ASPECTS OF SPATIAL STATISTICS, STOCHASTIC GEOMETRY AND MARKOV CHAIN MONTE CARLO METHODS

by

Jesper Møller
Department of Mathematical Sciences
Aalborg University
Fredrik Bajers Vej 7E
DK-9220 Aalborg Ø
Denmark
Email: `jm@math.auc.dk`

Denne afhandling er af Det Teknisk-Naturvidenskabelige Fakultet ved Aalborg Universitet antaget til forsvar for den naturvidenskabelige doktorgrad. Forsvaret finder sted på Aalborg Universitet i Auditorium B3-104, Fred. Bajers Vej 7 den 21. januar 2000 kl. 13.00 præcis.

Aalborg den 15. september 1999

Finn Kjærsdam
Dekan

This thesis has been accepted by the Faculty of Engineering and Science at Aalborg University for public defence in fulfilment of the requirements for the doctoral degree in Natural Sciences. The defence will take place at Aalborg University in Lecture Room B3-104, Fred. Bajers Vej 7 on January 21, 2000 at 1 p.m. punctually.

Aalborg, September 15, 1999

Finn Kjærsdam
Dean

Contents

Preface

This doctoral thesis is concerned with certain aspects of spatial statistics, stochastic geometry, and Markov chain Monte Carlo methods. It is based on a monograph [L] and twenty selected papers [A–K, M–V] produced over about the last decade; see the following two pages. The papers [A, G, H, I] and [95, 140] constituted my Ph.D. thesis from 1988 entitled *Stochastic Geometry and Markov Models*; [A, H, I] in revised form are included in the present thesis.

The thesis consists of three parts: (i) statistical models for lattice data, spatial (marked) point patterns, and spatial birth-and-death processes, (ii) Markov chain Monte Carlo methods and statistical inference for spatial models, and (iii) random tessellations. Parts (i) and (ii) are closely related and the focus is on recent theoretical and applied statistical aspects. Part (iii) is more within the traditional framework of stochastic geometry, though new results will be discussed as well.

I am very much indebted to many friends, collaborators, and colleagues throughout the world. A special thank you to my former colleagues at the Department of Theoretical Statistics, University of Aarhus, and my present colleagues at the Department of Mathematics, Aalborg University, for providing friendly and inspiring environments. In particular, Steffen Lilholt Lauritzen is acknowledged for his encouragement and request to me on preparing and submitting the present thesis. Finally, I owe a very special thank you to my wife, Birthe, and my sons, Søren and Peter.

Aalborg, March 1998 Jesper Møller

Thesis preface

Apart from updating a few references to articles which have appeared since the submission of the thesis, the following is essentially identical to a previous version of the manuscript.

Aalborg, October 1999 Jesper Møller

List of papers: This thesis is based on the following papers and monograph

[A] A. Baddeley and J. Møller. Nearest-neighbour Markov point processes and random sets. *International Statistical Review*, 2:89–121, 1989.

[B] A.J. Baddeley, M.N.M. van Lieshout, and J. Møller. Markov properties of cluster processes. *Advances in Applied Probability (SGSA)*, 28:346–355, 1996.

[C] C.J. Geyer and J. Møller. Simulation procedures and likelihood inference for spatial point processes. *Scandinavian Journal of Statistics*, 21:359–373, 1994.

[D] O. Häggström, M.N.M. van Lieshout, and J. Møller. Characterisation results and Markov chain Monte Carlo algorithms including exact simulation for some spatial point processes. *Bernoulli*, 5:641–659, 1999.

[E] H. Högmander and J. Møller. Estimating distribution maps from atlas data using methods of statistical image analysis. *Biometrics*, 51:393–404, 1995.

[F] J.L. Jensen and J. Møller. Pseudolikelihood for exponential family models of spatial point processes. *Annals of Applied Probability*, 3:445–461, 1991.

[G] J. Møller. A simple derivation of a formula of Blaschke and Petkantschin. Research Report 138, Department of Theoretical Statistics, Aarhus University, 1985.

[H] J. Møller. On the rate of convergence of spatial birth-and-death processes. *Annals of the Institute of Statistical Mathematics*, 3:565–581, 1989.

[I] J. Møller. Random tesselations in $\mathbf{R}^d$. *Advances in Applied Probability*, 21:37–73, 1989.

[J] J. Møller. Random Johnson-Mehl tessellations. *Advances in Applied Probability*, 24:814–844, 1992.

[K] J. Møller. Contribution to the discussion of N.L. Hjort and H. Omre (1994): Topics in spatial statistics. *Scandinavian Journal of Statistics*, 21:346–349, 1994.

[L] J. Møller. *Lectures on Random Voronoi Tessellations.* Lecture Notes in Statistics. Springer-Verlag, New York, 1994.

[M] J. Møller. Generation of Johnson-Mehl crystals and comparative analysis of models for random nucleation. *Advances in Applied Probability (SGSA)*, 27:367–383, 1995.

[N] J. Møller. Perfect simulation of conditionally specified models. *Journal of the Royal Statistical Society*, B 61:251–264, 1999.

[O] J. Møller. A review on probabilistic models and results for Voronoi tessellations. In P. Engel and H. Syta, editors, *Voronoi's Impact on Modern Science*, volume 1, pages 254–265, 1998. Institute of Mathematics of the National Academy of Sciences, Kyiv.

[P] J. Møller. Markov chain Monte Carlo and spatial point processes. In O.E. Barndorff-Nielsen, W.S. Kendall and M.N.M. van Lieshout, editors, *Stochastic Geometry: Likelihood and Computation*, pages 141–172, Bocan Raton, 1999. Chapman and Hall/CRC, London.

[R] J. Møller. Topics in Voronoi and Johnson-Mehl tessellations. In O.E. Barndorff-Nielsen, W.S. Kendall and M.N.M. van Lieshout, editors, *Stochastic Geometry: Likelihood and Computation*, pages 173–198, Bocan Raton, 1999. Chapman and Hall/CRC, London.

[S] J. Møller and M. Sørensen. Parametric models of spatial birth-and-death processes with a view to modelling linear dune fields. *Scandinavian Journal of Statistics*, 21:1–19, 1994.

[T] J. Møller, A.R. Syversveen, and R.P. Waagepetersen. Log Gaussian Cox processes. *Scandinavian Journal of Statistics*, 25:451–482, 1998.

[U] J. Møller and R.P. Waagepetersen. Markov connected component fields. *Advances in Applied Probability (SGSA)*, 30:1–35, 1998.

[V] J. Møller and S. Zuyev. Gamma-type results and other related properties of Poisson processes. *Advances in Applied Probability (SGSA)*, 28:662–673, 1996.

1 Introduction

Spatial statistics and stochastic geometry are concerned with probabilistic and statistical modelling and analysis of spatial data such as digital images, spatial point patterns, and more general random geometrical structures. Stochastic geometry emerged from integral geometry and geometrical probability in the late sixties (see D.G. Kendall's foreword in [180]); historically, Buffon's [31] needle problem from 1777 is considered as the start of geometrical probability theory, while the closely related area of integral geometry was founded by Blaschke and his school in the 1930s. The history of spatial statistics goes at least back to the beginning of this century when Student [186] analysed a certain point pattern by considering a Poisson hypothesis for the counts of points within small squares (incidentally one of the earliest applications of a χ^2-test for goodness of fit). However, a substantial progress in spatial statistics was first initiated with the introduction of Markov random fields and Markov chain Monte Carlo (MCMC) methods in the seventies, and the access to fast and cheap computers during the last decades.

Presently, the phrases stochastic geometry and spatial statistics are often used synonymously. Both areas have topics such as spatial point processes and more general random sets models (particularly the Boolean model) in common. It still appears that random tessellations and stereology are mostly studied within a truly stochastic geometrical framework, whilst random field models (especially in relation to Bayesian image analysis) and geostatistical methods are more applied in connection to spatial statistical analysis. Many ideas in spatial statistics originate from statistical physics, and as mentioned Markov chain Monte Carlo computations play a dominant role. Thereby the focus on non-parametric methods for very general spatial models has changed over the years towards exploiting more and more flexible and complex parametric statistical models, which are analysed within a Bayesian or likelihood approach by means of MCMC methods. In my view many stochastic geometers are still not so much aware of the potential use of MCMC methods, though these techniques play an increasing role in the second edition of the monograph [180].

MCMC methods originated from statistical physics with the seminal paper by Metropolis *et al.* [132] from 1953, where the first MCMC algorithm for simulating a Gibbs point process was proposed. Actually in statistics some of the earliest and most important applications of MCMC methods appear to be within spatial statistics (see the discussion in [19]). The Gibbs sampler [61], which is now widely used in statistics, has first been introduced implicitly in spatial statistics (see [187, 168, 169]) and in statistical physics (see, e.g., [44] where it is named 'the heat bath algorithm'). Indeed, today MCMC procedures are applied on a broad spectrum of complex problems in statistics, particularly in Bayesian inference.

The monograph [180] by Stoyan, Kendall and Mecke provides an excellent introduction to the current stage of stochastic geometry and spatial statistics; other useful books on these topics include [43, 45, 46, 125, 126, 130, 170, 171, 181]; MCMC methodology is treated in [72, 12], where [12] is concerned with stochastic geometry and spatial statistics as well. A book which collects and unifies recent theoretical advances and examples of applications in spatial statistics, stochastic geometry, and MCMC computations is still missing.

This thesis deals only with certain aspects of current research in spatial statistics, stochastic geometry, and MCMC methods. The thesis is based on the contributions [A–V], which may naturally be divided in accordance to the type of spatial process that is mainly considered:

random field models/lattice processes:	[E, N, U];
point processes:	[A, B, C, D, F, K, P, T];
spatial birth-and-death processes:	[H, S];
random tessellations:	[G, I, J, L, M, O, R, V].

(In fact [G] is concerned with a purely integral geometric result and [V] with some general probabilistic properties of Poisson processes, so the results of both papers find applications in many other areas than just random tessellations.)

Section 2 contains a survey on statistical models for lattice data, spatial point patterns, and spatial birth-and-death processes, while MCMC procedures and statistical inference for such models are discussed in Section 3. Both probabilistic and some statistical aspects of random tessellations, particularly Voronoi and Johnson-Mehl tessellations, are reviewed in Section 4. A few illuminating pictures are shown in Sections 2–4; many others can be found in [A–V] and the textbooks mentioned above.

2 Statistical models for lattice and point processes

In this section it is discussed how to construct usable models for lattice data and spatial point patterns, with greatest focus on the latter type of data. What is meant by 'usable' depends, of course, very much on the objective of the statistical analysis and the type of data which is available. Various classical and more recent types of Markovian models are studied in Sections 2.1–2.2. In Section 2.3 we consider briefly spatial birth-and-death processes. Finally, Section 2.4 deals with spatial heterogeneity and the class of log Gaussian Cox processes.

2.1 Modelling spatial interaction in lattice processes with a view to Bayesian image analysis

The primary impetus for modelling spatial interaction comes from statistical physics, where one attempts to explain the macroscopic behaviour of a system by its microscopic structure. In statistical mechanics, one considers systems so large that it makes sense to study infinite systems called Gibbs measures (introduced by Dobrushin, Lanford and Ruelle in the late sixties). One interesting property of a Gibbs measure for a given type of interaction is that it may fail to be unique as briefly exemplified below by means of the celebrated Ising [94] process (more precisely, phase transition behaviour is an interesting property in statistical mechanics, while as explained later on it is something we usually want to avoid for applications in spatial statistics and stochastic geometry). In the following we discuss firstly the role of the Ising and other Markov random field models in the context of Bayesian image analysis, and secondly we review some properties of another class of lattice processes called Markov connected component fields.

Consider a binary image made up of a finite rectangular array $I \subset \mathbf{Z}^2$ of pixels so that the state space of possible images is given by $\{0,1\}^I$. We refer to the pixels of zeros as the white pixels, and the others as the black pixels. In the homogeneous case (in the absence of an external magnetic field, etc.), the probability mass function of the *Ising model* is given by

$$p(x) \propto \exp(-\beta \sum_{\{i,j\} \subset I : i \sim j} \mathbf{1}[x_i \neq x_j]) \tag{1}$$

where β is a real parameter, $\sim$ is a symmetric relation defined on I, let us say the first-order relation $\sim = \sim^1$ defined by $i \sim^1 j \Leftrightarrow ||i - j|| = 1$, and $\mathbf{1}[\cdot]$ denotes indicator function. Usually, β is assumed to be positive (the attractive case) so that neighbouring pixels tend to take the same value; we make the same assumption in the following.

By a so-called local specification, it is possible to extend the definition of the Ising model to a binary process on the entire lattice $\mathbf{Z}^2$: Consider the local characteristics,

$$p(x_i | x_{\partial i}) = \exp(-\beta \sum_{j : j \sim i} \mathbf{1}[x_i \neq x_j]) / c(x_{\partial i}), \quad i \in \mathbf{Z}^2, \; x = (x_j)_{j \in \mathbf{Z}^2} \in \{0,1\}^{\mathbf{Z}^2} \tag{2}$$

where, as $\sim = \sim^1$, $\partial i = \{j \in \mathbf{Z}^2 : ||i - j|| = 1\}$ consists of the four horizontal and vertical neighbours to i, $x_{\partial i} = (x_j)_{j \in \partial i}$, and

$$c(x_{\partial i}) = \exp(-\beta \sum_{j : j \sim i} \mathbf{1}[x_j = 0]) + \exp(-\beta \sum_{j : j \sim i} \mathbf{1}[x_j = 1])$$

is the normalizing constant. It can be shown (see, e.g., [110]) that any probability mass function p^{ext} defined on $\{0,1\}^{\mathbf{Z}^2}$ and with local characteristics (2) — that is $p^{ext}(x_i | x_j, j \neq i) = p(x_i | x_{\partial i})$ for all $i \in \mathbf{Z}^2$ and $x \in \{0,1\}^{\mathbf{Z}^2}$ — is a convex combination of two stationary probability mass functions p^+ and p^-. Here stationarity means invariance under translations on $\mathbf{Z}^2$ and, for example, p^+ is obtained as the limiting measure by considering a sequence of increasing boxes with 1s on the boundary. Moreover, there is a critical value $\tilde{\beta} \approx .88$ so that $p^+ = p^-$ if $0 \leq \beta \leq \tilde{\beta}$ and $p^+ \neq p^-$ otherwise (the case of a phase transition). Hence, when $\beta > \tilde{\beta}$, the boundary has an influence even in the limit.

Returning to the Ising model (1) defined on a finite large lattice, as β increases a phase transition behaviour occurs as an abrupt transition from unimodality to bimodality of the distribution of the number of black pixels. This means that typical realisations with a subcritical value of β look very chaotic, while in the supercritical case most realisations are dominated by one (or a few) component(s) of the same colour together with a lot of very small components. Nevertheless, the Ising model has frequently and successfully been used as a prior model in Bayesian image analysis for low level tasks such as image restoration and segmentation problems (see [61, 21]). In general, when restoration procedures like maximum a posteriori estimation, maximum marginal a posteriori and iterative conditional modes are applied, the Ising model serves not as a realistic prior, but as a penalizing term which locally favours smooth boundaries, whereas the large scale properties of the estimated image are mostly (or hopefully) determined by the data via the likelihood term of the posterior. If a more detailed Bayesian analysis is requested, and there is a certain structure in the true image, one needs to construct more complex models such as the Markov connected component models to be discussed below.

