specifically, we assume  $\Gamma$  to be either a Gaussian Lévy basis with parameters  $\mu \in \mathbb{R}$  and  $\sigma^2 > 0$ , so that

$$\Gamma(A) \sim \mathcal{N}(\mu \lambda(A), \sigma^2 \lambda(A)), \quad (\text{C.3})$$

or a gamma Lévy basis with shape  $\kappa > 0$  and rate  $\tau > 0$ , so that

$$\Gamma(A) \sim \text{Ga}(\kappa \lambda(A), \tau), \quad (\text{C.4})$$

where  $\lambda(A)$  denotes the surface measure of a Borel set  $A \subseteq \mathbb{S}^2$ , with  $\lambda(\mathbb{S}^2) = 4\pi$ . We assume that the kernel function  $K$  is isotropic, in that  $K(v, u) = k(\vartheta(v, u))$  depends on the points  $v, u \in \mathbb{S}^2$  through their great circle distance  $\vartheta(v, u) \in [0, \pi]$  only. As  $\cos \vartheta(v, u) = u \cdot v$ , this is equivalent to assuming that the kernel depends on the inner product  $u \cdot v$  only. Jónsdóttir et al. (2008) show that in the isotropic case the mean function  $E(X(u))$  and the variance function  $\text{Var}(X(u))$  are constant, that is,

$$\mu_X = E(X(u)) = \mu_\Gamma c_1 \quad \text{and} \quad \sigma_X^2 = \text{Var}(X(u)) = \sigma_\Gamma^2 c_2,$$

for  $u \in \mathbb{S}^2$ , where we assume that

$$c_n = \int_{\mathbb{S}^2} k(\vartheta(v, u))^n dv$$

is finite for  $n = 1, 2$ . The values of the mean and variance parameters  $\mu_\Gamma$  and  $\sigma_\Gamma^2$  depend on the Lévy basis, as summarised in Table C.1. Depending on the choice of the Lévy basis,  $X(u)$  might not be positive and thus may not be usable for determining distances. We then use the cut-off version  $X_c(u) = \max\{c, X(u)\}$  for  $c > 0$ , which generates the random particle  $Y_c$ . For further discussion of the relevant properties of Lévy bases see Hellmund et al. (2008).

Note that  $X$  can be interpreted as a stochastic process on the sphere (Jones, 1963), whose covariance function is given by

$$\text{Cov}(X(u_1), X(u_2)) = \sigma_\Gamma^2 \int_{\mathbb{S}^2} k(\vartheta(v, u_1)) k(\vartheta(v, u_2)) dv, \quad u_1, u_2 \in \mathbb{S}^2,$$

**Table C.1:** Mean and variance parameters for Gaussian and gamma bases for the Lévy particle process.

Lévy basis $\Gamma$	$\mu_\Gamma$	$\sigma_\Gamma^2$
Normal	$\mu$	$\sigma^2$
Gamma	$\kappa/\tau$	$\kappa/\tau^2$

Under an isotropic kernel, the random field  $X$  is stationary and isotropic, and it is readily seen that  $\text{Cor}(X(u_1), X(u_2)) = C(\vartheta(u_1, u_2))$ , where

$$C(\theta) = \frac{2}{c_2} \int_0^\pi \int_0^\pi k(\theta) k(\arccos(\sin \theta \sin \eta \cos \phi + \cos \theta \cos \eta)) \, d\phi \sin \eta \, d\eta, \quad (\text{C.5})$$

for  $0 \leq \theta \leq \pi$ , is the correlation function of the random field  $X$ . In particular, the correlation structure does not depend on the choice of the Lévy basis.

### C.3 Hausdorff dimension

The Hausdorff dimension of a set  $Z \subset \mathbb{R}^d$  is defined as follows (Hausdorff, 1918; Falconer, 1990). For  $\varepsilon > 0$ , an  $\varepsilon$ -cover of  $Z$  is a finite or countable collection  $\{B_i : i = 1, 2, \dots\}$  of balls  $B_i \subset \mathbb{R}^d$  of diameter  $|B_i|$  less than or equal to  $\varepsilon$  that covers  $Z$ . With

$$H^\delta(Z) = \lim_{\varepsilon \rightarrow 0} \inf \left\{ \sum |B_i|^\delta : \{B_i : i = 1, 2, \dots\} \text{ is an } \varepsilon\text{-cover of } Z \right\}$$

denoting the  $\delta$ -dimensional Hausdorff measure of  $Z$ , there exists a unique nonnegative number  $\delta_0$  such that  $H^\delta(Z) = \infty$  if  $\delta < \delta_0$  and  $H^\delta(Z) = 0$  if  $\delta > \delta_0$ . This number  $\delta_0$  is the Hausdorff dimension of the set  $Z$ .

For the remainder of the section, we assume that the Lévy basis  $\Gamma$  is the Gaussian basis (C.3). Then  $X$  has Gaussian finite dimensional distributions and thus is a Gaussian process. While there is a wealth of results on the Hausdorff dimension of the graphs of stationary Gaussian random fields on Euclidean spaces, which is determined by the infinitesimal behaviour of the correlation function at the origin, as formalised by the fractal index (Hall and Roy, 1994; Adler, 1981), we are unaware of any extant results for the graphs of random fields on spheres, or for the surfaces of star-shaped random particles.

We now state and prove such a result. Toward this end, we define the fractal index of an isotropic random field  $X$  on the sphere with associated correlation function (C.5) as the number

$$\begin{aligned} \alpha &= \sup \left\{ \beta > 0 : C(0) - C(\theta) = o(\theta^\beta), \theta \downarrow 0 \right\} \\ &= \inf \left\{ \beta > 0 : \theta^\beta = o(C(0) - C(\theta)), \theta \downarrow 0 \right\}. \end{aligned} \quad (\text{C.6})$$

The fractal index exists for essentially all correlation functions of practical interest, and it is always true that  $\alpha \in (0, 2]$ . The following theorem relates the

Hausdorff dimension of the graph of an isotropic Gaussian random field  $X$  on the sphere  $\mathbb{S}^2$  to its fractal index. The proof employs stereographic projections that allow us to draw on classical results in the Euclidean case.

**Theorem C.1.** *Let  $X$  be an isotropic Gaussian random field on  $\mathbb{S}^2$  with fractal index  $\alpha \in (0, 2]$ . Consider the random surface*

$$Z_c = \{ (u, X_c(u)) : u \in \mathbb{S}^2 \},$$

where  $X_c(u) = \max\{c, X(u)\}$  with  $c > 0$ . Then with probability one either of the following alternatives holds:

- (a) *If  $\max_{u \in \mathbb{S}^2} X(u) \leq c$ , the realisation of  $Z_c$  is the sphere with radius  $c$  and so its Hausdorff dimension is 2.*
- (b) *If  $\max_{u \in \mathbb{S}^2} X(u) > c$ , the realisation of  $Z_c$  has Hausdorff dimension  $3 - \frac{\alpha}{2}$ .*

*Proof.* The claim in alternative (a) is trivial. To prove the statement in alternative (b), we assume without loss of generality that  $X(u_0) > c$ , where  $u_0 = (0, 0, 1)$ . The sample paths of  $X$  are continuous almost surely according to Gangolli (1967, Theorem 7.2). Thus, there exists  $\varepsilon \in (0, 1)$  such that  $X(u) > c$  for  $u$  in the spherical cap  $\mathbb{S}_\varepsilon^2 = \{u \in \mathbb{S}^2 : \vartheta(u, u_0) \leq \varepsilon\}$  of radius  $\varepsilon$  centred at  $u_0$ . Let  $\Pi : \mathbb{S}_\varepsilon^2 \rightarrow B_\varepsilon$  denote a stereographic projection that maps  $(0, 0, 1)$  to  $(0, 0)$ , where  $B_\varepsilon = \{x = (x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq \varepsilon^2\}$ . A stereographic projection is a local diffeomorphism,  $\Pi$  is thus differentiable and has a differentiable inverse  $\Pi^{-1}$ , which is locally bi-Lipschitz (do Carmo, 1976). We may therefore assume that  $\varepsilon$  is small enough so that for all  $x, x' \in B_\varepsilon$  there exists a constant  $A \geq 1$  with

$$\frac{1}{A} \|x - x'\| \leq \|\Pi^{-1}(x) - \Pi^{-1}(x')\| \leq A \|x - x'\|, \quad (\text{C.7})$$

where  $\|\cdot\|$  denotes the Euclidean norm on  $\mathbb{R}^2$  or  $\mathbb{R}^3$ , respectively. Let the Gaussian random field  $W$  on  $B_\varepsilon \subset \mathbb{R}^2$  be given by  $W(x) = X(\Pi^{-1}(x))$ . From Xue and Xiao (2009, Theorem 5.1) we get that the graph  $\text{Gr } W = \{(x, W(x)) : x \in B_\varepsilon\}$  has Hausdorff dimension  $3 - \frac{\alpha}{2}$  almost surely if there exists positive constants  $\delta_0, M_0 \geq 1$  such that for all  $x, x' \in B_\varepsilon$  with  $\|x - x'\| \leq \delta_0$

$$\frac{1}{M_0} \sum_{j=1}^2 |x_j - x'_j|^\alpha \leq \mathbb{E}(W(x) - W(x'))^2 \leq M_0 \sum_{j=1}^2 |x_j - x'_j|^\alpha. \quad (\text{C.8})$$

Let  $x, x' \in B_\varepsilon$  be such that  $x \neq x'$ , and let  $\vartheta(x, x') = \vartheta(\Pi^{-1}(x), \Pi^{-1}(x'))$ . Then it holds that  $\mathbb{E}|W(x) - W(x')|^2 = 2\sigma^2 c_2 [C(0) - C(\vartheta(x, x'))]$ . As chord length and great circle distance are bi-Lipschitz equivalent metrics, there exists a constant  $c \geq 1$  such that

$$\frac{1}{c} \|\Pi^{-1}(x) - \Pi^{-1}(x')\| \leq \vartheta(x, x') \leq c \|\Pi^{-1}(x) - \Pi^{-1}(x')\|. \quad (\text{C.9})$$

Let  $\tilde{\varepsilon} > 0$  be given, and let  $\beta < \alpha$  and  $\tau > \alpha$  be such that  $\alpha + \tilde{\varepsilon} > \tau$  and  $\alpha - \tilde{\varepsilon} < \beta$ . As  $X$  is of fractal index  $\alpha$ , there exists for every  $K > 0$  a  $\delta > 0$

such that  $\vartheta(x, x') < \delta$  implies  $C(0) - C(\vartheta(x, x')) \leq K\vartheta(x, x')^\beta$  for  $\beta < \alpha$ , and  $\vartheta(x, x')^\tau \leq K[C(0) - C(\vartheta(x, x'))]$  for  $\tau > \alpha$ . As the mapping  $x \mapsto x^{\alpha/2}$  is concave, we may apply Jensen's inequality, equations (C.7) and (C.9) and the considerations above to show that

$$\sum_{j=1}^2 |x_j - x'_j|^\alpha \leq 2^{1-\alpha/2} \sigma^2 c_2 A^\alpha C^\alpha K \vartheta(x, x')^{-\tilde{\varepsilon}} [C(0) - C(\vartheta(x, x'))].$$

Let  $M_0 = 2^{-\alpha/2} A^\alpha C^\alpha K$ . For  $\tilde{\varepsilon} \rightarrow 0$ , we further obtain

$$\frac{1}{M_0} \sum_{j=1}^2 |x_j - x'_j|^\alpha \leq 2\sigma^2 c_2 [C(0) - C(\vartheta(x, x'))],$$

where we assume that  $\vartheta(x, x') \leq 1$ , which can be guaranteed by selecting  $\mathbb{S}_\varepsilon^2$  sufficiently small. Similarly,

$$2\sigma^2 c_2 [C(0) - C(\vartheta(x, x'))] \leq M_0 \sum_{j=1}^2 |x_j - x'_j|^\alpha,$$

which concludes the proof of the inequalities in (C.8).

Now, consider the mapping  $\zeta$  from  $B_\varepsilon \times \mathbb{R}$  to  $\mathbb{S}_\varepsilon^2 \times \mathbb{R}$  defined by  $\zeta(x, r) = (\Pi^{-1}(x), r)$ , so that  $\zeta(\text{Gr } W) = \{(u, X(u)) : u \in \mathbb{S}_\varepsilon^2\}$ . The identity

$$\|\zeta(x, r) - \zeta(x', r')\|^2 = \|\Pi^{-1}(x) - \Pi^{-1}(x')\|^2 + |r - r'|^2.$$

along with (C.7) implies  $\zeta$  to be bi-Lipschitz. Therefore by Corollary 2.4 of Falconer (1990), the partial surface  $\{(u, X(u)) : u \in \mathbb{S}_\varepsilon^2\}$  has Hausdorff dimension  $3 - \alpha/2$  almost surely. Invoking the countable stability property of the Hausdorff dimension (Falconer, 1990, p. 29), we see that the full surface  $Z_c = \{(u, X_c(u)) : u \in \mathbb{S}_\varepsilon^2\}$  also has Hausdorff dimension  $3 - \alpha/2$  almost surely.  $\square$

## C.4 Isotropic kernels

It is often important that the surface of the particle process possesses the same Hausdorff dimension as that of the real-world particles to be emulated (Mandelbrot, 1982; Orford and Whalley, 1983). With this in mind, we introduce and study three one-parameter families of isotropic kernels for the Lévy particle process (C.1). The families yield interesting second order structures, and we derive the asymptotic behaviour of their correlation functions at zero, which determines the Hausdorff dimension of the Gaussian particle surface.

### C.4.1 Von Mises–Fisher kernel

Here, we consider  $k$  to be the unnormalised von Mises–Fisher density,

$$k(\theta) = \exp\{a \cos \theta\}, \quad 0 \leq \theta \leq \pi, \quad (\text{C.10})$$

with parameter  $a \in \mathbb{R}$ . The von Mises–Fisher density with parameter  $a > 0$  is widely used in the analysis of spherical data (Fisher et al., 1987), and in this context  $a$  is called the precision.

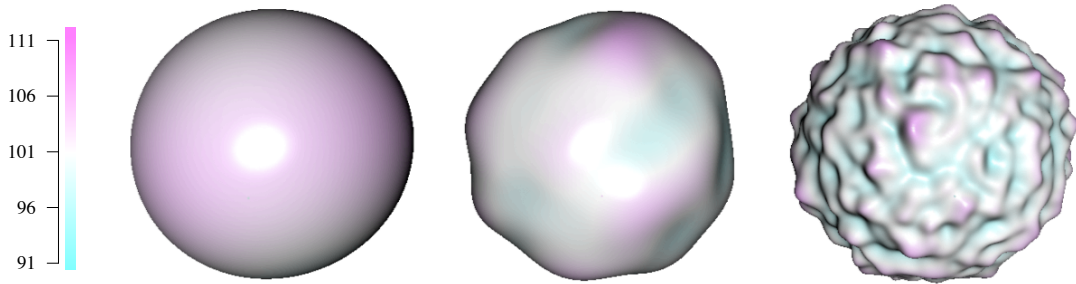
To obtain a closed form expression for the correlation function, we consider an alternative to the representation in equation (C.5). For  $\theta \in [0, \pi]$  let  $u_\theta = (\sin \theta, 0, \cos \theta) \in \mathbb{S}^2$ . Then

$$\begin{aligned} C(\theta) &= \frac{2}{c_2} \int_{\mathbb{S}^2} k(\vartheta(v, u_0)) k(\vartheta(v, u_\theta)) \, dv \\ &= \frac{2}{c_2} \int_{\mathbb{S}^2} \exp\left\{a \|u_0 + u_\theta\| \left(v \cdot \frac{u_0 + u_\theta}{\|u_0 + u_\theta\|}\right)\right\} \, dv \\ &= \frac{4}{c_2 \pi} \int_0^\pi \exp\{a \|u_0 + u_\theta\| \cos \theta\} \sin \theta \, d\theta \\ &= \frac{8\pi}{c_2} \frac{\sinh(a \|u_0 + u_\theta\|)}{a \|u_0 + u_\theta\|}. \end{aligned}$$

As  $\|u_1 + u_2\|^2 = 2[1 + \cos \vartheta(u_1, u_2)]$  for  $u_1, u_2 \in \mathbb{S}^2$ , we get

$$C(\theta) = \frac{2}{\sinh(2a)} \frac{\sinh(a \sqrt{2(1 + \cos \theta)})}{\sqrt{2(1 + \cos \theta)}}, \quad 0 \leq \theta \leq \pi, \quad (\text{C.11})$$

from which it is readily seen that the fractal index is  $\alpha = 2$ . The surfaces of the corresponding Gaussian Lévy particles are smooth and have Hausdorff dimension 2, independently of the value of the parameter  $a \in \mathbb{R}$ , as illustrated in Figure C.2.



**Figure C.2:** Lévy particles with mean  $\mu_X = 100$  and variance  $\sigma_X^2 = 10$ , using a gamma Lévy basis and the von Mises–Fisher kernel (C.10) with  $a = 3$  (left),  $a = 30$  (middle) and  $a = 300$  (right).

Perhaps surprisingly, the von Mises–Fisher correlation function (C.11) can be extended to a stationary and isotropic correlation function on the Euclidean

space  $\mathbb{R}^d$  for any  $d \geq 3$ . To see this, we define the function

$$\varphi(t) = \frac{{}_1F_1\left(1, \frac{3}{2}, a^2\left(1 - \frac{t}{4}\right)\right)}{{}_1F_1\left(1, \frac{3}{2}, a^2\right)}, \quad t \geq 0,$$

in terms of the classical confluent hypergeometric function. Using well-known results (Abramowitz and Stegun, 1965, pp. 505 and 507), we see that

$$(-1)^n \frac{\partial^n}{\partial t^n} \varphi(t) = \frac{a^{2n}}{4^n} \frac{(1)_n}{{(\frac{3}{2})}_n} \frac{{}_1F_1\left(n+1, n+\frac{3}{2}, a^2\left(1 - \frac{t}{4}\right)\right)}{{}_1F_1\left(1, \frac{3}{2}, a^2\right)} \geq 0, \quad t \geq 0,$$

where  $(x)_n = x(x+1)\cdots(x+n)$ , whence  $\varphi$  is completely monotone. By the key result of Schoenberg (1938), the function  $\psi(x) = \varphi(\|x\|^2)$ ,  $x \in \mathbb{R}^d$  is a correlation function in any dimension  $d \geq 1$ . Restricting  $\psi$  to  $\mathbb{S}^2$  and substituting spherical distance for Euclidean distance yields the von Mises–Fisher correlation function (C.11) with parameter  $a \in \mathbb{R}$ , thereby proving the claim.

### C.4.2 Uniform kernel

We now let the kernel  $k$  be uniform, in that

$$k(\theta) = \mathbb{1}_{\{\theta \leq r\}}, \quad 0 \leq \theta \leq \pi, \quad (\text{C.12})$$

with cut-off parameter  $r \in (0, \pi/2)$ . As shown in the appendix of Tovchigrechko and Vakser (2001), the associated correlation function is

$$C(\theta) = \frac{1}{\pi(1 - \cos r)} \left( \pi - \arccos(\cos \theta \csc^2 r - \cot^2 r) - 2 \cos r \arccos(\csc \theta \cos r \csc r - \cot \theta \cot r) \right) \mathbb{1}_{\{\theta \leq 2r\}}, \quad 0 \leq \theta \leq \pi. \quad (\text{C.13})$$

In particular, if  $r = \pi/2$  then  $C(\theta) = 1 - \theta/\pi$  decays linearly throughout. Taylor expansions imply that the correlation function has fractal index  $\alpha = 1$  for all  $r \in (0, \pi/2)$ , so that the corresponding Gaussian Lévy particles have non-smooth boundaries of Hausdorff dimension  $5/2$ . Examples of Gaussian Lévy particles under this kernel are shown in Figure C.3.

### C.4.3 Power kernel

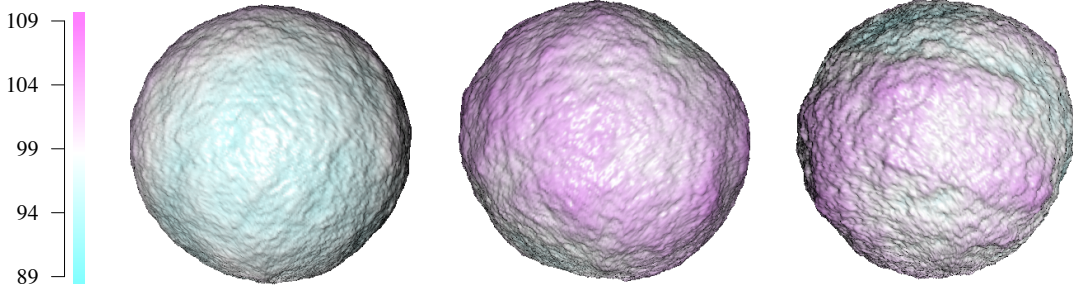
We now introduce the power kernel, which allows for Lévy particles with boundaries of any desired Hausdorff dimension. Specifically, let the isotropic kernel  $k$  be defined as

$$k(\theta) = (\theta/\pi)^{-q} - 1, \quad 0 < \theta \leq \pi, \quad (\text{C.14})$$

with power parameter  $q \in (0, 1)$ . The associated correlation function (C.5) takes the form

$$C(\theta) = \frac{1}{c_2} \int_0^\pi (\pi^q \lambda^{-q} - 1) \sin \lambda \int_{A(\lambda)} (\pi^q a(\theta, \lambda, \phi)^{-q} - 1) d\phi d\lambda, \quad (\text{C.15})$$





**Figure C.3:** Lévy particles with mean  $\mu_X = 100$  and variance  $\sigma_X^2 = 10$ , using a Gaussian Lévy basis and the uniform kernel (C.12) with  $r = 1.5$  (left),  $r = 1.0$  (middle) and  $r = 0.5$  (right). The fractal or Hausdorff dimension of the particle surfaces equals 2.5.

where

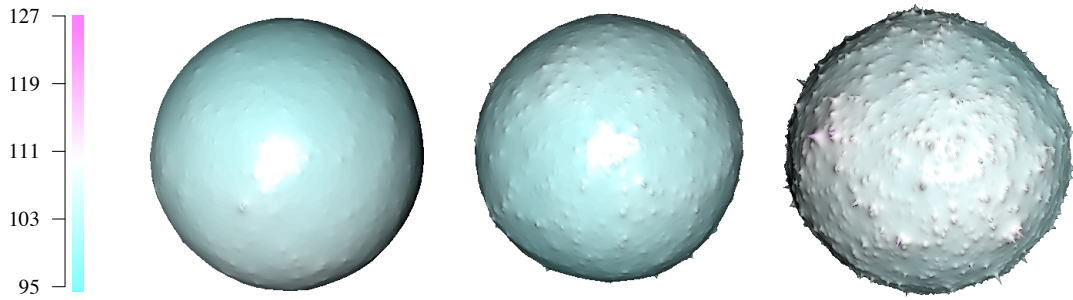
$$a(\theta, \lambda, \phi) = \arccos(\sin \theta \sin \lambda \cos \phi + \cos \theta \cos \lambda) \quad (\text{C.16})$$

and

$$A(\lambda) = \{ \phi \in [0, \pi] : 0 < a(\theta, \lambda, \phi) \leq \pi \}. \quad (\text{C.17})$$

Figures C.1 and C.4 show Lévy particles under the power kernel using Gaussian and gamma bases, respectively. The surface structure for the different bases resemble each other, even though the particles exhibit more pronounced spikes under the gamma basis.

Our next result shows the asymptotic behaviour of the correlation function in (C.15). Provided that the integral  $A_q$  in (C.18) is strictly positive, the correlation function has fractal index  $\alpha = 2 - 2q$ , so that the corresponding Gaussian Lévy particles illustrated in Figure C.1 have surfaces with Hausdorff dimension  $2 + q$ .



**Figure C.4:** Lévy particles with mean  $\mu_X = 100$  and variance  $\sigma_X^2 = 10$ , using a gamma Lévy basis and the power kernel (C.14) with  $q = 0.05$  (left),  $q = 0.25$  (middle) and  $q = 0.5$  (right).

**Theorem C.2.** *Let  $0 < q < 1$  and*

$$A_q = 2\pi^{2q} \int_0^\infty x^{1-q} \int_0^\pi (x^{-q} - (x^2 + 1 - 2x \cos \phi)^{-q/2}) d\phi dx \in [0, \infty). \quad (\text{C.18})$$

Provided that  $A_q > 0$ , the correlation function (C.15) admits the expansion

$$C(0) - C(\theta) \sim A_q \theta^{2-2q} \quad \text{as } \theta \downarrow 0. \quad (\text{C.19})$$

In particular, the correlation function has fractal index  $\alpha = 2 - 2q$ .

*Proof.* Let  $0 \leq \theta \leq \pi$ , and let the functions  $a$  and  $A$  be defined in (C.16) and (C.17). From (C.5), we get

$$\begin{aligned} \frac{1}{2}[C(0) - C(\theta)] &= \int_0^\pi (\pi^q \lambda^{-q} - 1) \sin \lambda \\ &\quad \times \left\{ \int_0^\pi (\pi^q \lambda^{-q} - 1) d\phi - \int_{A(\lambda)} (\pi^q a(\theta, \lambda, \phi)^{-q} - 1) d\phi \right\} d\lambda. \end{aligned}$$

Since  $A(\lambda) = [0, \pi]$  for  $\lambda \in (0, \pi - \theta]$  and  $A(\lambda) \subset [0, \pi]$  for  $\lambda \in (\pi - \theta, \pi)$ , we decompose the integral on the right-hand side as  $I_{1q}(\theta) + I_{2q}(\theta)$ , where  $I_{1q}(\theta)$  and  $I_{2q}(\theta)$  correspond to the integral with respect to  $\lambda$  over  $(0, \pi - \theta)$  and  $(\pi - \theta, \pi)$ , respectively.

As for the first term, substituting  $\lambda = \theta x$  yields

$$\begin{aligned} I_{1q}(\theta) &= \pi^{2q} \int_0^{\pi-\theta} (\lambda^{-q} - \pi^{-q}) \sin \lambda \\ &\quad \times \left\{ \int_0^\pi (\lambda^{-q} - \pi^{-q}) d\phi - \int_0^\pi (a(\theta, \lambda, \phi)^{-q} - \pi^{-q}) d\phi \right\} d\lambda \\ &= \theta^{2-2q} \pi^{2q} \int_0^{\frac{\pi-\theta}{\theta}} \frac{\sin(\theta x)}{\theta} (x^{-q} - \pi^{-q} \theta^q) \int_0^\pi (x^{-q} - a(\theta, \theta x, \phi)^{-q} \theta^q) d\phi dx. \end{aligned}$$

Noting that

$$\frac{\arccos(t)}{\theta} = \frac{\arccos(1 - y^2)}{y} \frac{y}{\theta} \Big|_{y=(1-t)^{1/2}}$$

for  $t \in (0, 1)$ , we find from (C.16) that

$$\begin{aligned} \lim_{\theta \downarrow 0} \frac{a(\theta, \theta x, \phi)}{\theta} &= \lim_{\theta \downarrow 0} \frac{\arccos(\sin \theta \sin(\theta x) \cos \phi + \cos \theta \cos(\theta x))}{\theta} \\ &= \frac{d}{dy} \arccos(1 - y^2) \Big|_{y=0} \lim_{\theta \downarrow 0} \left( \frac{1 - \cos \theta \cos(\theta x)}{\theta^2} - \frac{\sin \theta \sin(\theta x)}{\theta^2} \cos \phi \right)^{1/2} \\ &= \sqrt{2} \left( \frac{1 + x^2}{2} - x \cos \phi \right)^{1/2} \\ &= (x^2 + 1 - 2x \cos \phi)^{1/2}. \end{aligned}$$

An application of the Lebesgue Dominated Convergence Theorem with the dominating function

$$h(\phi, x) = \begin{cases} x^{1-q}(x^{-q} + 1) \left( \left( \frac{x}{1-x} \right)^q + 1 \right), & 0 < x < 1, \\ x^{1-q}(x^{-q} + 1) \left( \left( \frac{x}{x-1} \right)^q - 1 \right), & x > 1, \end{cases}$$

now shows that

$$\begin{aligned} I_{1q}(\theta) &= \theta^{2-2q} \pi^{2q} \int_0^{\frac{\pi-\theta}{\theta}} \frac{\sin(\theta x)}{\theta} (x^{-q} - \pi^{-q} \theta^q) \int_0^\pi (x^{-q} - a(\theta, \theta x, \phi)^{-q} \theta^q) d\phi dx \\ &\sim \theta^{2-2q} \pi^{2q} \int_0^\infty x^{1-q} \int_0^\pi (x^{-q} - (x^2 + 1 - 2x \cos \phi)^{-q/2}) d\phi dx = A_q \theta^{2-2q} \end{aligned}$$

as  $\theta \downarrow 0$ , with the constant  $A_q$  of equation (C.18), provided that  $A_q > 0$ .