One example, among many others, where an Ising prior seems natural, is in the breeding bird atlases studied in [E]. No structure in the atlas maps is present except that similarity between nearby 'pixels' is expected. Atlas mapping is a common method to present biogeographical ranges, and the use of statistical methods of Bayesian image analysis in this context seems first to be introduced in [E]; as discussed in [E], though the estimation of biogeographical distributions can be regarded as a restoration of a binary image, there are many features different from the standard applications of image analysis. This has been further developed in [84].

The Ising model is one example of a *Markov random field* (MRF). Restricting attention to the finite state space case $\Omega = \{0, 1, \ldots, k\}^I$ of a random field $X = (X_i)_{i \in I}$ with an arbitrary but finite index set I, we say that X (and its distribution) is an MRF if, for any $i \in I$ and $x \in \Omega$

with $P(X_{-i} = x_{-i}) > 0$, $p(x_i|x_{-i}) = P(X_i = x_i|X_{-i} = x_{-i})$ depends on $x_{-i} = (x_j)_{j:j\neq i}$ only through $x_{\partial i}$. Here and elsewhere in the following, the neighbourhood of site i, $\partial i = \{j \in I \setminus \{i\} : j \sim i\}$ is defined with respect to an arbitrary symmetric relation $\sim$ on I. According to the Hammersley-Clifford theorem, if the positivity condition $p(x) = P(X = x) > 0$ is satisfied for all $x \in \Omega$, then X is an MRF if and only if

$$p(x) \propto \prod_{C \in \mathcal{C}} \psi_C(x_C) \quad \forall x \in \Omega \tag{3}$$

where $\mathcal{C}$ is the set of non-empty cliques defined with respect to $\sim$ and ψ_C is a positive function. Clifford [36] discusses the history and further development of this characterisation result which goes back to 1970.

Besag [19] is responsible for introducing MRFs in the context of spatial statistics. Mostly pairwise interaction models with a simple parameterization have been used, i.e. when $\psi_C(x_C) = 1$ whenever $|C| \geq 3$ (where $|\cdot|$ denotes cardinality), but it is increasingly acknowledged that such models are not appropriate to describe the global structures in real images. Tjelmeland and Besag [193] demonstrate that interesting and flexible models can be obtained when higher order interaction terms are included, but then the number of parameters to be estimated may grow rapidly. Künsch, Geman and Kehagias [113] have showed that any stationary discrete random field can be approximated arbitrarily closely by hidden MRFs, but they admit that their examples of applications are not so convincing.

The *Markov connected component fields* (MCCF) introduced in [U] provide models which are flexible and model the global structure more directly without necessarily needing a high-dimensional parameterization. The general form of an MCCF is given by

$$p(x) \propto \prod_{K \in \mathcal{K}(x)} \Psi_K(l(x_K)) \tag{4}$$

where $\mathcal{K}(x)$ is the set of maximal connected components in the graph with vertices $V = I$ and edges $E = \{\{i,j\} : i \sim j,\, x_i = x_j,\, i \neq j\}$, $l(x_K)$ is the common value of x_i for $i \in K$, and Ψ_K is a non-negative function. In fact, under the positivity condition $p(\cdot) > 0$, any binary MRF is a MCCF. If in (4), $\Psi_K(0) = 1$ for all connected components K, then we have an MCCF with background colour 0, which can be seen to be a special case of the nearest-neighbour Markov point processes introduced in [A]; see also Section 2.2. For such MCCFs there is a characterisation result similar to (3) or more precisely to (14) below (see [U, Theorem 3]).

For example, the Ising model (1) (with $\sim = \sim^1$) is an MCCF with $\Psi_K(l) = \exp(-\beta u(K))$, where $u(K) = |\{(i,j) \in K \times (\mathbf{Z}^2 \setminus K) : i \sim^1 j\}|$ is the perimeter of K (absorbing 1/2 into β). This can be modified to the *penalized Ising model* given by $\Psi_K(l) = \exp(-\beta u(K) - \gamma/a(K))$, where $a(K) = |K|$ is the area of K. Taking $\gamma > 0$, small components are penalized. Another example of a MCCF of considerable intrinsic interest is the random cluster model [55, 77], which is closely connected to the Ising model. The penalized Ising model and the random cluster model are not MRFs.

It is demonstrated in [U] how the large scale variability of a microscopic image of a rock sample is well described by the penalized Ising model but not by the Ising model. Consider the binary 62×67 image in Figure 1 (top left) taken from [U, Figure 1]. The maximum likelihood estimate of β (obtained by MCMC methods as discussed in Section 3.5) is supercritical under the Ising model ($\hat{\beta} = .481$), but according to [196] subcritical under the penalized Ising model ($\hat{\beta} = .483$ and $\hat{\gamma} = 17.643$). The estimated Ising model is a poor representation of the variability present in the rock sample image, as realisations will be dominated by one of the colours and feature a large number of small components. Simulations of the estimated penalized Ising model (one example is shown in Figure 1 (top right)) show that the model captures the variability in the true image rather well, though the boundaries in the rock sample image appear to be somewhat smoother. Further, as a Bayesian image analysis experiment, we considered in [U] simulated data obtained by adding a Gaussian white noise with variance 1.25 to the the rock sample image; the expected rate of misclassified pixels is 32.9%. As priors we used the estimated Ising and penalized Ising models, respectively (using the maximum likelihood estimates as prior parameters corresponds to the case where training data is available). Figure 1 shows the

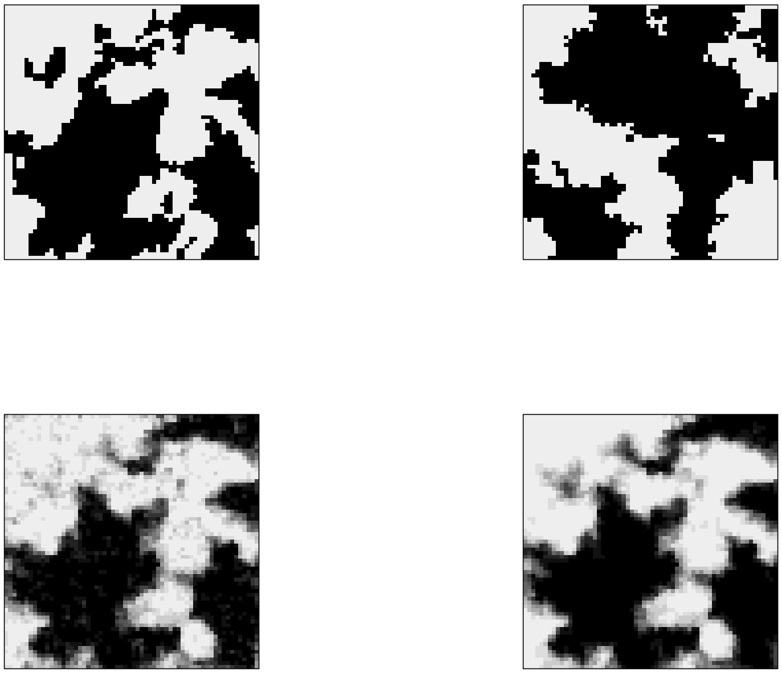

Figure 1: Top left: Rock sample image. Top right: A realisation of the estimated penalized Ising model. Bottom left: Marginal posterior probabilities of observing a black pixel under the Ising prior. Bottom right: As bottom left but under the penalized Ising prior.

marginal posterior probabilities of observing a black pixel (again obtained by MCMC methods) under the Ising prior (bottom left) and the penalized Ising prior (bottom right). Clearly, the results for the penalized Ising model are in much better accordance with those obtained by the Ising model. One may object that better results might be obtained with another value of β in the Ising prior, but it is not clear at all according to which criteria it should be chosen, and if training data is available, estimation of prior parameters from such data seems very natural. So this example stresses the importance of a realistic prior modelling in a 'true' Bayesian analysis.

The intersection between the classes of MRFs and MCCFs is characterised in [U, Theorem 1]. For instance, suppose that $\sim = \sim^k$ is the kth-order relation given by $i \sim^k j$ if and only if $||i-j||^2 \leq k,\ k = 1, 2, \dots$, and consider an MCCF which is also a homogeneous MRF in the sense that (3) holds with $\psi_C = \psi_{\acute{C}}$ whenever C and $\acute{C}$ are cliques which are isomorphic with respect to motions. Then the density is of an exponential family type with a sufficient statistic given by certain geometrical characteristics: If $\sim = \sim^1$, then

$$p(x) \propto \prod_{K \in \mathcal{K}(x)} \exp(\alpha_{l(x_K)} a(K) + \beta_{l(x_K)} u(K))$$

for some real parameters $\alpha_l,\ \beta_l,\ l = 0, \dots, k$. If $\sim = \sim^2$, then the form of the density is extended by adding the following term in the exponent above,

$$\gamma_{l(x_K)} k_+(K) + \delta_{l(x_K)} d(K) + \epsilon_{l(x_K)} \chi(K)$$

where $\gamma_l,\ \delta_l,\ \epsilon_l$ are real parameters and $k_+(K),\ d(K),\ \chi(K)$ are the number of concave corners, number of $\sim^1$ discontinuities, and Euler-Poincaré characteristic of K, respectively. These models have some similarity to the Quermass interaction processes for random sets studied in [8]. In fact, $(a(K), u(K), k_+(K), d(K), \chi(K))$ is in a linear one-to-one correspondence with $(n_0(K), \dots, n_4(K))$, where $n_0(K) = a(K)$, $n_1(K)$ is the number of pairs of horizontal and vertical neighbouring pixels contained in K, $n_2(K)$ is the number of pairs of diagonal neighbouring pixels contained in K, $n_3(K)$ is the number of triplets of second-order neighbouring pixels contained in K, and $n_4(K)$ is the number of 2×2 squares of pixels contained in K. Hence, if a pixel is added to or removed from a component for which the geometrical characteristics are known, the calculation of the geometrical characteristics of the new component requires only 'local information'. This makes it advantageous to use MCCFs with $(a(K), u(K), k_+(K), d(K), \chi(K)),\ K \in \mathcal{K}(x)$ as a sufficient statistic when single-site updating Metropolis-Hastings algorithms [81] are employed.

Further properties of MCCFs are investigated in [U]. Briefly, we study local and global Markov properties of MCCFs, extensions to infinite lattices, and estimation and modelling of images using MCCFs in various examples of applications. The discussion on infinite volume Gibbs measures is rather technical as connected components may be unbounded, and the main problems are to give a meaningful definition of a local specification as a natural limit of conditional MCCFs defined on a sequence of increasing bounded lattices and to establish quasi-locality of this specification so that existence results in [63, 161] can be applied. Uniqueness of Gibbs measures for the random cluster and penalized Ising models are investigated in [77, 196], but a study of phase transition behaviour for other types of MCCFs which are not MRFs remains to be done.

The application of MCCFs is somewhat more computationally demanding than the application of MRFs. By the local Markov property the calculations for updating one pixel only involves the components adjacent to the pixel and as mentioned above the calculation of the geometrical characteristics of the new component is not computationally demanding. The computational expense is due to deciding whether a pixel update leads to the division of a component or to the unification of a number of components. Sweeny [188] gives a very efficient algorithm for the random cluster model, but it is not applicable for MCCFs in general. We use in [U] an algorithm which works for any MCCF and which is based on standard graph algorithms [40].

In conclusion MCCFs are applicable for images with a relative few number of colours/grey levels and where <u>some</u> prior knowledge is available about e.g. size and shape characteristics of the connected components of constant colour in the image. Compared to MRFs, global information can more easily be incorporated in MCCFs.

One drawback of the discrete setup is the high dimensionality in many applications of image analysis. If more prior knowledge is available, e.g. that the scene consists of geometrical objects (or deformations) of a particular type like discs or polygons, other approaches may be more appropriate including the following: the marked point process approach in [10]; or the more flexible but technical approach of deformable template matching in [76]; or the elements of both these approaches as combined in [175]; or other types of polygonal models as investigated in [2, 150, 159].

2.2 Modelling spatial interaction in point processes

In the following we consider for simplicity finite point processes defined on a Borel set $S \subset \mathbf{R}^d$ with $0 < \lambda(S) < \infty$ where λ denotes Lebesgue measure. But it should be noticed that most concepts and results presented can be generalized to a more general setting with, for example (and covering most situations of interest), S a Polish space and λ a finite positive measure. Thereby point processes with multiplicities and marked point processes can be included as well; point processes with multiplicities are not so important from an applied point of view but they are useful for technical reasons; but marked point processes such as multitype point patterns, line segment processes, disc processes, etc., find numerous examples of applications; see, for example, [A] and [46, 180]. Extensions of point processes to $\mathbf{R}^d$ will be briefly discussed later on in this section.

The Poisson process plays a fundamental role as a reference process. Let $n(x) = |x|$ denote the cardinality of a finite point configuration $x = \{x_1, \dots, x_n\} \subset S$; if $n = 0$ we write $x = \emptyset$ for the empty configuration. Letting $\Omega = \Omega_S = \{x \subset S : n(x) < \infty\}$, the state space of simple finite point configurations contained in S (and equipped with a natural σ-field, see [A] and [32]), then the unit rate Poisson process $\mu = \mu_S$ on S is the probability measure on Ω so that if $X \sim \mu$, then $n(X)$ follows a Poisson distribution with mean $\lambda(S)$ and conditional on $n(X)$ the points in X are i.i.d. and uniformly distributed on S. Because of the latter property, the

Poisson process is said to exhibit no interaction (or CSR — complete spatial randomness). In this section we consider point process models $X \sim P$ which are absolutely continuous with respect to μ and we denote the density by $f = \mathrm{d}P/\mathrm{d}\mu$.

Usually point patterns are broadly classified as clustered or regular compared to CSR (see, e.g., [46, 180]). Often a regular or clustered behaviour is expected to be explained by some sort of a local Markov property expressed in terms of the Papangelou conditional intensity [102] given by

$$\lambda^*(x, \xi) = f(x \cup \{\xi\})/f(x), \quad x \in \Omega, \ \xi \in S \setminus x$$

(taking $a/0 = 0$ for $a \geq 0$). Thus the role of the Papangelou conditional intensity is similar to that of the local characteristics of a lattice process as discussed in Section 2.1. Typically, we interpret

$$y \subset x \ \Rightarrow \ \lambda^*(y, \cdot) \geq \lambda^*(x, \cdot) \tag{5}$$

as the *repulsive case* (implying regularity) and

$$y \subset x \ \Rightarrow \ \lambda^*(y, \cdot) \leq \lambda^*(x, \cdot) \tag{6}$$

as the *attractive case* (implying clustering or aggregation).

Henceforth, for ease of exposition we assume that f is hereditary: $f(x) > 0 \Rightarrow f(y) > 0$ whenever $y \subset x \in \Omega$. Then there is a one-to-one correspondence between f and λ^*, and the hereditary condition replaces the positivity condition in the Hammersley-Clifford characterisation result (3). We shall often just specify f up to proportionality $f \propto g$ and implicitly assume that f is well-defined, that is, g is integrable with respect to μ. A sufficient (and in many cases a necessary) condition is Ruelle stability [176] meaning that $g(x) \leq \alpha\beta^{n(x)}$ for some positive constants α, β and all $x \in \Omega$. A stronger condition than Ruelle stability is *local stability*: there is a constant $K > 0$ so that

$$\lambda^*(x, \xi) \leq K \quad \forall x \in \Omega, \ \xi \in S \setminus \{x\}. \tag{7}$$

Local stability is usually easily verified. It is satisfied by many point process models (see [67, 109]) including the specific models discussed below, but one example, where Ruelle but not local stability may be satisfied, is a pairwise interaction point process with density

$$f(x) \propto \beta^{n(x)} \prod_{\{\xi,\eta\}\subseteq x:\xi\neq\eta} g(||\xi-\eta||)$$

where $\beta > 0$ is a parameter and g is a strictly positive function with $g(r)$ very close to 0 for small values of $r > 0$ and strictly larger than 1 for $0 < \epsilon < r < R$. Incidentally, as discussed further in Section 3, local stability also plays an important role in the simulation algorithms studied in [C] and [67, 109].

In the classical Ripley-Kelly [172] *Markov point processes* (MPP), $\lambda^*(x,\xi)$ depends on x only through the neighbours $\eta \sim \xi$ in x to ξ, where $\sim$ is a given symmetric and reflexive relation. Frequently, one considers interactions of finite range $R \in]0,\infty[$ so that $\xi \sim \eta$ if and only if $||\xi-\eta|| \leq R$. A characterisation result for general MPPs, which is similar to the Hammersley-Clifford theorem (3) is verified in [172]: f is Markov (in the Ripley-Kelly sense) if and only if

$$f(x) = \prod_{y\subseteq x} \phi(y) \quad \forall x \in \Omega \tag{8}$$

where ϕ is an interaction function, that is, $\phi(y) = 1$ if y is not a clique (i.e., $\xi \not\sim \eta$ for some $\xi, \eta \in y$), and $f(\emptyset)$ is a normalizing constant. Ripley and Kelly used (8) to establish a spatial Markov property of MPPs.

The standard example of an MPP is the *Strauss process* [184, 105] with density

$$f(x) \propto \beta^{n(x)}\gamma^{s(x)} \tag{9}$$

(taking $0^0 = 1$) where $s(x) = |\{\{\xi,\eta\} \subseteq x : \xi \sim \eta, \xi \neq \eta\}|$ is the number of unordered pairs of neighbours in x and $\beta > 0, 0 \leq \gamma \leq 1$ are parameters (as shown in [105] the model is not well-defined for $\gamma > 1$). The Strauss process is repulsive: if $\gamma = 1$, we get a Poisson process on S of rate β; as γ decreases, realisations tend to be more and more regular; and for $\gamma = 0$ we have a (Markov) hard-core point process. Incidentally, in a current research project [124], we have shown that in contrast to many other types of hard-core point process models like Matérn hard-core point processes and the simple sequential inhibition process, the planar Strauss process with $\gamma = 0$ can produce highly regular patterns with a packing density as close as one wishes to the maximum possible packing density (about 90.7 percent when $d = 2$).

Actually, Strauss [184] wanted to construct a model for clustering, so he considered the attractive case $\gamma > 1$ which makes sense if we fix the number of points to $n(x) = n$ (with $n > 0$ a given integer) and consider (9) as a density with respect to Lebesgue measure on S^n (note that this 'conditional' density depends only on the parameter γ). However, as discussed in [K, P] and [69] the 'attractive Strauss process' is a rather uninteresting model: as γ increases, simulated realisations have an abrupt transition from 'CSR-like' patterns (γ close to 1) to tightly clustered patterns.

It is well known that MPPs provide a rich class of models for regular patterns, but not so much for modelling clustering as discussed in [P, Section 4.2]. One exception is the *area-interaction process* [11, 197] (or more generally, the Quermass interaction processes in [8]), which has interactions of arbitrarily high order. The density of the area-interaction process is

$$f(x) \propto \beta^{n(x)} \gamma^{-U_x} \tag{10}$$

where U_x is the area (or volume) of the union of balls with radii R and centers given by the points in x, and where $\beta > 0, \gamma > 0$ are parameters. If $\gamma < 1$ we have repulsion; the hard-core process may be considered as the limiting case $\gamma \to 0$; and $\gamma > 1$ is the attractive case (see Figure 2, Section 3.2).

The *nearest-neighbour Markov point processes* (NNMPP) introduced in [A] include the class of MPPs, but also many other types of point processes including certain cluster processes as shown in [B]. In [P, Section 4.3] it is argued that NNMPPs can be used as flexible models for clustered as well as regular point patterns.

The definition of an NNMPP is slightly more technical than that for an MPP. Omitting a few technical details (see [A]), suppose that for each $x \in$

Ω, a symmetric reflexive relation $\sim_x$ is defined on the points in x, where $\sim_x$ is allowed to depend on the point configuration x. This situation includes the Ripley-Kelly case where $\sim_x=\sim$ does not depend on x. Examples where the dependence of the configuration is taken into account include sequential neighbours on the real line ($\xi \sim_x \eta \Leftrightarrow]\xi,\eta[\cap x = \emptyset$); Dirichlet or Voronoi neighbours ($\xi \sim_x \eta \Leftrightarrow C(\xi|x) \cap C(\eta|x) \neq \emptyset$ where $C(\xi|x) = \{\kappa : ||\kappa - \xi|| \leq ||\kappa - \eta||\ \forall \eta \in x\}$ is the Voronoi cell associated to the nucleus ξ in the Voronoi tessellation generated by x); and the connected component relation defined by

$$\xi \sim_x \eta \Leftrightarrow \xi = y_1 \sim y_2 \sim \cdots \sim y_n = \eta \quad \text{for some} \quad y = \{y_1, y_2, \ldots, y_n\} \subseteq x \tag{11}$$

where $\sim$ is a given symmetric reflexive relation on S — this is very similar to the neighbour relation used for MCCFs (see Section 2.1). Now, a density f specifies an NNMPP if $\lambda^*(x,\xi)$ depends only on ξ, $N(\xi|x \cup \{\xi\})$ and the restrictions of the relations $\sim_x$, $\sim_{x\cup\{\xi\}}$ to $N(\xi|x \cup \{\xi\}) = \{\xi\} \cup \{\eta \in x : \xi \sim_{x\cup\{\xi\}} \eta\}$, the $x \cup \{\xi\}$ neighbourhood to ξ.

One example of an NNMPP which is not an MPP is the *continuum random cluster process* independently introduced in [K] and (for a particular case) in [111, 33] (see also [D, P]). Its density is

$$f(x) \propto \beta^{n(x)} \gamma^{-c(x)} \tag{12}$$

where $c(x)$ is the number of maximal cliques (or 'clusters') defined by $\sim_x$ in (11), and where β, γ are positive parameters; for $0 < \gamma < 1$ we have regularity (see Figure 2, Section 3.2), and for $\gamma > 1$ clustering (though neither (5) nor (6) holds). Simulations reported in [P] indicate that the behaviour of the continuum random cluster process with fixed $n(x) = n$ does not undergo the abrupt transition observed in the conditional Strauss process when increasing the parameter γ from 1 to infinity.

A characterisation of NNMPPs of the Hammersley-Clifford type is established in [A, Theorem 4.13]. Omitting again some details, certain natural consistency conditions are required (and satisfied by the examples of relations $\sim_x$ mentioned above). Then f is Markov (in the sense of Baddeley-Møller) if and only if

$$f(x) = \prod_{y \subseteq x} \phi(y)^{\chi(y|x)} \quad \forall x \in \Omega \tag{13}$$

where ϕ satisfies some natural conditions and $\chi(y|x) = \mathbf{1}[\xi \sim_x \eta \ \forall \xi, \eta \in y]$ is a clique indicator function. This characterisation result includes (8) as a special case and has been used in [107] to establish a spatial Markov property of NNMPPs (see also [P, Section 4.3]). In the special case where $\sim_x$ is the Dirichlet relation and $d = 2$, (13) implies that the density of an NNMPP contains interaction terms corresponding to the nuclei, edges, and vertices of the Voronoi cells (see [A] and [14]).

Another application of (13) is [B, Lemma 1] which states that a point process is an NNMPP with respect to the connected component relation (11) if and only if

$$f(x) \propto \prod_{y \subseteq \mathcal{C}(x)} \Phi(y) \tag{14}$$

where $\mathcal{C}(x)$ is the set of 'clusters' defined by $\sim_x$ and $\Phi \geq 0$ satisfies certain regularity conditions (if $f > 0$ then $\Phi > 0$ is the only condition). Clearly, the continuum random cluster process (12) is of the form (14). In [B], equation (14) is the basis for showing that if an MPP or NNMPP (with respect to (11) and when $\sim$ is the usual finite range relation) is used as the parent process for a cluster process, and the clusters are uniformly bounded and non-empty, then the cluster process is again an NNMPP (if the parent process is Poisson, the result remains true even if the clusters are allowed to be empty). Notice that the density (14) is analogous to that for a Markov connected component field (4) with background colour 0.

For modelling purposes it is useful to establish further characterisation results in terms of the Papengelou conditional intensity. Kelly and Ripley [105] verified (under a weak condition on S and $\sim$) that the Strauss process (9) is uniquely characterised by the property that $\lambda^*(x, \xi)$ depends only on the number of neighbours to ξ in x. Similar results have been established for the continuum random cluster model (12) in [D, K, P] and for the area-interaction process (10) in [D].

We turn now briefly to Gibbs point processes defined on $\mathbf{R}^d$ using, as in Section 2.1, a local specification in the sense of [161]. For example, for an MPP with interaction functions ϕ of finite range $R > 0$, the local specification is given by

$$f_B(x_B|x_{\partial B}) = c_B(x_{\partial B}) \prod_{y \subseteq x_B \cup x_{\partial B}:\, y \cap x_B \neq \emptyset} \phi(y),$$

for all $B \in \mathcal{B}$, $x_B \in \Omega_B$, $x_{\partial B} \in \Omega_{\partial B}$. Here $f_B(x_B|x_{\partial B})$ is a density with respect to μ_B, $\partial B = \{\eta \in B^c : \eta \sim \xi$ for some $\xi \in B\}$ is the neighbourhood to B, $c_B(x_{\partial B})$ is a normalizing constant, and $\mathcal{B}$ is the class of bounded Borel sets in $\mathbf{R}^d$. A point process on $\mathbf{R}^d$ with conditional densities $f_B(x_B|x_{B^c})$ so that, for any locally finite point configurations $x_B \subset B$ and $x_{B^c} \subset B^c$, $f_B(x_B|x_{B^c}) = f_B(x_B|x_{\partial B})$, is said to be Gibbsian. Using the results in [161] it is not so hard to verify that both the Strauss process (9) and the area-interaction process (10) can be extended to stationary Gibbs measures (see, e.g., [11, Theorem 1]), where stationarity means invariance of the measure with respect to translations in $\mathbf{R}^d$. The special case of the continuum random cluster process (12) with $1/\gamma = 2, 3, \ldots$ can be derived from the Widom-Rowlinson model (extended to $1/\gamma$ types of points; see Section 3.2 and [P, Section 4.3]), and we know that the Widom-Rowlinson model and hence also the continuum random cluster process can be extended to stationary Gibbs measures. The same is true for many other NNMPPs as recently established in [15].

Another related question is phase transition (defined in the same way as in Section 2.1). To the best of my knowledge, it is only known that a phase transition occurs in the Widom-Rowlinson model (and hence in the corresponding continuum random cluster model mentioned above, but also in the attractive area-interaction process) and the more general continuum Potts models studied in [64]; see also Section 3.2.

2.3 Spatial birth-and-death processes

A spatial point process can often be viewed as the equilibrium distribution of a spatial birth-and-death process. This can be used as the basis of Monte Carlo simulation as discussed in Sections 3.1–3.2. In [S] it is demonstrated that spatial birth-and-death processes can also be useful for modelling certain dynamical spatial processes. Basic properties of spatial birth-and-death processes and the models studied in [S] are briefly discussed below.

Preston [162] introduced a *spatial birth-and-death process* as a time-homogeneous Markov process $\{X(t) : t \geq 0\}$ defined on a state space $\Omega = \Omega_S$ of finite point configurations, where the transitions are of two types: either an addition ('birth') of a point in S to the process or a deletion ('death') of an existing point of the process. Simple conditions ensuring unique existence and convergence to a limit distribution of a spatial birth-and-death process can be found in Preston's paper; see also the reviews in [A, Section 2] and [H, Section 2]. Further simple conditions for establishing geometrically fast convergence to equilibrium are given in [H, Section 3]; to the best of my knowledge similar results have only been established in [116] for the special and simpler case of a hard-core birth-and-death process. The proofs of Preston's and my own results are based on a coupling to an ordinary 'dominating' birth-and-death process defined on the non-negative integers — a similar coupling is used in the recent papers [108, 109] on perfect simulation (see Section 3.2).

To be more specific, consider the common situation where, as at the beginning of Section 2.2, we are given a density f of a point process on a bounded Borel set $S \subset \mathbf{R}^d$. Assume that existing points of $X(t) = x$ are deleted with rate 1, while a newborn point $\xi \in S$ is added to x with rate $B(x) = \int_S \lambda^*(x, \eta) \mathrm{d}\eta$, where ξ has density $\lambda^*(x, \cdot)/B(x)$ and λ^* is the Papangelou conditional intensity for f. More precisely, suppose that we condition on the event that $X(t) = x = \{x_1, \ldots, x_n\}$ and let $E_0, E_1, \ldots, E_n$ be mutually independent (and independent of past history) exponentially distributed variates with means $1/B(x), 1, \ldots, 1$, respectively (for $B(x) = 0$ we set $E_0 = \infty$). Then the next transition happens at time $t + E$, where

$E = \min\{E_0, E_1, \ldots, E_n\}$. If $E = E_0$ then $X(t+E) = x \cup \{\xi\}$ represents a birth of the point ξ. If $E = E_i$ with $i > 0$ then $X(t+E) = x \setminus \{x_i\}$ represents the death of the point x_i. All conditions in [162] and [H] are then satisfied if the local stability condition (7) holds, and so $X(t)$ converges geometrically fast towards f for any given initial state $X(0) \in \{f > 0\}$ (see [H, Theorem 2.1 and Corollary 3.1]). Moreover, $X(t)$ is time-reversible and f specifies the unique invariant distribution of $X(t)$.

In [S] spatial birth-and-death processes with another construction of the birth and death rates are considered. Briefly, a flexible and statistically tractable parametric class of spatial birth-and-death processes, defined on the real line, is presented and related to modelling linear dune fields (with data in the form of aerial photographs from the Kalahari Desert). Viewed in the downwind direction of the prevailing wind (the 'time axis'), new dunes start to grow and some of the old dunes will tend to end due to the limited sand supply. The birth and death rates depend only on nearest left and right neighbour interactions, whereby, in the case of detailed balance, the equilibrium distribution becomes an NNMPP with respect to the sequential neighbour relation mentioned in Section 2.2. We return to this model in Section 3.5.

2.4 Spatial heterogeneity and log Gaussian Cox processes

In Section 2.2 we discussed clustering in a spatial point pattern caused by spatial interaction. Another possible source for clustering is spatial heterogeneity. The class of Cox processes provides flexible and useful models for aggregated point patterns where the aggregation is due to a stochastic environmental heterogeneity [43, 45, 46, 180]. In this section we focus mostly on a particular subclass of Cox processes, viz. the log Gaussian Cox processes introduced in [T].

Recall first the definition of a Cox process X on $\mathbf{R}^d$ directed by a random intensity process $\Lambda = \{\Lambda(s) : s \in \mathbf{R}^d\}$ which satisfies the local finiteness condition, $\int_B \Lambda(s)\mathrm{d}s < \infty$ for all $B \in \mathcal{B}$ (assuming, of course that the integral exists; recall that $\mathcal{B}$ denotes the class of bounded Borel sets in $\mathbf{R}^d$). Then X conditional on Λ is a Poisson process with intensity measure

$\nu(B) = \int_B \Lambda(s)\mathrm{d}s$, that is, for any $B \in \mathcal{B}$, $n(X \cap B)|\Lambda$ follows a Poisson distribution with mean $\nu(B)$, and conditionally on $n(X \cap B) = n$, if $\nu(B) > 0$, the n points in $X \cap B$ are i.i.d. with distribution $\nu(\cdot \cap B)/\nu(B)$.