As regards the second term, the first mean value theorem for integration implies that there exists a  $t \in (\pi - \theta, \pi)$  such that

$$\begin{aligned} I_{2q}(\theta) &= \int_{\pi-\theta}^\pi (\pi^q \lambda^{-q} - 1) \sin \lambda \\ &\quad \times \left\{ \int_0^\pi (\pi^q \lambda^{-q} - 1) d\phi - \int_{A(\lambda)} (\pi^q a(\theta, \lambda, \phi)^{-q} - 1) d\phi \right\} d\lambda \\ &= \theta (\pi^q t^{-q} - 1) \sin t \left\{ \int_0^\pi (\pi^q t^{-q} - 1) d\phi - \int_{A(t)} (\pi^q a(\theta, t, \phi)^{-q} - 1) d\phi \right\}. \end{aligned}$$

Hence,  $I_{2q}(\theta)$  decays faster than  $O(\theta^2)$  as  $\theta \downarrow 0$ , which completes the proof of the asymptotic expansion (C.19). The claim about the fractal index then is immediate from Theorem C.1.  $\square$

The power kernel (C.14) has a negative exponent and thus is unbounded. While positive exponents are feasible, they are of less interest, as the associated correlation functions have fractal index  $\alpha = 2$ , thereby generating smooth particles only.

## C.5 Examples

Here, we demonstrate the flexibility of the Lévy particle framework in simulation examples. First, we introduce a simulation algorithm. Then we simulate celestial bodies whose surface properties resemble those of Earth, Moon, Mars and Venus, as reported in the planetary physics literature. Furthermore, we study and simulate the planar Lévy particles that arise from the two-dimensional version of the three-dimensional Lévy particle model (C.1).

### C.5.1 Simulation algorithm

To sample from the Lévy particle model (C.1), we utilise the property that a Lévy basis is independently scattered. Specifically, for every sequence  $(A_n)$  of disjoint Borel subsets of  $\mathbb{S}^2$ , the random variables  $\Gamma(A_n)$ ,  $n = 1, 2, \dots$ , are independent and  $\Gamma(\cup A_n) = \sum \Gamma(A_n)$  almost surely (Jónsdóttir et al., 2008). Let  $(A_n)_{n=1}^N$  denote an equal area partition of  $\mathbb{S}^2$ , so that  $\lambda(A_n) = 4\pi/N$  for  $n = 1, \dots, N$ .

The random field  $X$  in (C.1) can then be decomposed into a sum of integrals over the disjoint sets  $A_n$ , in that

$$X(u) = \sum_{n=1}^N \int_{A_n} k(v, u) \Gamma(\mathrm{d}v), \quad u \in \mathbb{S}^2.$$

For  $n = 1, \dots, N$  fix any point  $v_n \in A_n$ . We can then approximate the random field  $X$  by setting

$$x(u) = \sum_{n=1}^N k(v_n, u) \Gamma(A_n), \quad u \in \mathbb{S}^2.$$

Let us denote the common distribution of  $\Gamma(A_1), \dots, \Gamma(A_N)$ , which derives either (C.3) or (C.4), depending on the choice of the Lévy basis, by  $F_N$ . To simulate a realisation  $y$  of the Lévy particle  $Y$  in (C.2), we use the following algorithm.

**Algorithm C.1.**

1. Set  $M = m_1 m_2$ , where  $m_1$  and  $m_2$  are positive integers, and construct a grid  $u_1, \dots, u_M$  on  $\mathbb{S}^2$ . Using spherical coordinates, let  $u_m = (\theta_m, \phi_m)$  and put  $\theta_m = i\pi/m_1$  and  $\phi_m = 2\pi j/m_2$ , where  $m = im_2 + j$  for  $i = 0, 1, \dots, m_1 - 1$  and  $j = 1, \dots, m_2$ .
2. Apply the method of Leopardi (2006) to construct an equal area partition  $A_1, \dots, A_N$  of  $\mathbb{S}^2$ .
3. For  $n = 1, \dots, N$ , let  $v_n$  have spherical coordinates equal to the mid range of the latitudes and longitudes within  $A_n$ .
4. For  $n = 1, \dots, N$ , generate independent random variables  $\Gamma_n$  from  $F_N$ .
5. For  $m = 1, \dots, M$ , set  $x(u_m) = \sum_{n=1}^N k(v_n, u_m) \Gamma_n$ .
6. Set  $y$  to be the convex hull of  $\{(u_m, x(u_m)) : m = 1, \dots, M\}$ .

When a Gaussian Lévy basis is used, we use the aforementioned modification, in which  $x(u_m)$  is replaced by  $\max\{c, x(u_m)\}$  for some  $c > 0$ . This simulation procedure has been implemented in R (R Development Core Team, 2009), and code is available from the authors upon request. It can be considered an analogue of the moving average method (Oliver, 1995; Cressie and Pavlicová, 2002) for simulating Gaussian random fields on Euclidean spaces. The quality of the realisations depends on the choice of  $M$  and  $N$ , and the usual trade-off applies, in that large values of  $M$  and  $N$  result in accurate simulations, at the expense of prolonged computation times.

For the simulations in Figure 1–4, we used  $m_1 = 200$ ,  $m_2 = 400$  and  $N = 10^5$ .

## C.5.2 Celestial bodies

In the literature, there are several articles concerning the Hausdorff dimension of the surface of the Earth and the other celestial bodies in our solar system,

see e.g. Mandelbrot (1982), Turcotte (1987) and Kucinskas et al. (1992). Here, we provide simulated version of the planets Earth, Venus, Mars, and of the Moon. According to Turcotte (1987) the Hausdorff dimension of the surface of these bodies is 2.5. We will assume the model (C.1) with  $k$  being the uniform kernel (C.12) and  $\Gamma$  being the Gaussian basis (C.3). We set  $r = \pi/3$  which gives the desired theoretical Hausdorff dimension 2.5. For this value of  $r$  we get  $c_n = 2\pi(1 - \cos(\pi/3)) = \pi$ , for  $n = 1, 2$ .

Further, we attempt to fix the parameter values of the Lévy basis such that the mean radius and the variation thereof is in accordance with reality. For this we use the information listed in Table C.2 obtained from Seidelmann et al. (2007), Wieczorek et al. (2006), Wikipedia<sup>1</sup>, The Mariana Trench<sup>2</sup>, Price (1988), Taylor (1998), Wikipedia<sup>3</sup>, geology.com<sup>4</sup>, Jones and Stofan (2008), Infoplease<sup>5</sup>. For each celestial body, we aim to have  $EX(u) = r_0$ . Hence, we fix the mean value parameter of the Lévy basis as  $\mu_\Gamma = r_0/c_1$ . As a heuristic measure of the deviation in the radius of a body we fix the variance parameter of the basis at  $\sigma_\Gamma^2 = (d_+ - d_-)/c_2$ .

Note that the values concerning Earth describe the 'Dry Earth'; that is, they do not take the oceans into account. From information on The Science Forum<sup>6</sup> we calculate the mean radius of Dry Earth from the radius of Earth with water (6371 km), the volume of water on Earth ( $1.3 \times 10^9 \text{ km}^3$ ), and the average ocean depth (3.79 km). To simulate Earth with water we first simulate the radial function for the Dry Earth. Then, we make a cut-off such that the oceans constitute 70.8 % of the surface of Earth (Pidwirny, 2006). Corresponds to  $Y_c$  with  $c = 6371$ .

We let  $m_1 = 200$  and  $m_2 = 400$  so that  $M = 80\,000$ , and the equal area partition used for the simulation has  $N = 10^6$ . The resulting celestial bodies are shown in Figure C.5, and the corresponding radial functions along the equator are displayed in Figure C.6.

### C.5.3 Planar Lévy particles

For ease of comparison, we reduce the dimension of the Lévy particle process (C.1) to two. This allows us to better assess the influence of the kernel and the basis choice on the resulting particles. Thus, we consider the planar random particle

$$Y_c = \bigcup_{u \in \mathbb{S}^1} \{ o + ru : 0 \leq r \leq \max(X(u), c) \} \subset \mathbb{R}^2,$$

<sup>1</sup> [http://en.wikipedia.org/wiki/Mount\\_Everest#cite\\_note-tas1982-8](http://en.wikipedia.org/wiki/Mount_Everest#cite_note-tas1982-8)

<sup>2</sup> [http://www.marianatrench.com/mariana\\_trench-oceanography.htm](http://www.marianatrench.com/mariana_trench-oceanography.htm)

<sup>3</sup> [http://en.wikipedia.org/wiki/Geology\\_of\\_Mars#cite\\_note-100](http://en.wikipedia.org/wiki/Geology_of_Mars#cite_note-100)

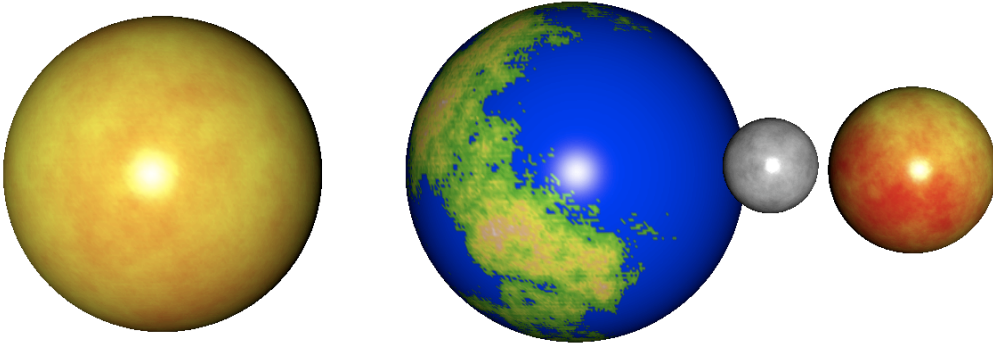
<sup>4</sup> <http://geology.com/articles/highest-point-on-mars.shtml#mars-low-point>

<sup>5</sup> <http://www.infoplease.com/ipa/A0004448.html>

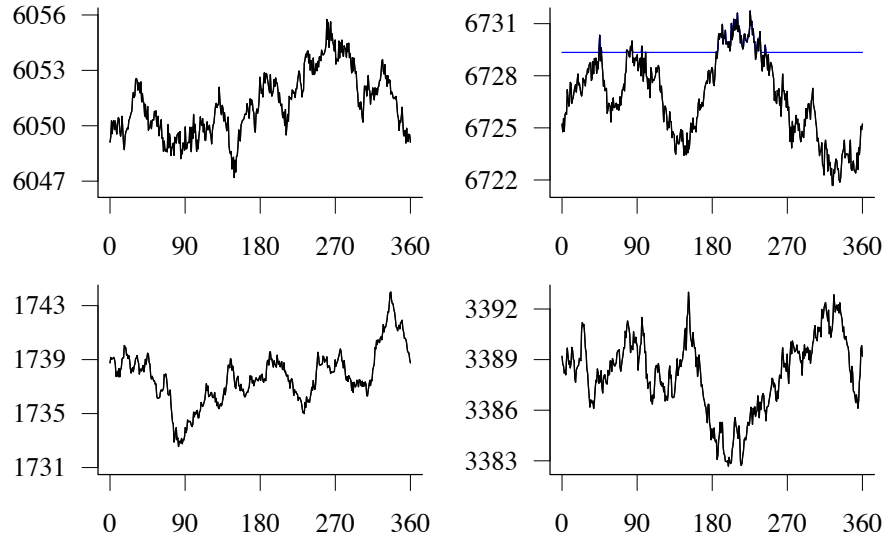
<sup>6</sup> <http://www.thescienceforum.com/viewtopic.php?p=232069>

**Table C.2:** Mean radius  $r_0$ , difference  $d_+$  between maximal and mean radius, and difference  $d_-$  between minimal and mean radius, for Venus, Dry Earth, Moon and Mars, in kilometres.

Body	Venus	Dry Earth	Moon	Mars
$r_0$	6051.8	6727.2	1737.1	3389.5
$d_+$	11.0	8.8	5.5	21.2
$d_-$	-3.0	-11.0	-12.0	-8.2



**Figure C.5:** Simulations of Venus, Earth, Moon and Mars in true relative size.



**Figure C.6:** Radial function along the equator for the simulated bodies in Figure C.5 in kilometres. Clockwise from upper left: Venus, Earth with ocean level at the blue line, Mars, and Moon.

where  $c > 0$  and  $o \in \mathbb{R}^2$  is an arbitrary centre, and the radial function  $X(u)$  is modelled as

$$X(u) = \int_{\mathbb{S}^1} K(v, u) \Gamma(\mathrm{d}v), \quad u \in \mathbb{S}^1,$$

with a suitable kernel function  $K : \mathbb{S}^1 \times \mathbb{S}^1 \rightarrow \bar{\mathbb{R}}$  and a Lévy basis  $\Gamma$  on the Borel subsets of the unit sphere  $\mathbb{S}^1 := \{x \in \mathbb{R}^2 : \|x\| = 1\}$ . As previously, we assume that the kernel function  $K$  is isotropic, in that  $K(v, u) = k(\vartheta(v, u))$  depends on the points  $v, u \in \mathbb{S}^1$  through their angular or circular distance  $\vartheta(v, u) \in [0, \pi]$  only. Table C.3 lists circular analogues of von Mises–Fisher, uniform and power kernels along with analytic expressions for the integrals

$$c_n = \int_{\mathbb{S}^1} k(\vartheta(v, u))^n \mathrm{d}v = 2 \int_0^\pi k(\eta)^n \mathrm{d}\eta,$$

where  $n = 1, 2$ , and the fractal index  $\alpha$ , as defined in (C.6), of the associated correlation function, namely

$$C(\theta) = \frac{1}{c_2} \left( \int_{\pi-\theta}^\pi k(\phi)k(2\pi - \phi - \theta) \mathrm{d}\phi + \int_0^{\pi-\theta} k(\phi)k(\theta + \phi) \mathrm{d}\phi \right. \\ \left. + \int_0^\theta k(\phi)k(\theta - \phi) \mathrm{d}\phi + \int_\theta^\pi k(\phi)k(\phi - \theta) \mathrm{d}\phi \right), \quad 0 \leq \theta \leq \pi.$$

For the von Mises–Fisher kernel with parameter  $a$ , the correlation functions admits the closed form

$$C(\theta) = I_0(a\sqrt{2(1 + \cos \theta)}) / I_0(2a), \quad 0 \leq \theta \leq \pi,$$

where  $I_0$  denotes a Bessel function. For the uniform kernel with cut-off parameter  $r \in (0, \pi/2]$ , the correlation functions admits the closed form

$$C(\theta) = (1 - \theta/(2r)) \mathbb{1}_{\{\theta \leq 2r\}}, \quad 0 \leq \theta \leq \pi.$$

For the power kernel with parameter  $q \in (0, 1/2)$  we find,

$$C(\theta) = \frac{\pi^q}{c_2} \frac{1}{1-q} \left[ (2\pi - \theta)^{-q} \left\{ {}_2F_1(1-q, q; 2-q; \pi/(2\pi - \theta)) \right. \right. \\ \left. \left. - \pi^q (\pi - \theta)^{1-q} {}_2F_1(1-q, q; 2-q; (\pi - \theta)/(2\pi - \theta)) \right\} \right. \\ \left. + \pi^q \theta^{-q} (\pi - \theta)^{1-q} {}_2F_1(1-q, q; 2-q; -(\pi - \theta)/\theta) \right. \\ \left. + \pi (\pi - \theta)^{-q} {}_2F_1(1, q; 2-q; \pi/(\pi - \theta)) \right] \\ - \frac{2\pi}{c_2} \frac{1+q}{1-q}.$$

If  $\max_{u \in \mathbb{S}^1} X(u) > c$ , the boundary of the Gaussian Lévy particle  $Y_c$  has Hausdorff dimension  $D = 2 - \alpha/2$  almost surely. We omit the proofs of these results, as they are analogous to those of the respective results for  $\mathbb{S}^2$ .

**Table C.3:** Analytic form, parameter range, constants and fractal index associated with for parametric families of isotropic kernels  $k : [0, 2\pi) \rightarrow \mathbb{R}$  on the circle  $\mathbb{S}^1$ . Here,  $I_0$  denotes a modified Bessel function of the first kind of order zero.

Kernel	von Mises–Fisher	Uniform	Power
Analytic Form	$k(\theta) = \exp(a \cos \theta)$	$k(\theta) = \mathbb{1}_{\{\theta \leq r\}}$	$k(\theta) = (\theta/\pi)^{-q} - 1$
Parameter	$a > 0$	$r \in (0, \pi/2)$	$q \in (0, 1/2)$
$c_1$	$2\pi I_0(a)$	$2r$	$2\pi \frac{q}{1-q}$
$c_2$	$2\pi I_0(2a)$	$2r$	$4\pi \frac{q^2}{1-3q+2q^2}$
Fractal Index	2	1	$1-2q$

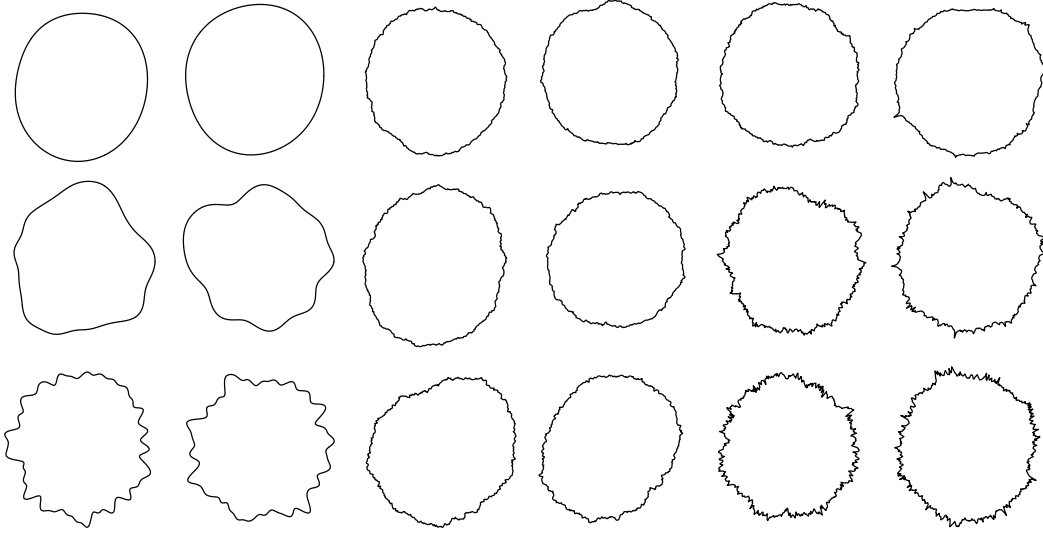
Figure C.7 shows simulated planar Lévy particles using Gaussian and gamma Lévy bases with von Mises–Fisher and power kernels, with parameters as listed in Table C.4. For the Gaussian basis, we further use a cut-off at zero. The simulation algorithm of Section C.5.1 continues to apply with natural adaptations, such as using a simulation grid  $u_m = 2\pi m/M$  for  $m = 1, \dots, M$ . We used  $M = 5000$  and  $N = 10^5$ .

**Table C.4:** Values of the parameter  $a$  for the von Mises–Fisher kernel and the parameter  $q$  for the power kernel used to generate the planar particles in Figure C.7.

Row	$a$	$r$	$q$
1	3	1.5	0.05
2	30	1.0	0.25
3	300	0.5	0.45

We see that the von Mises–Fisher kernel generates particles with a smooth, differentiable boundary of Hausdorff dimension 1, independently of the parameter  $a > 0$ . In contrast, the power kernel results in rough particles whose boundary is non-differentiable with Hausdorff dimension  $1.5 + q$ . Also, judging by the last row, it appears that the global particle shape under the von Mises–Fisher kernel and the uniform kernel can deviate more from a spherical shape than the global particle shape under the power kernel which seems to generate fairly spherical shapes regardless of the choice of Lévy basis. Furthermore, the plots indicate more pronounced oscillation of the radial function for the power kernel as the kernel parameter value increases and this seems to be independent of the choice of the Lévy basis. The gamma particles also appear more spiky than the Gaussian particles.





**Figure C.7:** Planar Lévy particles with mean  $\mu_X = 25$  and variance  $\sigma_X^2 = 10$ . Columns 1 and 2 show particles generated using a von Mises–Fisher kernel, columns 3 and 4 particles generated using a uniform kernel, and columns 5 and 6 particles generated using a power kernel with parameters varying by row as described in Table C.4. The particles in columns 1, 3, and 5 are generated under a Gaussian basis, those in columns 2, 4, and 6 under a gamma basis.

## C.6 Discussion

In this paper, we propose a flexible modelling framework for three-dimensional star-shaped particles. The particles are represented by their radial function which we model through a kernel smoothing of a Lévy basis on the sphere. We simulate particles under the model using both a gamma basis and a Gaussian basis. For the Gaussian particles, we show that a kernel function given by the unnormalised von Mises–Fisher density yields smooth particles with the Hausdorff dimension of the surface equal to that of the topological space. For the family of power kernels, on the other hand, we conjecture that we can obtain particles with boundaries of any Hausdorff dimension between 2 and 3.

The Hausdorff dimension of the surface is a function of the fractal index  $\alpha$  of the corresponding random field and  $\alpha$  can, in turn, be obtained from the asymptotic behaviour of the correlation function. Under the proposed model in (C.1) with an isotropic kernel function, the correlation function is a function of the great circle distance only and it does not depend on the choice of the Lévy base. Based on this and our simulation results which show only minor difference between the Gaussian particles and particles with a corresponding gamma base, we conjecture that results similar to Theorem C.1 will hold for a gamma Lévy base. However, such results are very difficult to obtain due to the lack of general theory on the Hausdorff dimension for non-Gaussian processes.

Could use non-isotropic kernels to get non-isotropic, and/or multi-fractal,

particles with spatially varying surface roughness and Hausdorff dimension. Gagnon et al. (2006) have argued that Earth is multi-fractal, with simulations on Planar Earth. Our work provides an avenue to multi-fractal simulations of Spherical Earth.

Our focus was on  $\mathbb{S}^2$ . Case of  $\mathbb{S}^1$  has been discussed in Section C.5.3, and a generalisation to  $\mathbb{S}^{d-1}$  for  $d \geq 2$  is possible. In this context, Estrade and Istas (2010) derive a recursive formula for the isotropic correlation function on  $\mathbb{S}^{d-1}$  that arises under a uniform kernel. In analogy to the terminology on Euclidean spaces (Gneiting, 1999), we might call this correlation function the spherical hat function with cut-off parameter  $r \in (0, \pi/2)$ . It has a linear behaviour at the origin and thus fractal index  $\alpha = 1$ . Estrade and Istas (2010) also show that scale mixtures of the spherical hat function provide correlation functions of any fractal index  $\alpha \in (0, 1]$ , similarly to the corresponding results of (Hammersley and Nelder, 1955) and (Gneiting, 1999) on Euclidean spaces.

We have not discussed inference under our modelling framework and consider this out of the scope of the present paper. In a Bayesian setting, inference could be performed in similar manner to the procedures proposed in Wolpert and Ickstadt (1998). Wolpert and Ickstadt model the intensity measure of a spatial point process using a construction similar to our random field model in (C.1). The observed data are then point counts which distribution is a function of the underlying intensity measure. They propose a simulated inference framework where the model parameters, the underlying random field, and the point process are updated in turn conditional on the current state of the other variables. A similar approach should be suitable here, where the observation is seen as a noisy version of the underlying radial function  $X$ .

## Acknowledgements

The authors thank Anders Rønn-Nielsen, Eva B. Vedel Jensen, Jens Ledet Jensen, and Werner Ehm for comments and discussions. This research has been supported by the Centre for Stochastic Geometry and Advanced Bioimaging, funded by a grant from the Villum Foundation, by the Alfried Krupp von Bohlen und Halbach Foundation, and by the German Research Foundation (DFG) within the programme “Spatio-/Temporal Graphical Models and Applications in Image Analysis”, grant GRK 1653.

## Bibliography

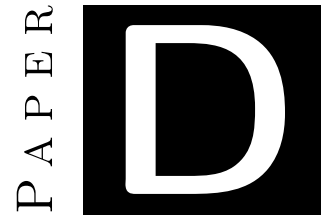
- Abramowitz, M. and Stegun, I. A. (1965). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York, ninth dover printing, tenth gpo printing edition.
- Adler, R. J. (1981). *The Geometry of Random Fields*. John Wiley and Sons, Chichester.

- Cressie, N. and Pavlicová, M. (2002). Calibrated spatial moving average simulations. *Statistical Modelling*, 2:267–279.
- Cruz-Orive, L. M. (1976). Particle size-shape distributions: the general spheroid problem. I. Mathematical model. *Journal of Microscopy*, 107:235–253.
- Cruz-Orive, L. M. (1978). Particle size-shape distributions: the general spheroid problem. II. Stochastic model and particle guide. *Journal of Microscopy*, 112:153–167.
- do Carmo, M. P. (1976). *Differential Geometry of Curves and Surfaces*. Prentice-Hall, Englewood Cliffs, NJ.
- Estrade, A. and Istas, J. (2010). Ball throwing on spheres. *Bernoulli*, 16:953–970.
- Falconer, K. (1990). *Fractal Geometry*. John Wiley and Sons, Chichester.
- Fisher, N. I., Embleton, B. J. J., and Lewis, T. (1987). *Statistical analysis of spherical data*. Cambridge University Press, Cambridge.
- Gagnon, J.-S., Lovejoy, S. L., and Schertzer, D. (2006). Multifractal earth topography. *Nonlinear Processes in Geophysics*, 13:541–570.
- Gangolli, R. (1967). Positive definite kernels on homogeneous spaces and certain stochastic processes related to levy's brownian motion of several parameters. *Annales de l'Institut Henri Poincaré Section B (NS)*, 3(2):121–226.
- Gneiting, T. (1999). Radial positive definite functions generated by Euclid's hat. *Journal of Multivariate Analysis*, 69:88–119.
- Grenander, U. and Miller, M. I. (1994). Representations of knowledge in complex systems. *Journal of the Royal Statistical Society Series B*, 56(4):549–603.
- Hall, P. and Roy, R. (1994). On the relationship between fractal dimension and fractal index for stationary stochastic processes. *Advances in Applied Probability*, 4(1):241–253.
- Hammersley, J. M. and Nelder, J. A. (1955). Sampling from an isotropic Gaussian process. *Proceedings of the Cambridge Philosophical Society*, 51:652–662.
- Hausdorff, F. (1918). Dimension und äußeres maß. *Mathematische Annalen*, 79:157–179.
- Hellmund, G., Prokešová, M., and Jensen, E. B. V. (2008). Lévy based cox point processes. *Advances in Applied Probability*, 40:603–629.
- Hobolth, A. (2003). The spherical deformation model. *Biostatistics*, 4(4):583–595.
- Hobolth, A., Kent, J. T., and Dryden, I. L. (2002). On the relation between edge and vertex modelling in shape analysis. *Scandinavian Journal of Statistics*, 29:355–374.
- Hobolth, A., Pedersen, J., and Jensen, E. (2003). A continuous parametric shape model. *Annals of the Institute of Statistical Mathematics*, 55:227–242.
- Jones, R. H. (1963). Stochastic processes on a sphere. *The Annals of Mathematical Statistics*, 34(1):213–218.
- Jones, T. and Stofan, E. (2008). *Planetology : Unlocking the secrets of the solar system*. Washington, D.C.: National Geographic Society.