If X is a Cox process with

$$\Lambda(s) = \exp(Y(s))$$

where $Y = \{Y(s) : s \in \mathbf{R}^d\}$ is a (continuous version of a) real-valued Gaussian process, then X is said to be a *log Gaussian Cox process* (LGCP). It is necessary to impose weak conditions on the mean function $m(s) = EY(s)$ and covariance function $c(s,t) = Cov(Y(s), Y(t))$ in order to get a well-defined LGCP as explained in [T, Section 2]. Examples of covariance functions together with simulated realisations of LGCPs and their underlying Gaussian processes are shown in [T, Section 2].

LGCPs possess various appealing properties as indicated below.

Theoretical results are easily derived. Letting $\rho^{(n)}$, $n = 1, 2, \ldots$, denote the nth order product densities for the reduced moment measures of an LGCP X, then since X is a Cox process, $\rho^{(n)}(s_1, \ldots, s_n) = E\prod_1^n \Lambda(s_i)$, $s_1, \ldots, s_n \in \mathbf{R}^d$. Moreover, as X is an LGCP,

$$\rho^{(n)}(s_1, \ldots, s_n) = \prod_{i=1}^{n} \rho(s) \prod_{1 \le i < j \le n} g(s_i, s_j) \tag{15}$$

where

$$\rho(s) = \rho^{(1)}(s) = \exp(m(s) + c(s,s)/2) \tag{16}$$

and

$$g(s,t) = \rho^{(2)}(s,t)/(\rho(s)\rho(t)) = \exp(c(s,t)) \tag{17}$$

are the intensity and the pair correlation functions of X, respectively. Hence there is a one-to-one correspondence between c and g and between (m, c) and (ρ, g). Consequently, the distribution of an LGCP is uniquely determined by its first- and second-order properties as given by the intensity and pair correlation functions. This makes parametric models easy to interpret and simple methods for parameter estimation and model checking (including third-order properties) become available as discussed in Section 3.6. Further theoretical results for LGCPs can be found in [T, Section 3].

Another very useful property of LGCPs is that the underlying Gaussian process and the intensity surface can easily be predicted by empirical Bayesian methods combined with efficient MCMC methods (see Section 3.6). Moreover, there is no problem with edge effects as the distribution of an LGCP restricted to a bounded subset is known.

Besides these properties which are rather characteristic for LGCPs, LGCPs are flexible models for clustering, easy to simulate, and the definition of an LGCP can easily be extended in a natural way to multivariate LGCPs as shown in [T]. Further, in [T], LGCPs are compared with Neyman-Scott processes (when the number of offspring is Poisson distributed), a popular and frequently used class of models which are simultaneously Poisson cluster and Cox processes. It is concluded that the product densities (15) of an LGCP are in general of a different and much simpler form than for Neyman-Scott processes. Moreover, it is observed that certain Thomas processes (a particular but widely used type of Neyman-Scott processes) may in practice be difficult to distinguish from LGCPs with a Gaussian covariance function.

Throughout [T] stationarity and often also isotropy is assumed — this may be a reasonable assumption for smaller point patterns as traditionally considered in the literature where the number of points is usually at most a few hundred, but today there is a need for analysing larger data sets. As indicated above the study of LGCPs can be extended to non-stationary and non-isotropic situations as well, and this aspect should be further explored, see Section 3.6. In general, at a physically large scale spatial heterogeneity may very well be the dominating factor, whereas small scale properties may mainly be explained by the interactions between the points. This is further studied in our recent joint research project [9], but primarily for another type of model obtained by thinning from a homogeneous Markov point process.

Finally, let me mention how LGCPs can naturally be extended to a space-time model. Conditional on a Gaussian process Y, consider a Poisson birth process $\{X(t) : t \geq 0\}$ with $X(0) = 0$ and birth rate $\exp(Y(s))b(t)$, where b is a non-negative integrable function. Then $X(t)$ is an LGCP driven by the Gaussian process $\{Y(s) + \log(\int_0^t b(u)\mathrm{d}u) : s \in \mathbf{R}^d\}$, and so by (17) the pair correlation function of $X(t)$ does not depend on t. Suppose now that $X(1), \dots, X(m)$ are observed (within some bounded window). Considered backwards in time, this is a Markov chain where $X(t-1)$ is obtained from $X(t)$ by independent thinning, i.e. when the points in $X(t)$ are deleted independently of each other with probability $p(t-1) = \int_{t-1}^t b(u)\mathrm{d}u / \int_0^t b(u)\mathrm{d}u,\ t = m, \dots, 1$. Since the pair correlation functions are identical for all times $t > 0$, we may start by analysing the largest observed point pattern $X(m)$ and next investigate the deletion probabilities $p(m-1), \dots, p(1)$. In the recent work [29], we are further extending this approach to multitype space-time LGCPs in a study of weeds.

3 Markov chain Monte Carlo methods and statistical inference for lattice and point processes

In Section 3.1, Metropolis-Hastings algorithms [132, 81] adapted for spatial processes are reviewed and compared with the competing approach of running a spatial birth-and-death process. The very recent ideas on how to make perfect (exact) simulation are discussed in Section 3.2. Section 3.3 deals with the celebrated Swendsen-Wang algorithm and other types of auxiliary variable algorithms. Then we turn to statistical analysis based on either the pseudo-likelihood function (Section 3.4) or the likelihood function (Section 3.5), where in the latter case MCMC methods play a prominent role. Finally, in Section 3.6 we discuss ways of doing statistical inference for log Gaussian Cox processes.

3.1 Metropolis-Hastings simulation

Lattice processes and spatial point processes with a fixed number of points are nothing but multivariate distributions so that common Monte Carlo algorithms such as the Gibbs sampler and the Metropolis algorithm may apply; in the case of lattice processes, see [179, 23, 17]; for point processes with a fixed number of points, see the review in [P, Section 3.1]. In the sequel I shall therefore concentrate on Metropolis-Hastings simulation for finite point processes with a random number of points. As in Section 2.2, we consider for simplicity the case of a hereditary density f for a finite simple point process defined on a Lebesgue set $S \subset \mathbf{R}^d$ with Lebesgue measure $0 < \lambda(S) < \infty$, though everything in the following is easily generalized to a more general setting as shown in the references mentioned below.

The traditional approach for simulating from the density f is to generate a sample from a spatial birth-and-death process $X(t), t \geq 0$, which converges towards f (see [162, 105, 168] and [A]). In Section 2.3 we considered the most common situation where existing points are deleted with rate 1, and if x is the current state of $X(t)$, a newborn point ξ is added with rate $B(x) = \int_S \lambda^*(x, \eta)\mathrm{d}\eta$, where ξ has density $\lambda^*(x, \cdot)/B(x)$. Also, assuming local stability, we noted that $X(t)$ is well-defined and converges geometrically fast to f. One immediate problem is the calculation of the integral $B(x)$ which in general requires numerical methods. However, local stability may be further used for obtaining the spatial birth-and-death process $X(t)$ by thinning from a dominating spatial birth-and-death process $D(t)$. For the dominating process we use again death rates 1 for existing points, but newborn points are uniformly distributed and the birth rate is constant $K\lambda(S)$, where K is the upper bound in the local stability condition (7). The dominating process is time-reversible and it converges geometrically fast towards its equilibrium distribution, which is a Poisson process on S with rate K. Thus it is straightforward to simulate $D(t)$ starting in equilibrium at time 0. From $D(t)$ we easily get $X(t)$ as follows. Suppose that $X(t-) = x$, where $t-$ means 'just before time t'. If $D(t) = D(t-) \cup \{\xi\}$, then $X(t) = x \cup \{\xi\}$ with probability $\lambda^*(x, \xi)/K$ and we set $X(t) = x$ otherwise. If $D(t) = D(t-) \setminus \{\eta\}$, then $X(t) = X(t-) \setminus \{\eta\}$. If $D(t) = D(t-)$, we do nothing.

An even simpler simulation procedure for point processes is based on the *Metropolis-Hastings algorithm* studied in [C, P] and [67]. If $x \in \{f > 0\}$ is the current state of the Markov chain generated by the algorithm, then we propose with probability $1/2$, (i) to add a uniformly distributed point $\xi \in S$, and else we attempt (ii) to delete a uniformly selected point η from x (if $x = \emptyset$ we do nothing in (ii)). In the case (i), with probability $\alpha_b(x,\xi) = \min\{1, r(x,\xi)\}$ we accept $x \cup \{\xi\}$ as the next state of the Markov chain, and otherwise we retain x; here

$$r(x,\xi) = \lambda^*(x,\xi)\lambda(S)/(n(x)+1)$$

is the 'Metropolis-Hastings ratio'. In the case (ii), with probability $\alpha_d(x,\eta) = \min\{1, 1/r(x \setminus \{\eta\}, \eta)^{-1}\}$ we accept $x \setminus \{\eta\}$ as the next state of the Markov chain, and otherwise we retain x.

This Metropolis-Hastings algorithm can be considered as a special case of Green's [74] reversible jump algorithm as explained in [67] ([C] appeared earlier than [74]). However, it is simpler to establish reversibility directly without referring to Green's paper due to the following detailed balance condition which is immediately seen to hold:

$$f(x \cup \{\xi\})\frac{1}{2}(n(x)+1)\frac{1}{n(x)+1}\alpha_d(x,\xi)\mathrm{d}\mu(x \cup \{\xi\})$$

$$= f(x)\frac{1}{2}\frac{\mathrm{d}\xi}{\lambda(S)}\alpha_b(x,\xi)\mathrm{d}\mu(x)$$

where

$$\mathrm{d}\mu(x) = \exp(-\lambda(S))/n!\ \mathrm{d}x_1 \cdots \mathrm{d}x_n \quad \text{for} \quad x = \{x_1, \dots, x_n\},\ n \geq 1$$

and $\mathrm{d}\mu(\emptyset) = \exp(-\lambda(S))$. In fact, an entire class of Metropolis-Hastings algorithms (including moves of existing points) is introduced in [C], but the particular algorithm described above seems simplest to analyse and implement; incidentally, for simulating the Strauss process we found in [C] that it was not useful to combine the algorithm with a Metropolis algorithm for moves of points.

Further desirable properties of the algorithm can be established assuming local stability: it is aperiodic, Harris recurrent, and satisfies a geometrical drift condition so that the Markov chain $\{X_t\}$ generated by the algorithm is geometrically ergodic in the sense of [133, Theorem 15.0.1] where f specifies the unique invariant distribution. This was first verified in [C] for the Strauss process and later extended in [67] to locally stable point processes along the same lines as in [C] but now combined with general concepts and results from [133]. The fact that the empty point configuration is an ergodic atom is utilized in the proofs in [C] and [67].

Geometrical ergodicity provides qualitative bounds on the rate of convergence. Quantitative bounds for the convergence rate are discussed in [P, Appendix B], where it is concluded that the obtained bounds (using the results which are so far available) seem too weak for real applications of our Metropolis-Hastings algorithm.

But at least as important, geometrical ergodicity implies a Central Limit Theorem for the Monte Carlo approximation $\bar{g}_n = \sum_{t=1}^{n} g(X_t)/n$ of the mean $Eg = \int g(x)f(x)\mathrm{d}\mu(x)$ for any real function g defined on Ω and with finite second-order moment, $Eg^2 < \infty$ (see [P, Appendix A] and [173, Corollary 3]). The variance in this CLT can be estimated by time series methods.

The two approaches of using either a spatial birth-and-death process or the Metropolis-Hastings algorithm are compared in [C, P]. The Metropolis-Hastings algorithm involves fewer calculations and is easier to program, but whether it runs faster depends of course on the implementation (see [37]). Further, its theoretical properties are easier to analyse. Also exploratory analysis for estimating the 'burn-in' (i.e., the number of transitions needed to be thrown away before collecting samples) and to detect badly mixing chains have to the best of my knowledge only been studied in the literature in connection to MCMC algorithms (see [41, 30], [C], [P, Appendix A], and the references therein). Furthermore, for simulation procedures based on running spatial birth-and-death processes, I'm not aware of anybody who has so far established the existence of a CLT for the Monte Carlo error and shown how to estimate the asymptotic variance in practice — of course apart from the fact that a spatial birth-and-death process sampled at equidistant discrete times forms a Markov chain. As discussed in [109] and Section 3.2 below, perfect simulation procedures based on each approach — and which make i.i.d. simulations feasible so that it is straightforward to establish a CLT for the Monte Carlo error and estimate the asymptotic variance — have recently been developed, but at the time of writing this thesis a comparative study of the two approaches remains to be done.

3.2 Perfect simulation

As mentioned in Section 3.1, a common problem in MCMC is that the available theoretical bounds for rates of convergence are not useful in practice. Therefore, it is highly remarkable that Propp and Wilson [163] recently devised a technique based on *coupling from the past* (CFTP) which delivers *perfect* (*or exact*) *simulation* from, for example, the Ising model and which works in practice even for fairly large systems and in the phase transition regime (see [163, 164]).

Briefly, Propp and Wilson assume the state space Ω to be finite and equipped with a partial ordering $\preceq$ so that Ω has a unique minimal element $\mathbf{O}$ and a unique maximal element $\mathbf{1}$, and the sampler (e.g. the Gibbs sampler) for simulating from a given target distribution π is monotone. It is convenient to express this in terms of a stochastic recursive sequence [24, 25, 56]: in general any Markov chain $\{X_t\}$ can be presented as

$$X_t = \phi(U_t, X_{t-1})$$

where ϕ is a deterministic measurable function and the U_t are i.i.d. In fact there are infinitely many such constructions, but in practice the 'updating rule' ϕ is often obvious and given by the way one writes the computer program for simulating π as the limit of the chain. Monotonicity means that $\phi(u,x) \preceq \phi(u,y)$ whenever $x \preceq y$. In other words, if we consider two paths generated by the updating rule, using the same 'random number sequence' and started in the states $x \preceq y$, respectively, then with respect to $\preceq$ the path started in x remains beneath the path started in y.

Now, in the Propp-Wilson algorithm, one uses repeatedly the updating rule to generate lower and upper Markov chains started further and further back in time < 0 in the states $\mathbf{O}$ and $\mathbf{1}$, respectively. By reusing a common 'random number sequence' (this is the CFTP), this is repeated until a pair of upper and lower chains coalesce at time 0, and then we return the result as a simulation from the target distribution. The monotonicity condition for the sampler ensures that any other chain started at time $t < 0$ and in an arbitrary state (reusing the 'random number sequence') sandwiches between any pair of lower and upper chains started at time $\geq t$. Under weak ergodicity conditions for the sampler, coalescence happens within finite time, and by considering a 'virtual simulation from time minus infinity', one can show that the output exactly follows the target distribution.

In the last three years Propp and Wilson ideas have been extended in various ways, particularly within spatial statistics as discussed below.

One idea is to combine CFTP with the thinning technique for generating a spatial birth-and-death from a dominating one (now generated backwards in time and started in equilibrium at time 0) as outlined in Section 3.1. This was first described in [108] for the case of the area-interaction process; we are giving the details for the general case of simulating locally stable point processes in [109].