- Jónsdóttir, K. Y., Schmiegel, J., and Jensen, E. B. V. (2008). Lévy based growth models. *Bernoulli*, 14:62–90.
- Kent, J. T., Dryden, I. L., and Anderson, C. R. (2000). Using circulant symmetry to model featureless objects. *Biometrika*, 87(3):527–544.
- Kucinskas, A. B., Turcotte, D. L., Huang, J., and Ford, P. G. (1992). Fractal Analysis of Venus Topography in Tinatin Planitia and Ovda Regio. *Journal of Geophysical Research*, 97(E8):13635–13641.
- Leopardi, P. (2006). A partition of the unit sphere into regions of equal area and small diameter. *Electronic Transactions on Numerical Analysis*, 25:309–327.
- Mandelbrot, B. B. (1982). *The Fractal Geometry of Nature*. Times Books, 1 edition.
- Miller, M. I., Joshi, S., Maffitt, D. R., McNally, J. G., and Grenander, U. (1994). Membranes, mitochondria and amoebae: shape models. *Journal of Applied Statistics*, 21(1):141–163.
- Muinenen, K., Nousiainen, T., Fast, P., Lumme, K., and Peltoniemi, J. I. (1996). Light scattering by Gaussian random particles: ray optics approximation. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 55(5):577–601.
- Muñoz, O., Volten, H., Hovenier, J. W., Nousiainen, T., Muinenen, K., Guirado, D., Moreno, F., and Waters, L. (2007). Scattering matrix of large Saharan dust particles: Experiments and computations. *Journal of Geophysical Research*, 112(D13):D13215.
- Oliver, D. S. (1995). Moving averages for Gaussian simulation in two and three dimensions. *Mathematical Geology*, 27:939–960.
- Orford, J. D. and Whalley, W. B. (1983). The use of the fractal dimension to quantify the morphology of irregular-shaped particles. *Sedimentology*, 30:655–668.
- Pidwirny, M. (2006). Fundamentals of physical geography (2nd edition). <http://www.physicalgeography.net>. Retrieved March 2007.
- Price, F. (1988). *The moon observer’s handbook*. Cambridge University Press.
- R Development Core Team (2009). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Schoenberg, I. J. (1938). Metric spaces and completely monotone functions. *The Annals of Mathematics*, 39(4):811–841.
- Seidelmann, P., Archinal, B., A’hearn, M., Conrad, A., Consolmagno, G., Hestroffer, D., Hilton, J., Krasinsky, G., Neumann, G., Oberst, J., Stooke, P., Tedesco, E., Tholen, D., Thomas, P., and Williams, I. (2007). Report of the IAU/IAG working group on cartographic coordinates and rotational elements: 2006. *Celestial Mechanics and Dynamical Astronomy*, 98:155–180.
- Taylor, G. J. (1998). The biggest hole in the solar system. Technical report.
- Tovchigrechko, A. and Vakser, I. A. (2001). How common is the funnel-like energy landscape in protein-protein interactions? *Protein Science*, 10(8):1572–1583.
- Turcotte, D. L. (1987). A fractal interpretation of topography and geoid spectra on the earth, moon, venus, and mars. *Journal of Geophysical Research*, 92:E597–E601.

- Wicksell, S. D. (1925). The corpuscle problem. A mathematical study of a biometric problem. *Biometrika*, 17:84–89.
- Wicksell, S. D. (1926). The corpuscle problem. Second memoir. Case of ellipsoidal corpuscles. *Biometrika*, 18:152–172.
- Wieczorek, M. A., Jolliff, B. L., Khan, A., Pritchard, M. E., Weiss, B. P., Williams, J. G., Hood, L. L., Richter, K., Neal, C. R., Shearer, C. K., McCallum, I. S., Tompkins, S., Hawke, B. R., Peterson, C., Gillis, J. J., and Bussey, B. (2006). The constitution and structure of the lunar interior. *Reviews in Mineralogy and Geochemistry*, 60(1):221–364.
- Wolpert, R. L. and Ickstadt, K. (1998). Poisson/gamma random field models for spatial statistics. *Biometrika*, 85(2):251–267.
- Xue, Y. and Xiao, Y. (2009). Fractal and Smoothness Properties of Space-Time Gaussian Models. <http://adsabs.harvard.edu/abs/2009arXiv0912.0285X>.





# Local stereology for Lévy particles

Linda Vadgård Hansen and Eva B. Vedel Jensen

*Centre for Stochastic Geometry and Advanced Bioimaging, Department of  
Mathematical Sciences, Aarhus University*

**Abstract:** Identifying a particle via its radial function and then modelling the radial function by a kernel smoothing of a Lévy basis provide a flexible parametric model of three-dimensional particles. Under this particle model the particle volume distribution is known and statements on the distribution of the local stereological volume estimators are obtained.

**Key words:** Lévy basis, local stereology, particle model, random field on the sphere, volume distribution





## D.1 Introduction

An advantage of local stereological techniques is that they can be applied without specific assumptions on the shape of the object under study (Jensen, 1998). This is a very important step forward compared to earlier methods that relied on restrictive shape assumptions such as spherical or ellipsoidal shape (Wicksell, 1925, 1926; Cruz-Orive, 1976, 1978). Local stereology allows the unbiased estimation of the mean particle size, such as volume and surface area, without specific shape assumptions on the particles belonging to the particle population under study. Unfortunately, these relaxed shape assumptions has the consequence that only very weak statements can be made about the distributional properties of the estimators. Since it is difficult to separate the variance due to estimation of the size of each sampled particle from the variance in the true particle size distribution, it is not possible to obtain an estimate of the true particle size distribution.

In this paper we propose a flexible parametric model of three-dimensional particles under which the particle volume distribution can be estimated. Also, stronger statements about the distribution of the local volume estimators can be provided under this model. Each particle is modelled as a compact and star-shaped random set via its corresponding radial function. The radial function is a random field on the unit sphere and it is modelled as the  $p$ 'th power of a kernel smoothing of a Lévy basis, defined on the unit sphere. We refer to this as the power  $p$  Lévy based model and to the random sets obtained under this model as Lévy particles. We note that for  $p = 1$  this purely spatial model is a special case of the spatio-temporal Lévy based growth model suggested in Jónsdóttir et al. (2008) for modelling particle growth over time. In a simulation study, Jónsdóttir et al. (2008) investigates the effect of the choice of Lévy basis on the shape of the particle. Furthermore, for  $p = 1$ , the model has been studied in Hansen et al. (2011b). They consider the fractal properties of the surface of the Lévy particle along with simulation examples.

As opposed to the earlier models suffering from severe shape restrictions the Lévy particle model proposed in this paper is very flexible concerning shape but yet it is statistically tractable. Performing parameter estimation is straightforward using e.g. the method of moments or maximum likelihood theory.

It is indeed very attractable to have such flexibility in a model. Though, for certain modelling purposes it might be desirable to favour specific shape characteristics. For instance fairly symmetric and possibly elongated shapes, without being perfect ellipsoids, say, often occur in studies of cell populations. We consider a class of isotropic kernels on the interval  $[0, \pi]$  that concentrate their mass around the end points 0 and  $\pi$ . This kernel type favours the above mentioned shape characteristics.

In this paper, we explore the advantages of the power  $p$  Lévy based models in connection with local stereological volume estimation. Under the power 3 Lévy based model a simple parametric form of the distribution of particle volume is derived. Furthermore, the considered local stereological volume predictors are

presented in a form such that their respective distributions are straightforward to simulate. If, in addition, we assume a certain simple choice of kernel (*the antipodal uniform kernel*) we obtain distributional results for the respective predictors.

The paper is organised as follows. Firstly, the model along with some theory is presented in Section D.2. Section D.3 concerns kernels of antipodal type and include a presentation of the two kernels to be considered in detail along with explicit expressions for the correlation functions they induce. Furthermore, examples of simulated particles are provided. Section D.4 presents the local stereological volume predictors to be considered and how their performance can be evaluated in term of mean square error. In Section D.5, the distribution of the volume predictors are derived under the power 3 Lévy based model. A discussion can be found in Section D.6 and, finally, some theoretical derivations are deferred to an Appendix.

## D.2 The model and theoretical background

Let  $Y$  be a random particle. We will model  $Y$  as a random compact subset of  $\mathbb{R}^3$  which is star-shaped at an interior point. Without loss of generalisation we assume this point to be the origin. Denote by  $R_Y$  the radial function of  $Y$  which is given by

$$R_Y(u) = \max\{r : ru \in Y\}, \quad u \in \mathbb{S}^2.$$

There is a 1–1 correspondence between  $Y$  and its radial function  $R_Y$ . The set  $Y$  can be obtained from  $R_Y$  in the following way

$$Y = \bigcup_{u \in \mathbb{S}^2} \{ru : 0 \leq r \leq R_Y(u)\}.$$

Here,  $\mathbb{S}^2$  denotes the unit sphere in  $\mathbb{R}^3$ , i.e.  $\mathbb{S}^2 = \{x \in \mathbb{R}^3 : \|x\| = 1\}$ . The radial function  $R_Y$  is modelled as the non-negative random field on  $\mathbb{S}^2$  given by

$$R_Y(u) = \left\{ \int_{\mathbb{S}^2} K(v, u) \Gamma(dv) \right\}^{1/p}, \quad \text{for } u \in \mathbb{S}^2, \quad p > 0. \quad (\text{D.1})$$

Here  $\Gamma$  is a positive Lévy basis on the Borel subsets of  $\mathbb{S}^2$  and  $K$  is a positive deterministic kernel function on  $\mathbb{S}^2 \times \mathbb{S}^2$ , that integrates to one, i.e.  $\int_{\mathbb{S}^2} K(v, u) dv = 1$ . Let  $c_1$  denote the normalising constant. We refer to the model (D.1) as the power  $p$  Lévy based model. If  $p = 1$  this purely spatial model is a special case of the linear spatio-temporal Lévy model presented in Jónsdóttir et al. (2008). We refer to this paper, and references therein, for a further description of Lévy bases.

In the following, we focus on two types of Lévy bases, a gamma basis and an inverse Gaussian basis, respectively. Furthermore, we assume the Lévy basis to be homogeneous, hence we have

$$\Gamma(dv) \sim \text{Ga}(\alpha dv, \beta), \quad \alpha, \beta > 0, \quad \text{and} \quad \Gamma(dv) \sim \text{IG}(\delta dv, \gamma), \quad \delta, \gamma > 0, \quad (\text{D.2})$$

in case of a gamma basis and an inverse Gaussian basis, respectively. There are several known parametrisations for both the gamma distribution and the inverse Gaussian distribution. We follow Jónsdóttir et al. (2008), that is  $\text{Ga}(\alpha, \beta)$  has density

$$\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x > 0,$$

and  $\text{IG}(\delta, \gamma)$  has density

$$\frac{\delta e^{\delta \gamma}}{\sqrt{2\pi}} x^{-3/2} \exp\left\{-\frac{1}{2}\left(\delta^2 \frac{1}{x} + \gamma^2 x\right)\right\}, \quad x > 0,$$

respectively. Here,  $\Gamma$  denotes the gamma function. The gamma and inverse Gaussian Lévy bases are examples of the more general G-measures presented in Brix (1999).

In order to utilise the moment result on linear spatio-temporal Lévy models presented in Jónsdóttir et al. (2008, (9), (10) and (11)), we often consider the power transformed random field  $X$  given by  $X(u) = R_Y(u)^p$ , for  $u \in \mathbb{S}^2$ . Then, we have the following first and second order results for  $X$ ,

$$\mathbb{E}(X(u)) = \mu_\Gamma,$$

$$\text{Var}(X(u)) = \sigma_\Gamma^2 \int_{\mathbb{S}^2} K(v, u)^2 dv, \quad (\text{D.3})$$

$$\text{Cov}(X(u_1), X(u_2)) = \sigma_\Gamma^2 \int_{\mathbb{S}^2} K(v, u_1) K(v, u_2) dv, \quad (\text{D.4})$$

where  $dv$  denotes the element of the Lebesgue measure on  $\mathbb{S}^2$ . The parameters  $\mu_\Gamma$  and  $\sigma_\Gamma^2$  depend on the choice of Lévy basis as described in Table D.1.

**Table D.1:** The table shows  $\mu_\Gamma$  and  $\sigma_\Gamma^2$  expressed in terms of the parameters for each of the Lévy bases under consideration.

	$\mu_\Gamma$	$\sigma_\Gamma^2$
$\text{Ga}(\alpha dv, \beta)$	$\alpha/\beta$	$\alpha/\beta^2$
$\text{IG}(\delta dv, \gamma)$	$\delta/\gamma$	$\delta/\gamma^3$

## D.3 Isotropic kernels

Now, and throughout, we assume that  $K$  is isotropic, in that  $K(v, u) = k(\vartheta(v, u))$  depends on the points  $v, u \in \mathbb{S}^2$  through the great-circle distance  $\vartheta(v, u) \in [0, \pi]$  only (or equivalently on the inner product  $v \cdot u \in [-1, 1]$ ). This restriction on the kernel and the fact that  $dv$  is rotation invariant imply that the integral in (D.3) does not depend on  $u$ . Therefore, the mean function and variance function are

both constant and we refer to them as  $\mu_X$  and  $\sigma_X^2$ , i.e.  $\mu_X = \mu_\Gamma$  and  $\sigma_X^2 = \sigma_\Gamma^2 c_2$ . Here,  $c_2$  denotes the integral over  $\mathbb{S}^2$  of the squared isotropic kernel, i.e.

$$c_2 = \int_{\mathbb{S}^2} k(\vartheta(v, u_0))^2 dv = 2\pi \int_0^\pi k(\theta)^2 \sin \theta d\theta.$$

Moreover, the covariance (D.4) only depends on  $\vartheta(u_1, u_2)$ . We let  $C(\eta)$  denote the correlation of the random field at two points on  $\mathbb{S}^2$  a distance  $\eta$  apart, i.e.

$$C(\eta) = \frac{1}{c_2} \int_{\mathbb{S}^2} k(\vartheta(v, u_0))k(\vartheta(v, u_\eta)) dv,$$

for  $0 \leq \eta \leq \pi$ . Here,  $u_\eta = (\sin \eta, 0, \cos \eta)$  for  $\eta \in [0, \pi]$ .

Below, we will in detail study two isotropic kernels. Both these kernels concentrate their mass at pairs of directions with mutual angle either around 0 or  $\pi$ . At first thought, the most natural thing would be to concentrate the mass of the kernel around 0. For many modelling purposes this would be completely reasonable. But because of the application we have in mind, modelling cells, we would like to have a model that produces sets that, on a global scale, are fairly symmetric and possibly elongated. We obtain this using such kernels that more or less equally distribute their mass around 0 and  $\pi$ . Such kernels are called *antipodal*.

### D.3.1 The antipodal uniform kernel

In this section we let  $k$  be the normalised sum of two indicator functions

$$k(\vartheta(v, u)) = c_1^{-1} (\mathbb{1}_{\{\vartheta(v, u) \leq r_1\}} + \mathbb{1}_{\{\vartheta(v, -u) \leq r_2\}}), \quad r_1, r_2 \in (0, \pi/2). \quad (\text{D.5})$$

For a fixed  $u \in \mathbb{S}^2$ , then  $k(\vartheta(v, u)) = 1/c_1$  if  $v \in B_{r_1}(u)$  or  $v \in B_{r_2}(-u)$ , where  $B_r(u) = \{v \in \mathbb{S}^2 : \vartheta(v, u) \leq r\}$  denotes the spherical cap on  $\mathbb{S}^2$  of radius  $r \in [0, \pi]$  centred at  $u$ . We will refer to the kernel (D.5) as the *antipodal uniform kernel*.

The kernel  $k$  given by (D.5) is itself the radial function of a set, say  $A_{r_1, r_2}$ . Moreover, as both  $r_1$  and  $r_2$  are less than  $\pi/2$  the Lebesgue measure of the intersection  $B_{r_1} \cap B_{r_2}$  is zero, and therefore  $k(\vartheta(v, u))^p = c_1^{1-p} k(\vartheta(v, u))$ . Thus, up to the constant  $c_1^{p-1}$ ,  $R_Y(u)$  denotes the quantity known within the field of stochastic geometry as *the  $p$ 'th radial mean body of  $A_{r_1, r_2}$*  (Gardner and Zhang, 1998).

Recall that the area of a spherical cap  $B_r(u)$  is  $2\pi(1 - \cos r)$ . Therefore, we get  $c_1 = 4\pi[1 - (\cos r_1 + \cos r_2)/2]$  and  $c_2 = 4\pi c_1^{-2}[1 - (\cos r_1 + \cos r_2)/2]$ .

Let  $A(d, r_1, r_2)$  denote the area of the intersection of two spherical caps a distance  $d$  apart and of radius  $r_1$  and  $r_2$ , respectively. A formula for the computation of  $A(d, r_1, r_2)$  is given by

$$A(d, r_1, r_2) = \begin{cases} 2\pi(1 - \cos(r_1 \wedge r_2)), & \text{if } r_1 \wedge r_2 \leq (r_1 \vee r_2) - d, \\ 0, & \text{if } r_1 + r_2 \leq d, \\ L(d, r_1, r_2), & \text{otherwise.} \end{cases}$$

Here,  $x \wedge y$  and  $x \vee y$  denote the minimum and maximum of  $x$  and  $y$ , respectively. The function  $L$  that determines  $A$  in the case of partial intersection is given by

$$L(d, r_1, r_2) = 2 \left[ \pi(1 - \cos r_1 - \cos r_2) - \arccos\left(\frac{\cos d - \cos r_1 \cos r_2}{\sin r_1 \sin r_2}\right) + \cos r_1 \arccos\left(\frac{-\cos r_2 + \cos d \cos r_1}{\sin d \sin r_1}\right) + \cos r_2 \arccos\left(\frac{-\cos r_1 + \cos d \cos r_2}{\sin d \sin r_2}\right) \right],$$

see e.g. Tovchigrechko and Vakser (2001) or Sander (2007) for a simplified version. Because  $A$  is symmetric in the second and third argument, that is  $A(d, r_1, r_2) = A(d, r_2, r_1)$ , the correlation function can be expressed as

$$C(\eta) = [A(\eta, r_1, r_1) + 2A(\pi - \eta, r_1, r_2) + A(\eta, r_2, r_2)] / c_2.$$

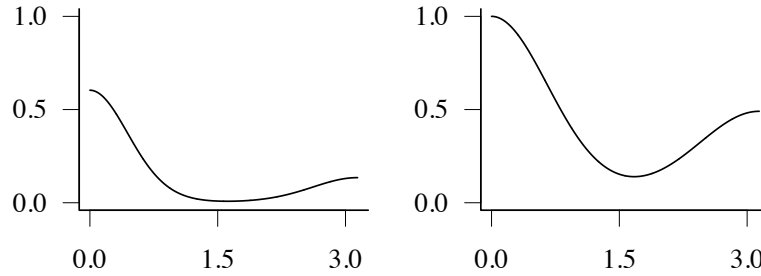
As we shall see later, the antipodal uniform kernel imply nice distributional results for the volume of a particle with radial function  $R_Y$  and for the volume predictors. A simulation study reveals that the antipodal uniform kernel produces particles with non-smooth boundary.

### D.3.2 The antipodal exponential kernel

For real valued shape parameters  $a_1, a_2$ , we consider the kernel given by

$$k(\vartheta(v, u)) = (e^{a_1(v \cdot u)} + e^{-a_2(v \cdot u)}) / c_1. \quad (\text{D.6})$$

This kernel is called *the antipodal exponential kernel*. Figure D.1 (left) shows the kernel graph for parameter values are  $a_1 = 5$  and  $a_2 = 3.5$ . The antipodal nature of the kernel is evident from this plot.



**Figure D.1:** Left. The antipodal exponential kernel  $k(\vartheta(v, u))$  (D.6) as a function of the angle between  $v$  and  $u$ . Right. The correlation function  $C(\eta)$  D.7. For both plots the kernel parameters are  $a_1 = 5$  and  $a_2 = 3.5$ .

Using polar decomposition we find

$$c_1 = 2\pi \int_0^\pi (e^{a_1 \cos \theta} + e^{-a_2 \cos \theta}) \sin \theta \, d\theta = 4\pi (\text{sinhc}(a_1) + \text{sinhc}(a_2)).$$

Here,  $\text{sinhc}(z) = \sinh(z)/z$  for  $z \neq 0$  and  $\text{sinhc}(0) = 1$ . Notice that

$$\begin{aligned} c_2 &= \frac{2\pi}{c_1^2} \int_0^\pi (e^{a_1 \cos \theta} + e^{-a_2 \cos \theta})^2 \sin \theta \, d\theta \\ &= \frac{4\pi}{c_1^2} (\text{sinhc}(2a_1) + \text{sinhc}(2a_2) + 2\text{sinhc}(a_1 - a_2)). \end{aligned}$$

Before deducing an expression for the correlation function, we make some initial observations. First notice that

$$\begin{aligned}
 c_1^2 \int_{\mathbb{S}^2} k(\vartheta(v, u_0)) k(\vartheta(v, u_\eta)) dv \\
 &= \int_{\mathbb{S}^2} (e^{a_1(v \cdot u_0)} + e^{-a_2(v \cdot u_0)}) (e^{a_1(v \cdot u_\eta)} + e^{-a_2(v \cdot u_\eta)}) dv \\
 &= f(a_1(u_0 + u_\eta)) + f(a_1 u_0 - a_2 u_\eta) + f(a_1 u_\eta - a_2 u_0) + f(-a_2(u_0 + u_\eta))
 \end{aligned}$$

where  $f(w) = \int_{\mathbb{S}^2} e^{v \cdot w} dv$ , for  $w \in \mathbb{S}^2$ . We get,

$$f(w) = \int_{\mathbb{S}^2} e^{v \cdot w} dv = \int_{\mathbb{S}^2} \exp\{\|w\| (v \cdot \frac{w}{\|w\|})\} dv = 4\pi \operatorname{sinhc}(\|w\|).$$

Furthermore,  $\|a u_0 + a' u_\eta\| = \sqrt{a^2 + a'^2 + 2aa' \cos \eta}$ , for  $a, a' \in \mathbb{R}$ . Thereby,

$$\begin{aligned}
 \frac{c_1^2}{4\pi} \int_{\mathbb{S}^2} k(\vartheta(v, u_0)) k(\vartheta(v, u_\eta)) dv \\
 &= \operatorname{sinhc}\left(a_1 \sqrt{2(1 + \cos \eta)}\right) + 2 \operatorname{sinhc}\left(\sqrt{a_1^2 + a_2^2 - 2a_1 a_2 \cos \eta}\right) \\
 &\quad + \operatorname{sinhc}\left(a_2 \sqrt{2(1 + \cos \eta)}\right).
 \end{aligned}$$

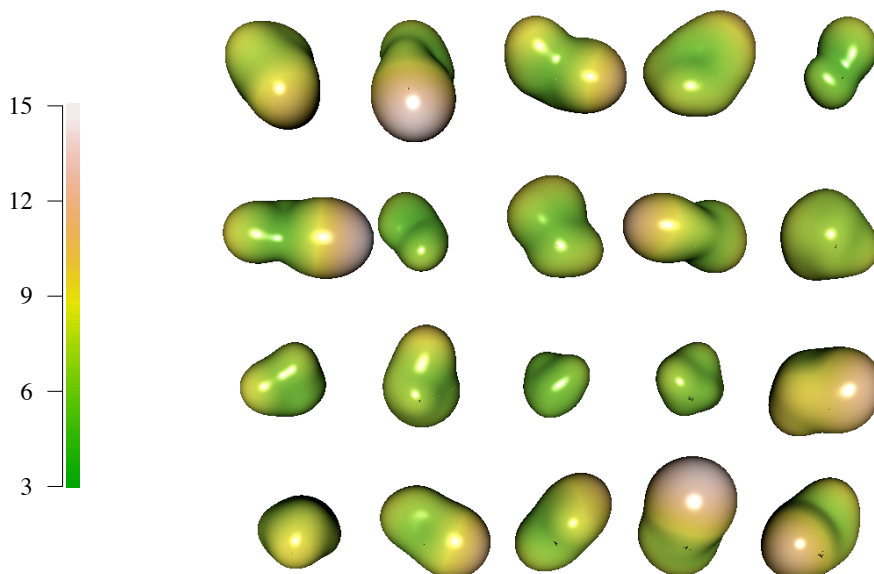
Thus, we have for the correlation function,

$$\begin{aligned}
 C(\eta) = \frac{4\pi}{c_2} \Big\{ \operatorname{sinhc}(a_1 \sqrt{2(1 + \cos \eta)}) + \operatorname{sinhc}(a_2 \sqrt{2(1 + \cos \eta)}) \\
 + 2 \operatorname{sinhc}(\sqrt{a_1^2 + a_2^2 - 2a_1 a_2 \cos \eta}) \Big\}. \quad (\text{D.7})
 \end{aligned}$$

For an illustrative plot of the correlation function we refer to Figure D.1 (right). Here, the kernel parameters are  $a_1 = 5$  and  $a_2 = 3.5$ . We see that the antipodal property of the kernel function is preserved by the correlation function.

Figure D.2 shows independent realisations of Lévy particles under the power 3 Lévy based model with the kernel function shown in Figure D.1 and a gamma basis (rows 1 and 2) and inverse Gaussian (rows 3 and 4), respectively. The parameter values of the Lévy basis are  $\alpha = 0.3$ ,  $\beta = 9 \times 10^{-4}$  and  $\delta = 10$ ,  $\gamma = 0.03$  for the gamma and inverse Gaussian basis, respectively. These basis parameter values assures that  $\mu_\Gamma$  and  $\sigma_\Gamma^2$  are the same under both choices Lévy bases. Based on Figure D.2 there are no obvious difference in the appearance of the particles in the four rows. Notice that, as suspected, the particles do not have a preferred orientation. Furthermore, some of the particles have a peanut-like shape due to the antipodal structure of the kernel function.

For the simulations we used the simulation algorithm proposed in Hansen et al. (2011b) but we raise the simulated values of the radial function to power  $1/p$  to fit our modelling set-up. The realisations in Figure D.2 were conducted for  $m_1 = 301$ ,  $m_2 = 300$  and  $N = 10^5$ . The small dark dots appearing on some



**Figure D.2:** Independent realisations of Lévy particles under the gamma basis (rows 1 and 2) and inverse Gaussian basis (rows 3 and 4), respectively. The kernel parameters are as in Figure D.1. The basis parameter values are  $\alpha = 0.3$ ,  $\beta = 9 \times 10^{-4}$  for the gamma basis and  $\delta = 10$ ,  $\gamma = 0.03$  for the inverse Gaussian basis, respectively. For each particles the length of the radial function is given by the colour according to the chart on the left.

of the particles are all artificial features due to the discretisation algorithm used for the simulations.

In Hansen et al. (2011b) the model (D.1) is considered in the special case where  $p = 1$ . In that context the unnormalised version of (D.6) with  $a_1 > 0$  and  $a_2 = 0$  is considered. The resulting kernel is the unnormalised von Mises–Fisher density often used in directional statistics. It is shown that under such assumptions the correlation function has the particularly nice form

$$\operatorname{sinhc}(a_1 \sqrt{2(1 + \cos \eta)}) / \operatorname{sinhc}(2 a_1), \quad 0 \leq \eta \leq \pi.$$

## D.4 Volume prediction

Since  $Y$  is star-shaped at the origin, its volume  $V(Y)$  can be expressed in terms of an integral of the radial function  $R_Y$  of  $Y$ ,

$$V(Y) = \frac{1}{3} \int_{\mathbb{S}^2} R_Y(u)^3 \, du,$$

see e.g. Baddeley and Jensen (2005, (8.25)). In the literature, there are several well-known estimators of  $V(Y)$ , see e.g. Jensen (2000). In the following we consider three such local stereological volume estimators, two of which are based

on line information and one on planar information. With the above integral expression of  $V(Y)$  in mind it is easily seen, that

$$\widehat{V}_1 = 4\pi R_Y(U)^3/3,$$

where  $U$  is a uniform direction on  $\mathbb{S}^2$ , is a conditional unbiased estimator of  $V(Y)$ , i.e.  $E(\widehat{V}_1|Y) = V(Y)$ . This type of estimator is often referred to as *the nucleator* (based on one ray), cf. Gundersen (1988). In order to reduce variability it has been suggested to replace  $U$  by a systematically set of directions, yielding the following conditionally unbiased estimator

$$\widehat{V}_{1,N} = 4\pi \sum_{i=1}^N R_Y(U_i)^3/(3N),$$

where  $U_1$  is uniform on  $\mathbb{S}^2 \cap E$  and  $U_i$  is a  $2\pi(i-1)/N$  anticlockwise rotation of  $U_1$  within  $\mathbb{S}^2 \cap E$ . Throughout the paper we denote by  $E$  an isotropic random plane through the origin, i.e.  $E$  is a two dimensional subspace with normal vector that is uniformly distributed on the hemisphere  $\mathbb{S}_+^2$ , cf. Baddeley and Jensen (2005, p. 124). We will refer to  $\widehat{V}_{1,N}$  as the *N-ray nucleator*. If we imagine using an infinite number of rays we obtain an estimator that utilises all the information in the section  $Y \cap E$ ,

$$\widehat{V}_2 = \frac{2}{3} \int_{\mathbb{S}^2 \cap E} R_Y(u)^3 du.$$

This estimator can also be derived from Baddeley and Jensen (2005, (8.41)). Following Hansen et al. (2011a) we refer to  $\widehat{V}_2$  as *the integrated nucleator*.