Another idea is to use *perfect Metropolis-Hastings sampling*. This was first described in [D] for the case of the Widom-Rowlinson [197] mixture model as discussed later on in this section. A perfect Metropolis-Hastings simulation algorithm for locally stable point processes in general, based on CFTP and a special case of the general Metropolis-Hastings algorithms in [C] (not the one considered in detail in Section 3.1), is constructed in [109]. As it is somewhat more technical to present than that based on spatial birth-and-death processes, a detailed description will not be given in this thesis. One very interesting observation concerning the Metropolis-Hastings algorithm in [109] is the fact that it is geometrically but not uniformly ergodic (unless there is an upper bound on the number of points); this should be contrasted to a recent result in [57, Theorem 4.2] which states that successful CFTP algorithms such as those of Propp and Wilson can be constructed if and only if the chain is uniformly geometrically ergodic.

Concerning perfect simulation of lattice processes, while Propp and Wilson in their examples require the target distribution to be attractive, it is shown in [80] how to make perfect simulation for the repulsive Ising process (1) with $\beta < 0$ and other 'anti-monotone' systems defined on a finite state space; here the CFTP is based on an idea in [108]. This is further extended in [N] to models with discrete and continuous state spaces and where some kind of attraction or repulsion is present, including locally specified exponential family models with applications in spatial statistics such as auto-binomial, auto-Poisson, auto-negative-binomial, auto-gamma models and certain pairwise-difference interaction models (see [19, 43, 75]). Briefly, the perfect simulation algorithms in [N] are based on Gibbs sampling, and the lower and upper chains are obtained by 'scaling' from a dominating Markov chain — indeed the techniques used in [N] are much inspired by our approach in [109]. Especially, for an auto-gamma model, as we are using Gibbs sampling on continuous distributions we never get coalescence of the lower and upper paths; however with an arbitrary good accuracy $\epsilon > 0$ we can deliver a perfect simulation within finite time (see [N, Section 3]). It is at least theoretically feasible to make simulations from an auto-gamma model with accuracy $\epsilon = 0$ by the perfect Metropolis-Hastings algorithm studied in [N, Section 3.3], but in practice this algorithm becomes extremely slow. The auto-gamma model has also been used in other papers on MCMC methods, in particular in connection to a Bayesian analysis of a dataset on pump reliability [60, 148]. The CFTP algorithms in [148] do not rely on any monotonicity property, but a larger number of paths are used in [148] than just the two paths for the lower and upper paths used in [N]. Possibly, a hybrid algorithm would be even better: first use my algorithm so that lower and upper paths become close enough and then use the multigamma coupler in [148].

Let us consider in more detail how we make perfect simulation of the *Widom-Rowlinson mixture model* in [D]. This is a two type point process (X, Y) defined on S and with density

$$f(x, y) \propto \beta_1^{n(x)} \beta_2^{n(y)} \mathbf{1}[d(x, y) > R] \tag{18}$$

with respect to $\mu \times \mu$, where β_1, β_2, R are positive parameters and $d(x, y)$ is the shortest distance between a point in x and a point in y. In other words, the distribution of (X, Y) is obtained by conditioning in two independent Poisson processes W, Z defined on S and with rates β_1, β_2, respectively, so that $d(W, Z) > R$. We can therefore easily sample from the conditional distributions $X|Y$ and $Y|X$ by thinning from (W,Z), whereby we get a two-component Gibbs sampler with updating rule $(x_{new}, y_{new}) = \phi((w,z),(x,y))$, where $x_{new} = w \backslash U_y$, $y_{new} = z \backslash U_{x_{new}}$, and U_x is defined as in the area-interaction process (10). This updating rule is monotone with respect to the partial ordering given by $(x,y) \preceq (\acute{x},\acute{y}) \Leftrightarrow x \subseteq \acute{x},\ y \supseteq \acute{y}$.

Now, the perfect simulation algorithm in [D] consists in thinning from independent Poisson processes $W_0, Z_0, W_{-1}, Z_{-1}, \ldots$ (playing the role of the 'random number sequence') with $(W_t, Z_t) \sim (W,Z)$: Let $0 > t_1 > t_2 > \ldots$ be a given sequence of integers; the nearly optimal choice is $t_i = -2^i$, c.f. [163]. Then for $i = 1, 2, \ldots$ we repeat to generate a lower chain $\{(X_t^{(i)}, Y_t^{(i)}) : t = t_i, \ldots, 0\}$ and an upper chain $\{(\acute{X}_t^{(i)}, \acute{Y}_t^{(i)}) : t = t_i, \ldots, 0\}$ by setting first

$$(X_{t_i}^{(i)}, Y_{t_i}^{(i)}) = (\emptyset, Z_{t_i}), \quad (\acute{X}_{t_i}^{(i)}, \acute{Y}_{t_i}^{(i)}) = (W_{t_i}, \emptyset),$$

and next for $t = t_i + 1, \ldots, 0$,

$$(X_t^{(i)}, Y_t^{(i)}) = (W_t \setminus U_{Y_{t-1}^{(i)}}, Z_t \setminus U_{X_t^{(i)}}), \quad (\acute{X}_t^{(i)}, \acute{Y}_t^{(i)}) = (W_t \setminus U_{\acute{Y}_{t-1}^{(i)}}, Z_t \setminus U_{\acute{X}_t^{(i)}}),$$

until $(X_0^{(i)}, Y_0^{(i)}) = (\acute{X}_0^{(i)}, \acute{Y}_0^{(i)})$. Letting

$$I = \inf\{i : (X_0^{(i)}, Y_0^{(i)}) = (\acute{X}_0^{(i)}, \acute{Y}_0^{(i)})\}$$

then as shown in [D], $I < \infty$ almost surely and the output $(X_0, Y_0) = (X_0^{(I)}, Y_0^{(I)})$ follows (18). Notice that in contrast to the situation in Propp and Wilson's work, the state space is neither finite nor countable and there are no maximum or minimum with respect to $\preceq$ (at least not within the state space — instead we start at quasi minimal and quasi maximal elements as described in [D, Section 5]).

As discussed in [D, Section 4], the marginal distribution of X under the Widom-Rowlinson model (18) is an attractive area-interaction process with parameters $(\beta, \gamma) = (\beta_1, \exp(\beta_2))$ (a similar result holds, of course, for the distribution of Y) and in the symmetric case $\beta_1 = \beta_2 = \beta$, we have that $X \sim Y$ and the superposition $X \cup Y$ follows a continuum random

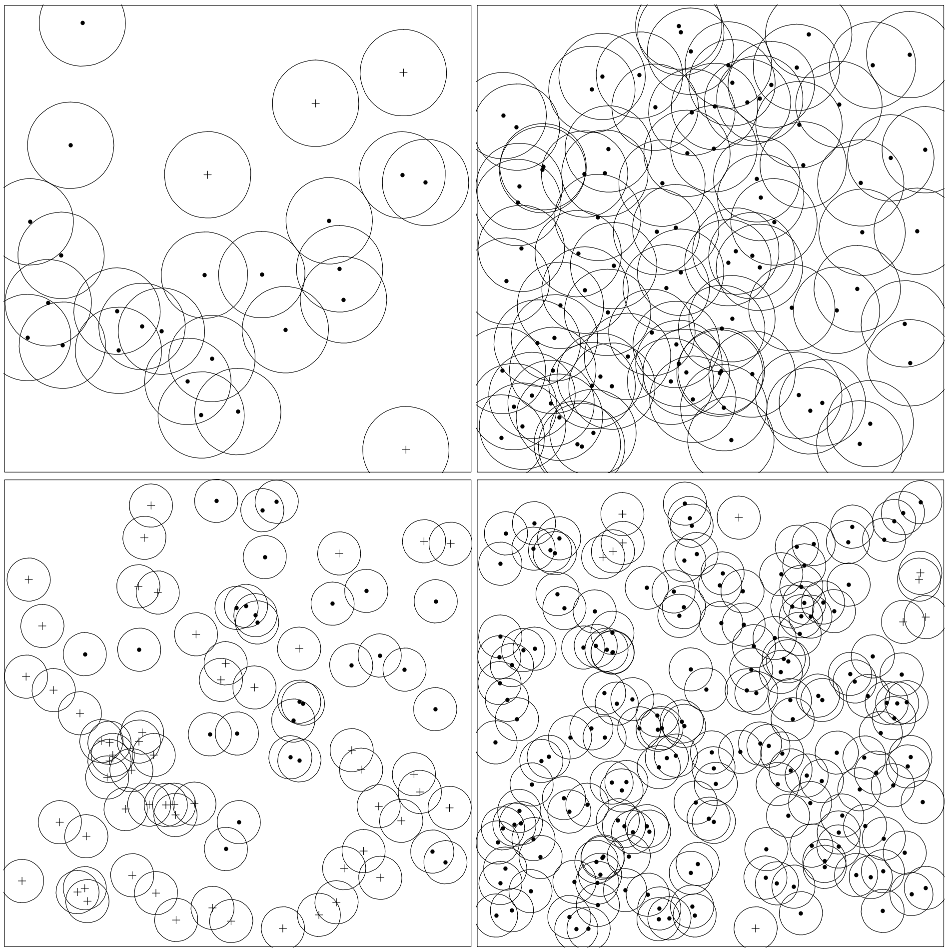

Figure 2: Realisations of the Widom-Rowlinson model using perfect simulation with $\beta = 30, R = .2$ (top left), $\beta = 80, R = .2$ (top right), $\beta = 100, R = .1$ (bottom left) and $\beta = 200, R = .1$ (bottom right).

cluster model with $\gamma = 1/2$ in (12). In the symmetric case there is a phase transition (see Section 2.2) when β is sufficiently large and no phase transition when β is sufficiently small (see [177, 33]), but it is yet an open problem to verify if there exists a so-called critical value of β. The simulated results in [D] (for the symmetric case and $d = 2$) indicate that it exists. The effect of phase transition is visible in Figure 2 (taken from [D, Figure 2]) which shows realisations of the symmetric Widom-Rowlinson model within a unit square and for different values of β and R (decreasing R corresponds to increasing the square window).

For extremely large values of β or $\lambda(S)$ there is of course the possibility that the perfect simulation algorithm in [D] fails to stop within practical constraints of time. As pointed out in [53], when using CFTP in general, the impatient user may then introduce a bias in the result if the running time (in our case I) and the output (in our case (X_0, Y_0)) are dependent. Recently, it has been shown in [192] how Fill's [53] rejection sampling algorithm can be extended from the finite state space case to the Widom-Rowlinson model (18) when using the two-component Gibbs sampler in [D]; then exact and unbiased results are produced. It is concluded in [192] that Fill's algorithm is approximately as efficient as ours in [D], but it needs more memory. Moreover, in [143] we show how Fill's algorithm for attractive finite state space models can be extended to general discrete conditionally specified repulsive models using sandwiching techniques as in [N].

Finally, a few remarks. So far perfect simulation algorithms in the context of spatial statistics, stochastic geometry, and statistical physics are only practicable for certain relative simple models. However, the possibility of combining perfect simulation with efficient simulation procedures such as simulated tempering (see Section 3.3) is explored in our recent work [142]. Another approach is perfect slice sampling as investigated in our ongoing research work [139]. One obvious advantage of using perfect simulation is the fact that there is no problem with the 'burn-in'. Further, i.i.d. sampling is now available so that e.g. asymptotic variances of Monte Carlo estimates can be calculated very directly. Moreover, we may assess the approximation error incurred by a 'non-perfect' algorithm via a comparison with a perfect variant of the algorithm. These aspects should be investigated in the future, in particular in connection to Bayesian and approximate likelihood inference.

3.3 Auxiliary variable algorithms

When simulating spatial processes with strong interactions, MCMC algorithms can fail to converge in practice. For example, in the Ising model (1) with β close or above the critical value, conventional local updating procedures such as the Gibbs sampler and the Metropolis algorithm mostly get stuck in a state with one dominating connected component (see Section 2.1). This problem was drastically reduced with a collective updating algorithm due to Swendsen and Wang [189], which relies on the fact that the Ising model (IM) and the random cluster model (RCM) of $0-1$ bonds can be defined simultaneously. In the *Swendsen-Wang algorithm* systematic Gibbs sampling is used by alternating between the conditional distributions $IM|RCM$ and $RCM|IM$. Simply, $IM|RCM$ is a product of uniform distributions on $\{0,1\}$ over the maximal connected components in RCM, and the bonds in RCM are conditionally independent given IM. Propp and Wilson [163] remark that there seems to be no useful monotonicity property in the Swendsen-Wang algorithm; in fact, in their perfect simulation algorithm they instead first simulate a realisation of the RCM using Gibbs sampling (which is monotone for RCM) and secondly they use this realisation to generate a realisation from $IM|RCM$. In either case we can consider the random cluster process as an auxiliary process used for simulating the Ising model.

Various extensions of the Swendsen-Wang algorithm have been suggested, most of them for applications in statistical mechanics (e.g., [50, 23, 54] and the references therein). A Swendsen-Wang type algorithm for the symmetric Widom-Rowlinson model (18) with $\beta_1 = \beta_2 = \beta$ is proposed in [D] — it was discovered independently in [34] — and consists of two steps: (i) if (x, y) is the current state of the chain, then we generate a new value $(\acute{x}, \acute{y})$ with $\acute{x} \cup \acute{y} = x \cup y$ by independently for each maximal connected component $z \subseteq x \cup y$ determining according to the uniform distribution on $\{0,1\}$ whether all the points in z should be of the first or second type; (ii) replace $\acute{y}$ by $\tilde{y}$ generated from a Poisson process of rate β on $S \setminus U_{\acute{x}}$; then $(\acute{x}, \tilde{y})$ is the next state of the chain. As in the original Swendsen-Wang algorithm, it seems difficult to find any useful monotonicity property in our Swendsen-Wang type algorithm. So it would be interesting to see if it is practicable first to use one of our perfect simulation algorithms in [109] for simulating the continuum random cluster model (12) with $\gamma = 1/2$ and secondly to make a simulation of the Widom-Rowlinson model by step (i) in the Swendsen-Wang type algorithm.

Another extension of the Swendsen-Wang algorithm is product slice sampling; in [139] we study perfect product slice sampling for Ising and other spatial auto-models.

Swendsen-Wang type algorithms seem only to be effective if a special structure is exploited (see [50, 23, 54, 139]). A more general auxiliary variable algorithm is *simulated tempering* [118, 69]. Suppose that $f_1(x) \propto h_1(x)$ is the target density (the 'cold' distribution) and we have some other densities $f_i(x) \propto h_i(x)$, $i = 2, \dots, m$ which are 'heated' versions of f_1 and defined on the same state space Ω as f_1; for the 'unnormalized densities', we may take $h_i(x) = h_1(x)^{T_i}$ with $T_1 = 1 > T_2 > \dots > T_m \geq -\infty$, but other forms of heating may be considered as well. Assume that for each i we have an MCMC algorithm MC_i with invariant density f_i and we have chosen a 'pseudo-prior' on the m distributions by specifying m positive numbers $\pi(1), \dots, \pi(m)$. Then in the simulated tempering algorithm we alternate between updating states $x \in \Omega$ and distributions $i \in \{1, \dots, m\}$. If (x, i) is the current state, then we (i) update x by MC_i to a state y, (ii) set $j = i \pm 1$ according to probabilities $q_{i,j}$, where $q_{1,2} = q_{m,m-1} = 1$ and $q_{i,i+1} = q_{i,i-1} = 1/2$ if $1 < i < m$, and (iii) with probability $\min\{1, (h_j(y)\pi(j)q_{j,i})/(h_i(y)\pi(i)q_{i,j})\}$, let (y, j) be the next state; (y, i) is the next state otherwise. Thereby a Markov chain (X_t, I_t) with invariant density $f(x, i) \propto h_i(x)\pi(i)$ is generated and, with obvious notation, $f(x|i) = f_i(x)$ so that (in the ergodic case and after an appropriate burn-in) we can use the states X_t with $I_t = 1$ as a sample from the target density. It is discussed in [69] how to specify in practice the number of distributions and the pseudo-prior. In particular, if MC_m for the 'hot' distribution is an independence sampler, then we have regeneration, i.e.

sample paths between the times t where $I_t = m$ are i.i.d. This simplifies estimation of Monte Carlo errors and has been used in [69] for simulating the attractive Strauss process (see Section 2.2) with the 'hot' distribution being a binomial process.

Simulated tempering can be applied for simulating other spatial models, such as hard-core point processes with a high packing degree (see [124]), and in missing data problems of marked point processes of geometrical objects, which require conditional simulations on a 'narrow' part of the state space; the latter aspect is currently under study in [28].

3.4 Pseudo-likelihood

The normalizing constant in the density of many parametric spatial models (including those considered in Section 2) cannot be evaluated explicitly. In order to avoid this problem, Besag [20] extended the 'coding' technique for lattice processes in [19] to obtain a pseudo-likelihood function. As in Section 2.1, consider for simplicity a probability mass function p_θ for a discrete model with unknown parameter $\theta \in \Theta$ to be estimated from data $x = (x_i)_{i \in I}$, where I is a finite index set. Letting $J \subseteq I$, then the pseudo-likelihood function on J is given by

$$PL_J(\theta; x) = \prod_{j \in J} p_\theta(x_j | x_i, i \in I \setminus \{j\}). \tag{19}$$

For instance, if $p_\theta(\cdot)$ specifies an MRF on $I \subset \mathbf{Z}^2$, 'edge effects' may be eliminated by taking, with obvious notation, $J = I \setminus \partial(I^c)$ (the 'border method'). The *maximum pseudo-likelihood estimate* (MPLE) is found by maximizing (19). In the special case of an auto-logistic model (including the Ising process), finding the MPLE is computationally equivalent to finding the maximum likelihood estimate (MLE) in a logistic-regression model, so a standard statistical computer package can be used for the computations (see [160, 185]). Consistency and asymptotic normality of the MPLE for lattice processes (in various settings) have been considered in [62, 78, 79, 70, 38, 39].

Besag [16] extended the definition of the pseudo-likelihood function to the Strauss process (9) by combining (19) with an approximation of the Strauss process given by an auto-Poisson lattice process (see [18]). Based on this derivation a general expression of the pseudo-likelihood for point processes is stated in [171].

The *pseudo-likelihood for point processes* (in a general setting including marked point processes) is derived by a direct argument in [F, Section 2]. Using a notation as in Section 2.2, suppose that for each $\theta \in \Theta$, f_θ is a hereditary and Ruelle stable density for a point process on S (e.g. a bounded Borel set) with respect to a Poisson process μ_S with a finite diffuse intensity measure λ (e.g. Lebesgue measure). For measurable $T \subseteq S$ and finite point configurations $x \in \Omega_S$ we define the pseudo-likelihood on T by

$$PL_T(\theta; x) = \exp(-\lambda(T)) \lim_{i \to \infty} \prod_{j=1}^{m_i} f_\theta(x_{A_{ij}} | x_{S \setminus A_{ij}}).$$

Here $\{A_{ij} : j = 1, \dots, m_i\}$, $i = 1, 2, \dots$, are nested subdivisions of T such that $m_i \to \infty$ and $m_i [\max_{1 \le j \le m_i} \lambda(A_{ij})]^2 \to 0$ as $i \to \infty$. Further,

$$f_\theta(x_{A_{ij}} | x_{S \setminus A_{ij}}) = \frac{f_\theta(x_{A_{ij}} \cup x_{S \setminus A_{ij}})}{\int f_\theta(y \cup x_{S \setminus A_{ij}}) \mathrm{d}\mu_{A_{ij}}(y)} \quad \text{if} \quad f_\theta(x_{S \setminus A_{ij}}) > 0$$

and 0 otherwise. By [F, Theorem 2.2], for μ_S almost all $x \in \Omega_S$ the pseudo-likelihood on T is well defined and given by

$$PL_T(\theta; x) = \exp\Big(- \int_T \lambda_\theta^*(x, \xi) \mathrm{d}\lambda(\xi) \Big) \prod_{\xi \in x_T} \lambda_\theta^*(x \setminus \{\xi\}, \xi) \qquad (20)$$

where λ_θ^* is the Papangeleou conditional intensity associated to f_θ (the pseudo-likelihood is unaltered whether λ_θ^* is based on f_θ or on the conditional density $f_\theta(x_T | x_{S \setminus T})$). Hence, if $\Theta \subseteq \mathbf{R}^k$ is open, the pseudo-likelihood equation on T becomes

$$\int_T \frac{\mathrm{d}}{\mathrm{d}\theta} \lambda_\theta^*(x, \xi) \mathrm{d}\lambda(\xi) = \sum_{\xi \in x_T} \frac{\mathrm{d}}{\mathrm{d}\theta} \log \lambda_\theta^*(x \setminus \{\xi\}, \xi) \qquad (21)$$

provided we can interchange the integration and the differentiation.

This also can be obtained as a moment type estimation equation using the Nguyen-Zessin [149] integral characterisation of a Gibbs point process X defined on $\mathbf{R}^d$,

$$E \sum_{\xi \in X} h(X \setminus \{\xi\}, \xi) = E \int \lambda_\theta^*(X, \xi) h(X, \xi) \mathrm{d}\xi \tag{22}$$

for measurable functions h so that the mean values exist (here the definition of λ^* is obvious if we consider a Markov process of finite range). Setting $h(X,\xi) = \mathbf{1}[\xi \in T] \frac{\mathrm{d}}{\mathrm{d}\theta} \log \lambda_\theta^*(X,\xi)$ and omitting the expectation signs in (22) we get (21) (in the case of marked point processes, see [73]). Using other 'test functions' h, (22) was used in [52, 190] to obtain other estimation equations (the so-called Takacs-Fiksel method).

Maximum pseudo-likelihood estimation based on (20) is computationally equivalent to maximum likelihood estimation in an inhomogeneous Poisson process. In practice the integral in (20) may be approximated by numerical methods. As noticed in [13, 7] the GLIM-package can be used to provide an approximate MPLE if λ^* is of a certain form.

If the density f_θ belongs to an exponential family, then $PL_T(\theta; x)$ is log concave (see [F, Proposition 2.3]). When, in addition, the density is Markov of finite range and essentially the density is either repulsive or a hard-core condition is satisfied, we use in [F] a coding idea to establish consistency of the MPLE; the proof is independent of existence and uniqueness of an underlying infinite volume Gibbs process. Strong consistency has later been established in [123] for planar point processes with pairwise-only interactions and satisfying certain stability conditions so that an infinite volume Gibbs process exists. Asymptotic normality of the MPLE is discussed in [78, 99, 100, 119].

In general, for lattice and spatial point processes with high dependence, the MPLE is inefficient. Intuitively, as the pseudo-likelihood only depends on the local dependence structure, global information may better be taken into account when using the likelihood function than the pseudo-likelihood function. For a conditional auto-regression model as considered in [104], the asymptotic variance of the MPLE can be nearly twice as large as that for the MLE. Similarly for the Ising model, where in the two-dimensional case the MLE 'senses' the critical point in a way that the MPLE does not (see [65, 68]); in the one-dimensional case the MPLE is not optimal within a natural class of estimating functions considered in [97]. Optimality of the MLE and inefficiency of the MPLE is established in [122] for a certain class of pairwise soft-core interaction point processes. Improvements of the maximum pseudo-likelihood method are discussed in [93].

3.5 Likelihood analysis based on MCMC methods

As the likelihood function is unavailable for most parametric spatial models, different approximate methods for finding the MLE have been proposed. One approach is to approximate the normalizing constant of the density based on Taylor expansions [182] or cluster expansions from statistical physics [151, 152, 153, 154, 183] or sparse-data approximations [156]. Another approach used in [156, 68, 66, 67] and [C] is based on MCMC approximations of the log-likelihood function or its derivatives. Yet another approach is iterative stochastic approximation of the MLE based on stochastic gradient algorithms [199, 145]. For reviews on many of these methods and their history, see [171, 68, 48, 67].

In the sequel we concentrate on the approach of estimating the likelihood function by MCMC methods as proposed by Geyer and Thompson [68]. Apart from seeming more efficient this has many benefits over the other methods (including the MCMC Newton-Raphson iteration method in [156] and the MCMC EM-algorithm). As the amount of Monte Carlo effort increases, consistency of the MCMCMLE, the *MCMC approximation to the MLE*, assuming the latter exists, is established in [68]. Further, there is a CLT for the Monte Carlo error, i.e. the deviation of the MCMCMLE from the MLE so that the accuracy of the approximation can be estimated (see [66, Theorem 7]) — this CLT is unrelated to the CLT for the MLE (as the 'observation window' increases) — asymptotic behaviour of the MLE for the lattice case is discussed in [120, 78, 158, 70, 38] and for spatial point processes in [121, 122, 99]. Moreover, estimates of Fisher information and likelihood ratio tests are provided by the MCMC likelihood method.

The *MCMC likelihood method* can be described in a general setup with data $X = x^*$ and densities $f_\theta(x) = c(\theta)h_\theta(x)$ with respect to the same reference measure $\mathrm{d}\mu(x)$ for all $\theta \in \Theta$, where the expression of the 'unnormalized density' h_θ is explicitly known. Suppose that $f(x) = ch(x)$ is another density with respect to μ such that

$$h(x) = 0 \;\Rightarrow\; h_\theta(x) = 0 \;\; \forall x, \theta$$

and we have generated a sample $X_1, \ldots, X_n$ from a Markov chain with limiting density f. Then one version of the log-likelihood function is

$$l(\theta) = \log(h_\theta(x^*)/h(x^*)) - \log(c(\theta)/c)$$

and the Monte Carlo approximation is given by

$$l_n(\theta) = \log \frac{h_\theta(x^*)}{h(x^*)} - \log\left[\frac{1}{n}\sum_1^n \frac{h_\theta(X_i)}{h(X_i)}\right]. \tag{23}$$

The MCMCMLE $\hat{\theta}_n$ maximizes (23). As (23) is based on importance sampling, f should not be too far from $f_{\hat{\theta}}$, where $\hat{\theta}$ is the MLE. Frequently one uses $f = f_\psi$ with ψ the MPLE or the approximate MLE obtained by a stochastic gradient algorithm. Alternatively an iterative procedure as described in [68] may be used. If f_θ belongs to an exponential family model, $l_n(\theta)$ is concave and the approximate score and Fisher information based on (23) are of a particular simple form — the mean and variance of the minimal canonical statistic are just replaced with Monte Carlo approximations obtained by using normalized importance sampling weights $(h_\theta(X_i)/h(X_i))/\sum_{j=1}^n (h_\theta(X_j)/h(X_j)),\ i = 1, \ldots, n$. Discussions on how to determine the burn-in and the spacing between samples can be found in [68, 67], [C], and the references therein.

Similarly for *missing data problems*. Suppose that we can write $X = (Y, Z)$ where only data $Y = y^*$ is observed. For example, in a point process setting, if X is a point process defined on $S = T \cup U$, where T is an 'observation window' and U is 'the rest', we may set $Y = X \cap T$ and $Z =$

$X \cap U$. Further, assume that $\mathrm{d}\mu(y,z) = \mathrm{d}\nu(y)\mathrm{d}\kappa(z)$ and we have generated samples $(Y_1, Z_1), \ldots, (Y_n, Z_n)$ and $Z_1^*, \ldots, Z_n^*$ from two Markov chains with limiting densities given by $f(y,z)$ and $f(z|y^*) = f(y^*, z)/\int f(y^*, w)\mathrm{d}\kappa(w)$, respectively. Then the Monte Carlo approximation to the log-likelihood becomes

$$l_n(\theta) = \log\left[\frac{1}{n}\sum_1^n \frac{h_\theta(y^*, Z_i^*)}{h(y^*, Z_i^*)}\right] - \log\left[\frac{1}{n}\sum_1^n \frac{h_\theta(Y_i, Z_i)}{h(Y_i, Z_i)}\right]. \qquad (24)$$

Though the MCMC likelihood method is easy to describe in theory, one should of course be careful in practice: as explained above it provides only a local approximation of the likelihood function; and one should carefully consider various diagnostics for convergence as briefly discussed at the end of Section 3.1 — perfect simulation procedures (Section 3.2) are still so new that they have yet not been used in connection to likelihood analysis. Further, multimodality of the likelihood function may cause problems for any optimization method including the MCMC likelihood method. Examples of applications of the MCMC likelihood method within spatial statistics and stochastic geometry are discussed in [C, U] and [67, 28]; many others may be expected to appear in the future when the possibility for making a detailed likelihood analysis in many situations is realised. For instance, many random set models in stochastic geometry are germ-grain models obtained by considering the union of simple geometrical objects (the grains) described by a marked point process, where the marked point process is unobserved and the random set is only observed within some bounded window. We show in [28] how MCMC likelihood methods for such missing data problems apply on the Boolean and other related models from stochastic geometry.

One question addressed in [C] is conditional and unconditional inference in the Strauss process (9). Traditionally, statistical inference for spatial point processes has been performed conditionally on the number n of points, probably mainly for computational convenience, but also because it is often claimed that n provides little information about the interaction parameter which is usually the feature of primary interest (see [59]). For example, in the case of the Strauss process (9), the density of the conditional Strauss process with respect to the binomial process, $f(x|n) \propto \gamma^{s(x)}$ depends only on the interaction parameter γ. Combining the MCMC likelihood method with the Metropolis-Hastings algorithm as described in Section 3.1, it is demonstrated in [C] that the unconditional and the conditional MLE of γ appear to be approximately identical. This seems to be explained by the fact that the Strauss process belongs to an exponential family model and its minimal canonical statistic $(n(X), s(X))$ appears to be approximately normally distributed. The latter observation is in accordance with the asymptotic result in [98, Theorem 1], though this result requires a condition to be satisfied which is practically enforcing the Poisson case as noticed in [C, Section 6]. It is shown in [121] that for a certain class of pairwise interaction point processes, the unconditional and the conditional MLE are asymptotically equivalent unconditionally, if the density of points is low enough. But apart from these observations it seems yet not known to what extent n can be considered as approximately ancillary for other types of point processes.

Variations of the MCMC likelihood method have also been considered. An interesting Bayesian smoothing method in the estimation of the pair potential function of pairwise interaction point processes is studied in [85]. As the full posterior in [85] is too demanding to simulate due to the unknown normalizing constant of the likelihood term, focus is on finding the posterior mode by means of algorithms similar to the MCMC likelihood method. It appears to me that many of the complications in [85] may be reduced if the likelihood term is replaced by a pseudo-likelihood function.