We notice that, for  $\widehat{V}$  any of the three estimators mentioned above we have  $E(\widehat{V}|Y) = V(Y)$ , i.e.  $\widehat{V}$  is a conditional unbiased estimator of  $V(Y)$  given  $Y$ . Since  $V(Y)$  is a random variable we will from now on refer to  $\widehat{V}_1$ ,  $\widehat{V}_{1,N}$  and  $\widehat{V}_2$  as design unbiased predictors of  $V(Y)$ .

Often the variance is used as a measure of the precision of an estimator. When dealing with predictors a more suitable performance measure is the mean square error (MSE). The MSE of a predictor  $\widehat{V}$  for  $V(Y)$  is given by

$$\text{MSE}(\widehat{V}) = E(\widehat{V} - V(Y))^2.$$

If  $\widehat{V}$  is design unbiased, then

$$\text{MSE}(\widehat{V}) = E(\widehat{V}^2) - E(V(Y)^2) = \text{Var}(\widehat{V}) - \text{Var}(V(Y)).$$

For the calculation of the MSE later on the following observation, regarding conditioning, becomes useful. We have  $\text{MSE}(\widehat{V}) = E(\text{MSE}(\widehat{V}|Y))$ , where

$$\text{MSE}(\widehat{V}|Y) := E((\widehat{V} - V(Y))^2|Y) = E(\widehat{V}^2|Y) - 2V(Y)E(\widehat{V}|Y) + V(Y)^2.$$

If  $\widehat{V}$  is design unbiased, then

$$\text{MSE}(\widehat{V}|Y) = E(\widehat{V}^2|Y) - E(\widehat{V}|Y)^2 = \text{Var}(\widehat{V}|Y), \quad (\text{D.8})$$

and thereby  $\text{MSE}(\widehat{V}) = E \text{Var}(\widehat{V}|Y)$ , cf. Hobolth and Jensen (2002).



## D.5 Volume prediction under the power 3 Lévy model

In this section, we will derive the distribution of  $V(Y)$  and the three predictors  $\widehat{V}_1$ ,  $\widehat{V}_{1,N}$  and  $\widehat{V}_2$  under the model (D.1) with  $p = 3$ . In this case,

$$V(Y) = \frac{1}{3} \int_{\mathbb{S}^2} \int_{\mathbb{S}^2} k(\vartheta(v, u)) \, du \, \Gamma(\mathrm{d}v) = \Gamma(\mathbb{S}^2)/3.$$

This shows that  $V(Y)$  follows the same distribution as the Lévy basis, independent of the choice of kernel. In particular, we have for the variance

$$\mathrm{Var}(V(Y)) = 4\pi\sigma_\Gamma^2/9.$$

If  $\Gamma$  is a gamma basis as in (D.2) then

$$V(Y) \sim \mathrm{Ga}(4\pi\alpha, \beta)/3 = \mathrm{Ga}(4\pi\alpha, 3\beta).$$

If, on the other hand,  $\Gamma$  is an inverse Gaussian basis as in (D.2) then

$$V(Y) \sim \mathrm{IG}(4\pi\delta, \gamma)/3 = \mathrm{IG}\left(\frac{4\pi}{\sqrt{3}}\delta, \sqrt{3}\gamma\right).$$

Furthermore, the predictors take the form

$$\begin{aligned} \widehat{V}_1 &= \frac{4\pi}{3} \int_{\mathbb{S}^2} k(\vartheta(v, U)) \, \Gamma(\mathrm{d}v), \\ \widehat{V}_{1,N} &= \frac{4\pi}{3} \frac{1}{N} \sum_{i=1}^N \int_{\mathbb{S}^2} k(\vartheta(v, U_i)) \, \Gamma(\mathrm{d}v), \\ \widehat{V}_2 &= \frac{2}{3} \int_{\mathbb{S}^2} \int_{\mathbb{S}^2 \cap E} k(\vartheta(v, u)) \, du \, \Gamma(\mathrm{d}v). \end{aligned}$$

Recall that, for the considered Lévy bases the distribution of  $\Gamma(A)$  only depends on  $A$  via its Lebesgue measure. Thus, as  $k$  is isotropic and the Lebesgue measure is rotation invariant, the distribution of the estimator  $\widehat{V}_1$  is independent of  $U$ . An equivalent distributional independence statement holds for the pairs  $\widehat{V}_{N,1}$  and  $\{U_1, \dots, U_N\}$  and  $\widehat{V}_2$  and  $E$ , respectively.

The MSE of  $\widehat{V}_1$ ,  $\widehat{V}_{1,N}$  and  $\widehat{V}_2$ , respectively, is of the following form

$$\mathrm{MSE}(\widehat{V}_1) = \sigma_\Gamma^2 \frac{4\pi}{9} (4\pi c_2 - 1) \quad (\text{D.9})$$

$$\mathrm{MSE}(\widehat{V}_{1,N}) = \sigma_\Gamma^2 \frac{4\pi}{9} \left( \frac{4\pi c_2}{N} \left[ C(0) + 2 \sum_{j=1}^{\lceil N/2 \rceil} C(\pi \frac{j}{N}) + C(\pi) \mathbb{1}_{\{N \text{ even}\}} \right] - 1 \right) \quad (\text{D.10})$$

$$\mathrm{MSE}(\widehat{V}_2) = \sigma_\Gamma^2 \frac{4\pi}{9} \left( 4c_2 \int_0^\pi C(\eta) \, \mathrm{d}\eta - 1 \right). \quad (\text{D.11})$$

The proof of these results can be found in the Appendix.

In the theorem below, we present distributional results on the volume predictors  $\widehat{V}_1$ ,  $\widehat{V}_{1,N}$  and  $\widehat{V}_2$  for the antipodal uniform kernel (D.5) and an arbitrary Lévy basis. For the formulation of the results we need some notation. Let  $\text{Band}_r(C)$  denote the great-circle band of radius  $r$  around the great-circle  $C$ ,

$$\text{Band}_r(C) := \{ v \in \mathbb{S}^2 \mid \text{there exists } u \in C : \vartheta(v, u) \leq r \}, \quad \text{for } r > 0.$$

In particular, let  $\text{Band}_r$  denote the equatorial band of radius  $r$ .

**Theorem D.1.** *Assume the power 3 Lévy model with the antipodal uniform kernel (D.5) and an arbitrary Lévy basis. Then,  $\widehat{V}_1$ ,  $\widehat{V}_{1,N}$  and  $\widehat{V}_2$  can be expressed as linear combinations of random variables following the same distribution as the Lévy basis,*

$$\begin{aligned} (i) \quad \widehat{V}_1 &= \frac{4\pi}{3c_1} \left[ \Gamma(B_{r_1}(u_0)) + \Gamma(B_{r_2}(-u_0)) \right], \\ (ii) \quad \widehat{V}_{1,N} &= \frac{4\pi}{3c_1} \frac{1}{N} \sum_{i=1}^N \left[ \Gamma(B_{r_1}(u_i)) + \Gamma(B_{r_2}(-u_i)) \right], \\ (iii) \quad \widehat{V}_2 &= \frac{2}{3c_1} \left[ 2 \Gamma(\text{Band}_{r_1 \wedge r_2}) + \Gamma(\text{Band}_{r_1 \vee r_2} \setminus \text{Band}_{r_1 \wedge r_2}) \right]. \end{aligned}$$

Here,  $u_0 = (0, 0, 1)$  and  $u_i = (\sin(\frac{i-1}{N}2\pi), 0, \cos(\frac{i-1}{N}2\pi))$ , for  $i = 1, \dots, N$ . Furthermore, the summands in (i) and (iii), respectively, are independent random variables. In (ii), for  $i = 1, \dots, N$ , the pair  $\Gamma(B_{r_1}(u_i))$  and  $\Gamma(B_{r_2}(-u_i))$  are independent.

*Proof.* As (i) is the special case of (ii) where  $N = 1$  we simply consider this more general case. We find,

$$\begin{aligned} \widehat{V}_{1,N} &= \frac{4\pi}{3c_1} \frac{1}{N} \sum_{i=1}^N \int_{\mathbb{S}^2} (\mathbb{1}_{\{\vartheta(v, U_i) < r_1\}} + \mathbb{1}_{\{\vartheta(v, -U_i) < r_2\}}) \Gamma(\mathrm{d}v) \\ &= \frac{4\pi}{3c_1} \frac{1}{N} \sum_{i=1}^N \int_{\mathbb{S}^2} (\mathbb{1}_{\{\vartheta(v, u_i) < r_1\}} + \mathbb{1}_{\{\vartheta(v, -u_i) < r_2\}}) \Gamma(\mathrm{d}v) \\ &= \frac{4\pi}{3c_1} \frac{1}{N} \sum_{i=1}^N [\Gamma(B_{r_1}(u_i)) + \Gamma(B_{r_2}(-u_i))]. \end{aligned}$$

For the second equality we utilised that

$$\begin{aligned} &\text{For any Borel subset } A \subseteq \mathbb{S}^2, \text{ the distribution of } \Gamma(A) \\ &\text{only depend on the area of } A \text{ and not on its position.} \end{aligned} \tag{D.12}$$

Both  $r_1$  and  $r_2$  are less than  $\pi/2$ . Thus, for fixed  $i$ ,  $B_{r_1}(u_i) \cap B_{r_2}(-u_i) = \emptyset$  and therefore  $\Gamma(B_{r_1}(u_i))$  and  $\Gamma(B_{r_2}(-u_i))$  are independent.

For  $i \neq j$ , the intersection  $B_{r_{i'}}(u_i) \cap B_{r_{j'}}(-u_j)$  might be non-empty in which case the corresponding variables  $\Gamma(B_{r_{i'}}(u_i))$  and  $\Gamma(B_{r_{j'}}(-u_j))$  might be correlated.

To show (iii) notice that

$$\begin{aligned}
 \widehat{V}_2 &= \frac{2}{3c_1} \int_{\mathbb{S}^2} \int_{\mathbb{S}^2 \cap E} [\mathbb{1}_{B_{r_1}(u)}(v) + \mathbb{1}_{B_{r_2}(-u)}(v)] \, du \, \Gamma(\mathrm{d}v) \\
 &= \frac{2}{3c_1} \int_{\mathbb{S}^2} [\mathbb{1}_{\text{Band}_{r_1}(E)}(v) + \mathbb{1}_{\text{Band}_{r_2}(E)}(v)] \, \Gamma(\mathrm{d}v) \\
 &= \frac{2}{3c_1} \int_{\mathbb{S}^2} [\mathbb{1}_{\text{Band}_{r_1}}(v) + \mathbb{1}_{\text{Band}_{r_2}}(v)] \, \Gamma(\mathrm{d}v) \\
 &= \frac{2}{3c_1} \int_{\mathbb{S}^2} [2 \mathbb{1}_{\text{Band}_{r_1 \wedge r_2}}(v) + \mathbb{1}_{\text{Band}_{r_1 \vee r_2} \setminus \text{Band}_{r_1 \wedge r_2}}(v)] \, \Gamma(\mathrm{d}v) \\
 &= \frac{2}{3c_1} \left[ 2 \Gamma(\text{Band}_{r_1 \wedge r_2}) + \Gamma(\text{Band}_{r_1 \vee r_2} \setminus \text{Band}_{r_1 \wedge r_2}) \right].
 \end{aligned}$$

Again, for the third equality, we utilised (D.12). To conclude, we notice that  $\text{Band}_{r_1 \wedge r_2}$  and  $\text{Band}_{r_1 \vee r_2} \setminus \text{Band}_{r_1 \wedge r_2}$  are non-overlapping thus we have that  $\widehat{V}_2$  can be written as a sum of two independent variables.  $\square$

Depending on the choice of Lévy basis we get as a corollary to Theorem D.1 the explicit distribution of the predictors  $\widehat{V}_1$  and  $\widehat{V}_2$ .

**Corollary D.1.** *Let the model be as in Theorem D.1. Then,*

(i) *If  $\Gamma(\mathrm{d}v) \sim \text{Ga}(\alpha \mathrm{d}v, \beta)$  then*

$$\begin{aligned}
 \widehat{V}_1 &\sim \text{Ga}(c_1 \alpha, \frac{3c_1}{4\pi} \beta) \\
 \widehat{V}_2 &\sim \text{Ga}\left(4\pi \cos\left(\frac{\pi}{2} - r_1 \wedge r_2\right) \alpha, \frac{3c_1}{4} \beta\right) \\
 &\quad + \text{Ga}\left(4\pi(\cos\left(\frac{\pi}{2} - r_1 \vee r_2\right) - \cos\left(\frac{\pi}{2} - r_1 \wedge r_2\right)) \alpha, \frac{3c_1}{2} \beta\right)
 \end{aligned}$$

(ii) *If  $\Gamma(\mathrm{d}v) \sim \text{IG}(\delta \mathrm{d}v, \gamma)$  then*

$$\begin{aligned}
 \widehat{V}_1 &\sim \text{IG}\left(\sqrt{\frac{4\pi c_1}{3}} \delta, \sqrt{\frac{3c_1}{4\pi}} \gamma\right) \\
 \widehat{V}_2 &\sim \text{IG}\left(4\pi \cos\left(\frac{\pi}{2} - r_1 \wedge r_2\right) \frac{2}{\sqrt{3c_1}} \delta, \frac{\sqrt{3c_1}}{2} \gamma\right) \\
 &\quad + \text{IG}\left(4\pi(\cos\left(\frac{\pi}{2} - r_1 \vee r_2\right) + \cos\left(\frac{\pi}{2} - r_1 \wedge r_2\right)) \sqrt{2/(3c_1)} \delta, \sqrt{3c_1/2} \gamma\right).
 \end{aligned}$$

Before conducting the proof, we notice that the area of a great-circle band of radius  $r$  is  $4\pi$  minus two times the area of a spherical cap of radius  $\pi/2 - r$ . Thus,

$$\text{area}(\text{Band}_r) = 4\pi - 2[2\pi(1 - \cos(\frac{\pi}{2} - r))] = 4\pi \cos(\frac{\pi}{2} - r), \quad r \leq \pi/2. \quad (\text{D.13})$$

*Proof.* First, we consider the distribution of  $\widehat{V}_1$ . The proof follows directly from summation and scaling properties of independent gamma and independent inverse Gaussian random variables, respectively.