Finally, let me comment briefly on likelihood inference for parametric models of spatial birth-and-death processes as outlined in [S], where MCMC methods are not needed. In the case of full observation of a spatial birth-and-death process on a bounded time interval, the likelihood function is explicitly known, but in order to take edge effects into account an approximate likelihood function L is derived in [S, Section 4]. This function L factorizes into five terms, and based on three of these terms a partial likelihood function in the sense of Cox [42] can be derived. Existence of maximum likelihood estimates and maximum partial likelihood estimates, model checking, confidence intervals, and log-likelihood ratios for various statistical hypotheses are then discussed. In conclusion, compared to lattice and spatial point processes, the likelihood analysis is much more tractable. However, for many more complicated space-time processes other than spatial birth-and-death-processes, it is still a challenge for statisticians to develop satisfactory statistical models and inference procedures.

3.6 Statistical inference for log Gaussian Cox processes

Before we discuss estimation, model checking, and empirical Bayesian inference for log Gaussian Cox processes as introduced in Section 2.4, it is appropriate to first give a short description on how to simulate from an LGCP X and its underlying (continuous) Gaussian process Y when both are restricted to a bounded region, let's say the unit square $W = [0,1[^2$.

There is an efficient way of simulating first the Gaussian process and secondly the LGCP given the intensity process. For $(i,j) \in I = \{0,\dots,M-1\}^2$, the restriction of Y to the square cell $D_{ij} = [\frac{i}{M}, \frac{i+1}{M}[\times[\frac{j}{M}, \frac{j+1}{M}[$ is approximated by the value of the midpoint $\tilde{Y}_{ij} = Y(\frac{i}{M} + \frac{1}{2M}, \frac{j}{M} + \frac{1}{2M})$, where M is a suitable value for the discretization. We can extend $\tilde{Y} = (\tilde{Y}_{ij})_{(i,j)\in I}$ to a larger Gaussian field $\tilde{Y}_{ext} = (\tilde{Y}_{ij})_{(i,j)\in I_{ext}}$ with $I_{ext} = \{0,\dots,2(M-1)-1\}^2$ so that the covariance matrix K of $\tilde{Y}_{ext}$ is circulant (the distribution of $\tilde{Y}_{ext}$ is not in accordance with the distribution of Y outside W). Then K can easily be diagonalized by means of the two-dimensional discrete Fourier transform (see [T, Section 6.1] and [198]). Using a lexicographic ordering of the indices, we can set

$$Y_{ext} = \Gamma Q + \mu_{ext} \tag{25}$$

where Γ follows a standard multivariate normal distribution of dimension $d = \mathrm{rank}(K)$, Q is a certain $d \times (2(M-1))^2$ matrix of rank d, and the restriction of μ_{ext} to I agrees with the mean of $\tilde{Y}$. Thereby a fast simulation algorithm for $\tilde{Y}_{ext}$ and hence $\tilde{Y}$ is obtained. If $\tilde{Y} = \tilde{y}$ is the output, then we simulate a homogeneous Poisson process on W with intensity $\tilde{\lambda}_{max} = \max_I \exp(\tilde{y}_{ij})$. Finally, we are thinning from this so that a Poisson point situated in D_{ij} is retained with probability $\exp(\tilde{y}_{ij})/\tilde{\lambda}_{max}$.

Now, suppose that we have observed a point pattern $X \cap W = x$. For simulation from the conditional distribution of the Gaussian process given the data, we use in [T] a *Metropolis-adjusted Langevin algorithm* (MALA) as suggested in [22] and further studied in [174]. This is a Metropolis-Hastings algorithm with collective updating inspired by the definition of a Langevin diffusion. Actually, we use a MALA for conditional simulation of Γ given x and transform the output by (25); provided K is strictly positive definite, there is another MALA for simulating $\tilde{Y}_{ext}$ given x, but for the examples considered in [T] this algorithm mixes more slowly. Omitting an additive constant depending on x only, the log conditional density of Γ given x is

$$-||\gamma||^2/2 + \sum_{(i,j)\in I_{ext}} (\tilde{y}_{ij} n_{ij} - e^{\tilde{y}_{ij}}/M^2) \tag{26}$$

where, in accordance with (25), $(\tilde{y}_{ij})_{(i,j)\in I_{ext}} = \gamma Q + \mu_{ext}$ and $n_{ij} = n\big(x \cap [\frac{i}{M}, \frac{i+1}{M}[\times[\frac{j}{M}, \frac{j+1}{M}[\big)$, $(i,j) \in I_{ext}$. The gradient of (26) becomes

$$\nabla(\gamma) = -\gamma + (n_{ij} - e^{\tilde{y}_{ij}}/M^2)_{(i,j)\in I_{ext}} Q^*$$

and the conditional density of Γ given x is seen to be strictly log-concave. If γ is the current state generated by the MALA, then we first propose a new state generated from a multivariate normal distribution with mean $\gamma + (h/2)\nabla(\gamma)$ and covariance matrix hI, where I is the identity matrix and $h > 0$ is a user specified parameter, and secondly we accept or reject the proposal in accordance to the Hastings ratio. It is advantageous that all coordinates of Γ are updated simultaneously and that gradient information is used in the proposal distribution. For further details and a truncated version of the MALA which is geometrically ergodic, see [T, Section 8].

We turn now to parameter estimation in the homogeneous case where Y has a constant mean $\mu \in \mathbf{R}$, a constant variance $\sigma^2 > 0$, and isotropic correlation function $r_\beta(a) = r(a/\beta)$ for distances $a \geq 0$ and where $\beta > 0$ is a scale parameter. Here the particular form of the class of covariance functions $c_{(\beta,\sigma^2)} = \sigma^2 r_\beta$ may be suggested by considering a plot of the logarithm of a non-parametric kernel estimate $\hat{g}$ of the isotropic pair correlation function as discussed in [T, Section 7]. As the Gaussian process is unobserved, consider the MCMC approximation of the likelihood function obtained from (24) (with y^*, Z_i^*, Y_i, Z_i replaced by x, Y_i^*, X_i, Y_i using an obvious notation). Since the conditional density $f(x|y)$ does not depend on $\theta = (\mu, \sigma^2, \beta)$ it is only the unnormalized Gaussian density which appears in the approximation (so we need only a conditional sample $\{Y_i^*\}$ generated by the MALA and an unconditional sample $\{Y_i\}$ from the Gaussian process using the discrete Fourier method). Hence there may be some analogue to the situation in [171] where the difficulties with multimodality in likelihood functions for spatial Gaussian processes are discussed.

Since the distribution of the LGCP is completely determined by its first- and second-order properties we suggest instead in [T] to base the inference on the corresponding summary statistics $\hat{\rho} = n$ and $\hat{g}$, where $\rho > 0$ is the intensity of the point process. The parameters β and σ^2 are estimated by a *minimum contrast method* based on comparing the pair correlation function $g_{(\beta,\sigma^2)} = \exp(c_{(\beta,\sigma^2)})$ with its non-parametric estimate $\hat{g}$. This procedure is computationally much simpler than the usual minimum contrast method based on the K-function as suggested in [46]. Combining the estimates of σ^2 and ρ with the relation $\rho = \exp(\mu + \sigma^2/2)$, we finally obtain an estimate of μ.

Having estimated the parameters, we check the model assumptions by comparing non-parametric estimates of various summary statistics including the common functions $F,\ G,\ L = \sqrt{K/\pi}$, but also a new edge-corrected third-order characteristic which is of a particular simple form for LGCPs (see [T, Sections 3 and 7]). This is illustrated by analysing a data set consisting of the locations of Scots pine saplings in a square plot, where we conclude that the estimated LGCP model fits the data better than a previously fitted Matérn cluster point process in [157, 181].

The unobserved Gaussian process Y or intensity process $\Lambda = \exp(Y)$ may be predicted by using an *empirical Bayesian approach*. Here the posterior distribution is obtained by considering the Gaussian distribution of Y as a prior which smoothes the intensity process, the conditional distribution of $X \cap W$ given Y is the likelihood term, and the parameters of the prior may be estimated as described above. Extending the approximation $\tilde{Y}$ of the Gaussian process on W to $\tilde{Y}_{ext}$ and using the MALA for Γ given x together with the linear relationship (25), we obtain MCMC estimates of the posterior mean, credibility intervals, posterior uncertainty, etc. for the Gaussian process and the intensity process restricted to W. Also maximum a posteriori estimates can be derived easily due to log-concavity of the posterior density. However, for the Scots pine saplings the MAP-estimate of the intensity surface turned out to be useless. This may be due to the skewness and heavytailedness of the log Gaussian distribution combined with the correlation structure of $\exp(\tilde{Y})$. For the Scots pine saplings, the posterior mean of the intensity surface is quite peaked and different from a non-parametric kernel estimate proposed in [47]. To study these estimates under known conditions, a simulation study is reported in [T, Section 8]. Both for the Scots pine saplings and in the simulation study, the posterior mean appeared to be the best estimate.

In fact, the MALA can easily be extended to the situation where X is only observed in a subwindow $T \subset W$. Thereby the intensity process can still be predicted on W, and so by conditional simulation of the LGCP given the intensity process we can also predict X on $W \setminus T$.

Future investigations of LGCPs may include the possibility of using MCMC likelihood methods (including the MCMC EM-algorithm) and as mentioned in Section 2.4 to see to what extent inhomogeneous LGCPs can be used for analysing large point patterns. If it is assumed that the inhomogeneity can be described by the intensity function but the pair correlation function $g(s,t) = g(||s-t||)$ is homogeneous, one may suggest first to obtain a non-parametric estimate $\hat{\rho}(s)$ of the intensity function and another for the pair correlation function and then use (17) to obtain a non-parametric estimate of the isotropic covariance function. Thereby we can estimate the mean function of the Gaussian process by combining the non-parametric estimates with (16). Finally, we may proceed as in the homogeneous case considered above, fitting parametric isotropic covariance models, doing a model check, and making predictions as further discussed in [29].

4 Random tessellations

A tessellation of the d-dimensional Euclidian space is a subdivision of $\mathbf{R}^d$ into d-dimensional non-overlapping sets called cells, crystals, regions, tiles, etc. depending on the particular application (most work on tessellations focuses on the space $\mathbf{R}^d$, but subdivisions of other spaces like the $(d-1)$-dimensional sphere are of course of interest too). Examples of real-life tessellations are cell aggregates, foam, crack patterns, and crystalline structures. Frequently, the cells are assumed to be bounded and convex, and the number of cells intersecting any bounded subset of $\mathbf{R}^d$ is finite. Then the cells become d-dimensional convex polytopes.

Random tessellations are a major subject in stochastic geometry, where a general theory for stationary tessellations has been established during the seventies and the eighties; see [I], [180], and the references therein. Mainly mean value relations between certain geometrical characteristics of the cells have been studied using either Palm measure theory or an ergodic approach. Combining this with stereological results gives relations between characteristics of a random tessellation and its intersection with an affine subspace. Many of these relationships are collected and systematically studied in [I] for stationary tessellations with convex cells; see also the review in [180, Section 10]; the non-convex case is treated in [200].

Explicit probability models are typically constructed from spatial processes of simple geometrical objects such as points and hyperplanes. Perhaps the most important examples are random Voronoi tessellations generated by a point process and Johnson-Mehl tessellations generated by a space-time process. Another example is a random hyperplane process, which in the Poisson case provides one of the mathematically most tractable classes of models for random tessellations (see, e.g., [126, 130]), but their importance for practical applications seem to be somewhat limited. The cells under hyperplane and Voronoi tessellations are convex, while Johnson-Mehl cells can be non-convex.

Probabilistic aspects of Voronoi tessellations are discussed in Sections 4.1–4.4, while Johnson-Mehl tessellations are briefly studied in Section 4.5. Statistical aspects of random tessellations are discussed in Section 4.6.

4.1 Introductory remarks on Voronoi tessellations

Since Voronoi cells can be interpret due to their nearest-nucleus definition as territories, market areas, crystals etc., Voronoi tessellations provide models for many natural phenomena or processes studied in science and technology; numerous examples of applications are discussed in [155, 180] and [L]. A comprehensive collection of results on Voronoi and related tessellations can be found in [155], while a detailed mathematical study is given in [L].

The following Sections 4.2–4.4 are nearly identical to the review in [O] on the probabilistic aspects of Voronoi and Delaunay tessellations. In Section 4.2, we study some general properties of stationary Voronoi tessellations. Certain geometric constraints and their implications are discussed in Section 4.2.1. In Section 4.2.2, mean value relations of geometrical characteristics of a stationary Voronoi tessellation are considered. The approach is based on Palm measure theory (as developed by Mecke and others) which seems simpler than the ergodic approach used in many of the pioneering works by Meijering, Gilbert, Miles, and others. Section 4.3 is concerned with the case where the nuclei form a stationary Poisson process. This is the simplest non-trivial model for a random Voronoi tessellation and it will be outlined how various results may be derived for Poisson-Voronoi and Poisson-Delaunay tessellations. Finally, Monte Carlo studies, non-Poissonian models and more general constructions of tessellations are briefly discussed in Section 4.4.

4.2 Stationary Voronoi tessellations

In the sequel, stationarity (as defined below in Section 4.2.1) is assumed. Stationarity is a basic assumption in stochastic geometry which loosely speaking means that the statistical properties do not depend on the location of the observer. This may seem to be a strict assumption for certain applications, but often it is needed to make any progress in the analysis. At least within sufficiently small 'homogeneous' regions it may be a reasonable assumption.

4.2.1 Geometric constraints and properties

Most probabilistic models for the point process of nuclei (which generates the Voronoi tessellation) are very different from the highly regular patterns of nuclei used in the seminal work of Dirichlet [49] and Voronoi [195]. In this section we discuss the conditions to be satisfied by the nuclei of a Voronoi tessellation in order to establish the mean value relations in Section 4.2.2. Apart from restricting attention to the stationary case the discussion follows mostly the same lines as in [L, Section 2.1], where further details and proofs can be found.

Let $\Phi = \{x_i\} \subset \mathbf{R}^d$ denote a non-empty point pattern of nuclei generating the *Voronoi cells* $C_i = C(x_i|\Phi)$ as defined by

$$C(x_i|\Phi) = \{y \in \mathbf{R}^d : ||x_i - y|| \leq ||x_j - y|| \ \forall j \neq i\} \tag{27}$$

where $||\cdot||$ is the usual Euclidean distance. Usually, Φ is assumed to be almost surely locally finite, i.e., with probability one, $\Phi(B) = n(\Phi \cap B)$ (the number of nuclei in B) is finite for any bounded subset $B \subset \mathbf{R}^d$. Then each point in $\mathbf{R}^d$ belongs to finitely many Voronoi cells so that

$$\mathcal{V}(\Phi) = \{C_i\} = \{C(x_i|\Phi)\}$$

becomes a *Voronoi tessellation* of $\mathbf{R}^d$ into convex closed sets (which are possibly not bounded; a bounded Voronoi cell becomes a convex polytope). A part of a planar Voronoi tessellation is shown in Figure 3.

If the distribution of Φ and hence $\mathcal{V}(\Phi)$ is invariant under translations in $\mathbf{R}^d$, the point process of nuclei and the Voronoi tessellation are said to be stationary. This implies boundedness of the Voronoi cells.