To show the distribution of  $\widehat{V}_2$  we utilise the observation (D.13) and the scaling properties of the respective distributions.  $\square$

As a special case of Corollary D.1 we obtain the following result for the distribution of  $\widehat{V}_2$ .

**Corollary D.2.** *Let the model be as in Theorem D.1 and assuming that  $r = r_1 = r_2$ . Then,*

$$\widehat{V}_2 \sim \text{Ga}(4\pi \cos(\pi/2 - r) \alpha, 3c_1/4 \beta) \quad \text{and} \quad \widehat{V}_2 \sim \text{IG}\left(\frac{8\pi}{\sqrt{3c_1}} \cos\left(\frac{\pi}{2} - r\right) \delta, \frac{\sqrt{3c_1}}{2} \gamma\right),$$

if  $\Gamma(\text{d}v) \sim \text{Ga}(\alpha \text{d}v, \beta)$  and  $\Gamma(\text{d}v) \sim \text{IG}(\delta \text{d}v, \gamma)$ , respectively.

## D.6 Discussion

In the proposed model (D.1), we restrict ourselves to positive Lévy bases in order to assure that the process  $R_Y$  is non-negative. An alternative is to consider instead a power  $p$  log-linear Lévy model of  $R_Y$ , i.e. for  $u \in \mathbb{S}^2$ ,

$$R_Y(u) = \left( \exp \left\{ \int_{\mathbb{S}^2} K(v, u) \Gamma(\text{d}v) \right\} \right)^{1/p} = \exp \left\{ \frac{1}{p} \int_{\mathbb{S}^2} K(v, u) \Gamma(\text{d}v) \right\}. \quad (\text{D.14})$$

A disadvantage of the model (D.14) is that the nice distributional results on the volume and the volume predictors do not seem to be obtainable as under the model (D.1). The reason is that we no longer are able to interchange the order of integration during the calculations of volume and volume predictors when  $R_Y(u)$  is modelled according to (D.14).

We have considered antipodal kernels to obtain fairly symmetric and possibly elongated particles. An alternative idea for obtaining such shape characteristics might be to introduce shape via a deterministic function  $f : \mathbb{S}^2 \mapsto [0, \infty)$  added directly to the kernel function in (D.1). Then, the kernel becomes  $f(u) + K(v, u)$  and for the model we have

$$R_Y(u) = \left( f(u) \Gamma(\mathbb{S}^2) + \int_{\mathbb{S}^2} K(v, u) \Gamma(\text{d}v) \right)^{1/p}. \quad (\text{D.15})$$

For  $p = 1$  this model provides particles that are random scalings (by  $\Gamma(\mathbb{S}^2)$ ) of the deterministic shape  $f$  with some added noise from the kernel smoothing integral term. The amount of noise is of course dependent on the kernel and the Lévy basis. As under the model (D.1) this model imply that the one-ray nucleator, the  $N$ -ray nucleator and the integrated nucleator are design unbiased predictors of  $V(Y)$ . The distribution of  $V(Y)$  again follows the same distribution as the Lévy basis but the distribution of the predictors  $\widehat{V}_1$ ,  $\widehat{V}_{1,N}$  and  $\widehat{V}_2$  will now longer be independent of  $U$ ,  $\{U_1, U_2, \dots, U_N\}$  and  $E$ , respectively. The simulation example of Figure D.2 illustrates that the power  $p$  Lévy based model does not favour any certain spatial orientation of the Lévy particle. This isotropic property will not be present for the model (D.16) unless the function  $f$  is independent of the direction  $u$ .

A slightly different way of incorporating the deterministic shape function  $f$  would be to add it directly to the integral in the model (D.1). Then,

$$R_Y(u) = \left\{ f(u) + \int_{\mathbb{S}^2} K(v, u) \Gamma(\mathrm{d}v) \right\}^{1/p}. \quad (\text{D.16})$$

This model only differs from the above model (D.15) by the feature that the deterministic shape  $f$  is no longer scaled. The induced predictors again become design unbiased. The distributional results are slightly different. For instance consider the distribution of  $V(Y)$  for  $p = 3$ . We have

$$V(Y) = \frac{1}{3} \int_{\mathbb{S}^2} f(u) \, \mathrm{d}u + \frac{1}{3} \Gamma(\mathbb{S}^2).$$

Therefore,  $V(Y)$  does not follow the same distribution as the Lévy basis but a shifted (by  $\frac{1}{3} \int_{\mathbb{S}^2} f(u) \, \mathrm{d}u$ ) version thereof.

In this paper we identify  $Y$  by its radial function. This choice was inspired by the possible application within local stereology due to the following neat property of the radial function. Let  $E$  be an isotropic random plane. Then

$$R_{Y \cap E}(u) = R_Y(u), \quad \text{for } u \in E.$$

As a consequence, we can make inference under the model (D.1) based on observations within one or more central sections.

If, in addition,  $Y$  is assumed convex then another approach might be worth pursuing. In this case  $Y$  can also be identified by its support function  $H_Y$  (Gardner, 2006). Then we might model  $H_Y(u)$  by the random field

$$H_Y(u) = \left\{ \int_{\mathbb{S}^2} h(v, u) \Gamma(\mathrm{d}v) \right\}^{1/p}, \quad (\text{D.17})$$

for a kernel function  $h$ . Here, it is required that, for each  $v \in \mathbb{S}^2$ ,  $h(v, \cdot)$  is a support function. Then the positivity of  $\Gamma$  together with the linearity of integration with respect to  $\Gamma(\mathrm{d}v)$  assure that the right side of D.17 becomes a support function for  $p = 1$  (Schneider, 1993, Theorem 1.7.1). Recall that, a function is sub-linear if and only if it is both convex and homogeneous. Thus, as  $x \mapsto x^{1/p}$  is convex for  $p \leq 1$ , the right hand side of (D.17) is a support function for  $p \leq 1$ . Let  $E_u$  denote a plane parallel to a plane through the origin containing  $u$ . Furthermore, let  $\text{proj}_Y(u)$  denote the projection of  $Y$  onto  $E_u$ . Then  $H_Y(u) = H_{\text{proj}_Y(u)}(u)$ . Thus, the model (D.17) might be of particular interest if observations are available on projection of the particle and not from central sections.

Moreover, an explicit relation between the radial function and the support function might allow for a comparison of local and pivotal volume estimators under the same Lévy particle model. Such a relation is at present unknown to us. Though, for a strict convex set  $Y$  a more implicit relation do exist, as  $R_Y(\nabla_{H_Y}(u)) = 1$ , where  $\nabla_{H_Y}(u)$  denotes the gradient of  $H_Y$  at  $u$  (Schneider, 1993, Corollary 1.7.3.).

The results of Theorem D.1 and its corollaries can be used for model checking. Say, we have performed measurements and obtained estimates  $\widehat{V}_1^1, \widehat{V}_1^2, \dots, \widehat{V}_1^n$ , based on the one-ray nucleator for each particle in a sample of  $n$  of cells from a cell population. Then one way to check the plausibility of the model would be to compare the empirical distribution of  $\widehat{V}_1^1, \widehat{V}_1^2, \dots, \widehat{V}_1^n$  to the distribution determined by simulating variables according to (i) in Theorem D.1 with parameter values estimated under the model. Similar model checking can of course be conducted based on measurement from either the  $N$ -ray nucleator or the integrator nucleator simply by applying the relevant parts of Theorem D.1.

## Bibliography

- Baddeley, A. and Jensen, E. B. V. (2005). *Stereology for statisticians*. Chapman & Hall, Boca Raton.
- Brix, A. (1999). Generalized gamma measures and shot-noise cox processes. *Advances in Applied Probability*, 31:929–953.
- Cruz-Orive, L. M. (1976). Particle size-shape distributions: the general spheroid problem. I. Mathematical model. *Journal of Microscopy*, 107:235–253.
- Cruz-Orive, L. M. (1978). Particle size-shape distributions: the general spheroid problem. II. Stochastic model and particle guide. *Journal of Microscopy*, 112:153–167.
- Gardner, R. J. (2006). *Geometric tomography*. Cambridge University Press, New York.
- Gardner, R. J. and Zhang, G. (1998). Affine inequalities and radial mean bodies. *American Journal of Mathematics*, 120(3):505–528.
- Gundersen, H. J. G. (1988). The nucleator. *Journal of Microscopy*, 151:3–21.
- Hansen, L. V., Nyengaard, J. R., Andersen, J. B., and Jensen, E. B. V. (2011a). The semi-automatic nucleator. *Journal of Microscopy*, 242(2):206–215.
- Hansen, L. V., Thorarinsdottir, T. L., and Gneiting, T. (2011b). Lévy particles: modelling and simulating star-shaped compact sets. *Manuscript*.
- Hobolth, A. and Jensen, E. B. V. (2002). A note on design-based versus model-based variance estimation in stereology. *Advances in Applied Probability*, 34(3):484–490.
- Jensen, E. B. V. (1998). *Local Stereology*. World Scientific Publishing, Singapore.
- Jensen, E. B. V. (2000). On the variance of local stereological volume estimators. *Image Analysis and Stereology*, 19:15–18.
- Jónsdóttir, K. Y., Schmiegel, J., and Jensen, E. B. V. (2008). Lévy based growth models. *Bernoulli*, 14:62–90.
- Sander, P. V. (2007). Ambient aperture lighting. In *I3D '07: Proceedings of the 2007 symposium on Interactive 3D graphics and games*, ACM. Press.
- Schneider, R. (1993). *Convex bodies: the Brunn-Minkowski theory*. University Press.
- Tovchigrechko, A. and Vakser, I. A. (2001). How common is the funnel-like energy landscape in protein-protein interactions? *Protein Science*, 10(8):1572–1583.

Wicksell, S. D. (1925). The corpuscle problem. A mathematical study of a biometric problem. *Biometrika*, 17:84–89.

Wicksell, S. D. (1926). The corpuscle problem. Second memoir. Case of ellipsoidal corpuscles. *Biometrika*, 18:152–172.

## Appendix

In the following we prove the explicit expressions (D.9), (D.10) and (D.11) for the MSE of  $\widehat{V}_1$ ,  $\widehat{V}_{1,N}$  and  $\widehat{V}_2$  under the power  $p$  Lévy based model with an isotropic kernel. For this we utilise equations (D.3) and (D.8).

For the proof, we only consider  $\widehat{V}_{1,N}$  and  $\widehat{V}_2$  as  $\widehat{V}_1$  is a special case of the first. For the  $N$ -ray nucleator we find for the conditional expectation given  $(E, U_1)$  that

$$\begin{aligned} \mathbb{E}((\widehat{V}_{1,N} - V(Y))^2 \mid U) &= \text{Var}\left(\int_{\mathbb{S}^2} \frac{4\pi}{3} \frac{1}{N} \sum_{i=1}^N k(\vartheta(v, U_i)) - \frac{1}{3} \Gamma(\mathrm{d}v) \mid E, U_1\right) \\ &= \sigma_{\Gamma}^2 \frac{1}{9} \int_{\mathbb{S}^2} \left(4\pi \frac{1}{N} \sum_{i=1}^N k(\vartheta(v, U_i)) - 1\right)^2 \mathrm{d}v \\ &= \sigma_{\Gamma}^2 \frac{4\pi}{9} \left(\frac{4\pi}{N^2} \int_{\mathbb{S}^2} \left(\sum_{i=1}^N k(\vartheta(v, U_i))\right)^2 \mathrm{d}v - 1\right). \end{aligned}$$

This is independent of  $(E, U_1)$ , as

$$\int_{\mathbb{S}^2} \left(\sum_{i=1}^N k(\vartheta(v, U_i))\right)^2 \mathrm{d}v = c_2 N \left(C(0) + 2 \sum_{j=1}^{\lfloor N/2 \rfloor} C(\pi \frac{j}{N}) + C(\pi) \mathbb{1}_{\{N \text{ even}\}}\right),$$

thereby justifying the result for the  $N$ -ray nucleator. Here,  $\lceil x \rceil$  denote the smallest integer not less than  $x$ .

Now consider the conditional expectation of  $\widehat{V}_2$  given  $E$ ,

$$\begin{aligned} \mathbb{E}((\widehat{V}_2 - V(Y))^2 \mid E) &= \text{Var}\left(\int_{\mathbb{S}^2} \frac{2}{3} \int_{\mathbb{S}^2 \cap E} k(\vartheta(v, u)) \mathrm{d}u - \frac{1}{3} \Gamma(\mathrm{d}v) \mid E\right) \\ &= \sigma_{\Gamma}^2 \frac{1}{9} \int_{\mathbb{S}^2} \left(2 \int_{\mathbb{S}^2 \cap E} k(\vartheta(v, u)) \mathrm{d}u - 1\right)^2 \mathrm{d}v \\ &= \sigma_{\Gamma}^2 \frac{4}{9} \left(\int_{\mathbb{S}^2} \left(\int_{\mathbb{S}^2 \cap E} k(\vartheta(v, u)) \mathrm{d}u\right)^2 \mathrm{d}v - \pi\right). \end{aligned}$$

Straight forward calculations show that the remaining integral can be reduced to an integral of the correlation function,

$$\int_{\mathbb{S}^2} \left(\int_{\mathbb{S}^2 \cap E} k(\vartheta(v, u)) \mathrm{d}u\right)^2 \mathrm{d}v = 4\pi c_2 \int_0^\pi C(\eta) \mathrm{d}\eta,$$

which concludes the argument.