Frequently, apart from local finiteness and stationarity, the nuclei are assumed to be in general quadratic position, that is, with probability one, no $k+1$ nuclei lie on a $(k-1)$-dimensional affine subspace of $\mathbf{R}^d$, $k = 2, \ldots, d$, and no $d+2$ nuclei lie on the boundary of a sphere. Then the facets of the Voronoi tessellation $\mathcal{V}(\Phi)$ defined by non-void intersections of

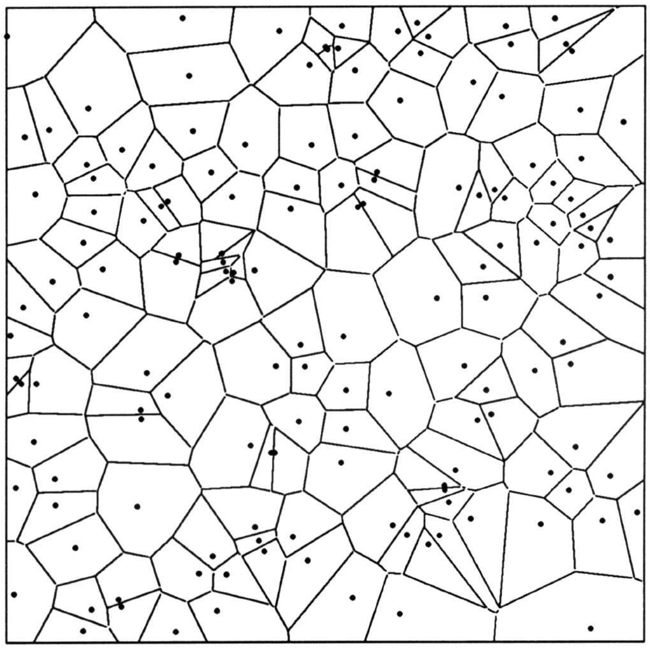

Figure 3: Planar Voronoi tessellation generated by a binomial process of 150 independent and uniformly distributed points on the unit square.

neighbouring cells,

$$F(x|\Phi) = \bigcap_{i=0}^{k} C(x_i|\Phi),\ 0 \leq k \leq d,\ x = \{x_0, \dots, x_k\} \subseteq \Phi \qquad (28)$$

have many desirable geometrical and topological properties: (i) Any $(d-k)$-facet $F(x|\Phi)$ defined by $k+1$ neighbouring Voronoi cells as in (28) is a bounded convex polytope of dimension $d-k$. (ii) Any $(d-k)$-facet lies in the relative boundaries of $(k+1)!/((l+1)!(k-l)!)$ intersecting $(d-l)$-facets, $0 \leq l \leq k \leq d$. (iii) The set of facets is locally finite as the number of facets intersecting a given bounded subset of $\mathbf{R}^d$ is finite. (iv) The set of $(d-k)$-facets, $1 \leq k \leq d$, coincides with the set of vertices of the Voronoi cells when $k = d$, edges of Voronoi cells when $k = d-1$, ..., and faces of Voronoi cells when $k = 1$. (v) The intersection between $d+2$ or more Voronoi cells is empty.

Similar properties hold for the dual *Delaunay tessellation* $\mathcal{T}(\Phi)$ when Φ is stationary, locally finite, and with points in general quadratic position. To each vertex $F(\{x_0, \dots, x_d\}|\Phi)$ of $\mathcal{V}(\Phi)$ we associate a *Delaunay cell* defined by the closed d-dimensional simplex $T(\{x_0, \dots, x_d\})$ with vertices given by the nuclei $x_0, \dots, x_d$. Hence the Delaunay tessellation

$$\mathcal{T}(\Phi) = \{T(\{x_0, \dots, x_d\}) : F(\{x_0, \dots, x_d\}|\Phi) \neq \emptyset,\ \Phi(\{x_0, \dots, x_d\}) = d+1\}$$

constitutes also a stationary tessellation of $\mathbf{R}^d$, where its facets of dimension $< d$ coincide with the vertices, edges, ..., faces of the Delaunay cells. Clearly, by duality, the k-facets of $\mathcal{T}(\Phi)$ are in one-to-one correspondence with the $(d-k)$-facets of $\mathcal{V}(\Phi)$.

Voronoi tessellations in general quadratic position are *normal* tessellations, which means that every $(d-k)$-facet of $\mathcal{V}(\Phi)$ lies in the boundaries of exactly $k+1$ Voronoi cells $(0 < k \leq d)$. Many real-life non-artificial tessellations have this property. Though Delaunay tessellations are not normal, many results for Voronoi tessellations carry over because of the duality.

A stationary Poisson process as considered in Section 4.3 is almost surely locally finite with points in general quadratic position. Hence, any point process model, which restricted to any bounded Borel set (or ball) B in $\mathbf{R}^d$ is absolutely continuous with respect to a stationary Poisson process, is also almost surely locally finite with points in general quadratic position. For instance, this is the case for the models considered in Section 4.4.

4.2.2 Palm distributions and mean value relations of geometrical characteristics

Henceforth suppose that $\Phi = \{x_i\}$ is stationary and almost surely locally finite with points in general quadratic position. In the following we define the concepts of 'typical' cells and facets. Further, we review certain mean value relations for the geometrical characteristics of the facets of $\mathcal{V}(\Phi)$ and $\mathcal{T}(\Phi)$. The 'practical' planar ($d = 2$) and spatial ($d = 3$) cases have been studied in [128, 129, 166, 180] and [L]; general definitions and results in $d \geq 1$ dimensions are given in [I]. These results allow us to parameterize all mean values considered by only a few of them, which in turn often may be easily estimated by non-parametric methods. In fact many of the relations hold in general for stationary tessellations with convex cells; others utilize only the normality of $\mathcal{V}(\Phi)$. The special structure of a Voronoi tessellation plays a more important role when considering more specific models as in Sections 4.3–4.4.

Typical cells and facets of Voronoi tessellations:
Let us first consider the simplest kind of characteristics. For $k = 0, \dots, d$, let

$$\Lambda_{d-k}(B) = E \sum_{x \subset \Phi : \Phi(x) = k+1} |F(x|\Phi) \cap B|_{d-k}$$

be the mean $(d-k)$-content of all Voronoi facets of dimension $d-k$ observed within a 'window' $B \subset \mathbf{R}^d$ with $0 < |B|_d < \infty$. So $\Lambda_0(B)$ is the mean number of Voronoi vertices in B, $\Lambda_1(B)$ is the mean length of edges in B, and so forth. By stationarity,

$$\Lambda_{d-k}(B) = \lambda_{d-k}|B|_d$$

where the constant λ_{d-k} is the *density* of $(d-k)$-facets. We assume that $0 < \lambda_{d-k} < \infty$ for $0 < k \le d$ (trivially, $\lambda_d = 1$).

We turn next to the definition of the typical Voronoi cell. Let λ be the intensity of Φ, that is the mean number of nuclei per unit volume in $\mathbf{R}^d$. Assume that $0 < \lambda < \infty$ and set $\Phi_\xi = \{x_i + \xi\}$, the translate of nuclei $\Phi = \{x_i\}$ by $\xi \in \mathbf{R}^d$. The *Palm distribution* of the *typical Voronoi cell* $\mathcal{C}$ is defined by the so-called *Campbell-Mecke formula*,

$$E \sum_i h(C(0|\Phi_{-x_i}), x_i) = \lambda \int Eh(\mathcal{C}, \xi) \mathrm{d}\xi \qquad (29)$$

for non-negative and measurable functions h (here defined on the product space of the set of d-polytopes and $\mathbf{R}^d$; further, E means expectation with respect to the distribution of Φ respective $\mathcal{C}$ in the left respective right hand side of (29)). Intuitively, $\mathcal{C}$ follows the conditional distribution of a Voronoi cell given that its nucleus is located at an arbitrary fixed point (here taken to be the origin 0). In the ergodic case, we can think of $\mathcal{C}$ as a 'uniformly picked' Voronoi cell in a region 'expanding to $\mathbf{R}^d$'. In the special case where Φ is a stationary Poisson process the definition of $\mathcal{C}$ has another easy interpretation, see (35), Section 4.3. Moreover, if h is invariant under translations of the cells, i.e. when $h(C_\xi, \eta) = h(C, \eta)$ for d-polytopes C and $\xi, \eta \in \mathbf{R}^d$, then

$$Eh(\mathcal{C}, \xi) = E[h(C^*, \xi)/|C^*|_d] E|\mathcal{C}|_d = E[h(C^*, \xi)/|C^*|_d]/\lambda \qquad (30)$$

where C^* denotes the almost surely well-defined Voronoi cell $C_i = C(x_i|\Phi)$ with $0 \in C_i$ (by stationarity, 0 may be replaced by any other fixed point in $\mathbf{R}^d$); we refer to C^* as a *point sampled Voronoi cell.*

Similarly, one defines a typical $(d-k)$-facet $\mathcal{F}_{d-k}$ whenever $0 \le k \le d$ (the typical d-facet is nothing but the typical Voronoi cell, $\mathcal{F}_d = \mathcal{C}$). For any $k+1$ nuclei $x = \{x_0, \dots, x_k\} \subset \Phi$, let $z(x|\Phi)$ be a given 'center' point or centroid associated to $F(x|\Phi)$ which is covariant under translations, i.e. $z(x_\xi|\Phi_\xi) = \xi + z(x|\Phi)$; in accordance with (29) we can simply take $z(\{x_0, \dots, x_k\}|\Phi) = \sum_0^k x_i/(k+1)$ which is clearly covariant (there are other ways of specifying the centroid which become more useful if only the cells and not the nuclei are observed, cf. e.g. [L]). Then

$$E \sum_{x \subset \Phi:\, F(x|\Phi) \neq \emptyset,\, \Phi(x)=k+1} h(F(x|\Phi)_{-z(x|\Phi)}, z(x|\Phi)) \\ = I_{d-k} \int \cdots \int Eh(\mathcal{F}_{d-k}, z)\, \mathrm{d}z \tag{31}$$

for non-negative and measurable functions h. Here

$$I_{d-k} = E \sum_{x \subset \Phi:\, F(x|\Phi) \neq \emptyset,\, \Phi(x)=k+1} \mathbf{1}[z(x|\Phi) \in B]/|B|_d$$

is the *intensity* of $(d-k)$-facets, where $\mathbf{1}[\cdot]$ denotes indicator function.

It can be shown that this definition of I_{d-k} does not depend on the choice of the centroid. Moreover, in (29)-(31), if h is invariant under translations of the facet, i.e. when $h(F_{-z}, x) = h(F, x)$, we have that the expressions do not depend on the choice of centroid. In particular, $\lambda = I_d$.

Numerous mean value relations between such translation invariant functions can be established using (31) in combination with Fubini's theorem and the geometrical and topological properties studied in Section 4.2.1 above (see [I, Section 5]). For instance, for stationary tessellations in general, we have that

$$\lambda_{d-k} = I_{d-k} E|\mathcal{F}_{d-k}|_{d-k}, \quad 0 \leq k \leq d.$$

Due to the normality of $\mathcal{V}(\Phi)$ we have for $0 \leq l \leq k \leq d$ that

$$\binom{k+1}{k-l} I_{d-k} E|\mathcal{F}_{d-k}|_{d-k} =$$

$$I_{d-l} E|\bigcup\{(d-k)\text{-facets contained in } \mathcal{F}_{d-l}\}|_{d-k}$$

and

$$\binom{k+1}{k-l} I_{d-k} = I_{d-l} E \mathsf{card}(\{(d-k)\text{-facets contained in } \mathcal{F}_{d-l}\}).$$

Using these relations it is possible to express all such mean values by λ and λ_1 if $d = 2$:

$$I_0 = 2\lambda, \quad I_1 = 3\lambda, \quad I_2 = \lambda, \quad E\mathsf{card}(\{\text{vertices of } \mathcal{C}\}) = 6, \tag{32}$$

$$E\mathsf{area}(\mathcal{C}) = 1/\lambda, \quad E\mathsf{perimeter}(\mathcal{C}) = 2\lambda_1/\lambda. \tag{33}$$

Results for the spatial case $d = 3$ and higher dimensions can be found in the references mentioned at the beginning of this section.

Typical Delaunay cell and facets:
In a similar way we define the typical Delaunay cell and facets; see also Section 4.3.2. Then the general results for stationary tessellations with convex cells apply. Let us here just utilize the duality between $\mathcal{V}(\Phi)$ and $\mathcal{T}(\Phi)$ to obtain (with obvious notation) that

$$I^{\mathcal{T}}_{d-k} = I^{\mathcal{V}}_k \ (= I_k), \quad 0 \le k \le d.$$

Results for Poisson-Delaunay tessellations are considered in Section 4.3.2.

Sectional Voronoi tessellations:
The intersection between $\mathcal{V}(\Phi)$ and a given affine subspace $L_n \subset \mathbf{R}^d$ of dimension $n < d$ forms a stationary and normal tessellation of L_n. This sectional Voronoi tessellation is in general not a Voronoi tessellation. An important counterexample is provided by a sectional Poisson-Voronoi tessellation (see [35]).

General formulae relating the characteristics of the sectional Voronoi tessellation to those of $\mathcal{V}(\Phi)$ can be found in [I, L] and [180]. Often isotropy is assumed in order to obtain useful relations, i.e. the distribution of either Φ or L_n is assumed to be invariant under rotations about the origin in $\mathbf{R}^d$. Stereological aspects of Voronoi tessellations are further discussed in [I, L] and [180, Section 1.5.4].

4.3 Poisson-Voronoi and Poisson-Delaunay tessellations

Throughout this section Φ is assumed to be a stationary Poisson process in $\mathbf{R}^d$ with intensity $\lambda \in (0, \infty)$. The usual characterisation of this distribution is given by that $\Phi(B) = n(\Phi \cap B)$ follows a Poisson distribution with mean $\lambda |B|_d$ for any bounded Borel set $B \subset \mathbf{R}^d$ (this property implies complete spatial independence in the sense that the points in $\Phi \cap B$ conditional on $\Phi(B)$ are statistically independent and uniformly distributed on B). Another less used but extremely fruitful characterisation is the so-called *Slivnyak-Mecke formula* established in [127] or its generalisation stated as follows:

$$E \sum_{x \subset \Phi:\, \Phi(x) = k+1} f(\Phi, x) = \tag{34}$$
$$\frac{\lambda^{k+1}}{(k+1)!} \int \cdots \int E f(\Phi \cup \{x_0, \ldots, x_k\}, \{x_0, \ldots, x_k\})\, \mathrm{d}x_0 \cdots \mathrm{d}x_k$$

for any integer $k \geq 0$ and non-negative measurable functions f. Notice that a stationary Poisson process is also isotropic.

Combining (29) and (34) we see that for the *typical Poisson-Voronoi cell*,

$$\mathcal{C} = C(0|\Phi \cup \{0\}) \tag{35}$$

meaning that it is distributed as the Voronoi cell with nucleus 0 if we have added 0 to Φ. Moreover, for translation invariant non-negative and measurable functions h defined on the space of $(d-k)$-dimensional convex polytopes, it can be verified using (31) and (34) that

$$I_{d-k} |B|_d E h(\mathcal{F}_{d-k}) =$$
$$\frac{\lambda^{k+1}}{(k+1)!} \int \cdots \int E\{\, h(F(\{x_0, \ldots, x_k\} | \Phi \cup \{x_0, \ldots, x_k\})) \tag{36}$$
$$\mathbf{1}[F(\{x_0, \ldots, x_k\} | \Phi \cup \{x_0, \ldots, x_k\}) \neq \emptyset,\ \sum_0^k x_i/(k+1) \in B]\,\}$$
$$\mathrm{d}x_0 \cdots \mathrm{d}x_k$$

where $F(\{x_0, \ldots, x_k\}|\Phi \cup \{x_0, \ldots, x_k\}) = \{y \in \mathbf{R}^d : ||y - x_0|| = \ldots = ||y - x_d|| \leq ||y - \xi|| \, \forall \xi \in \Phi\}$. The latter mean value $E\{\cdots\}$ in (36) can be evaluated for certain polytope charactersitics $h(F)$, indicating why (34) is so applicable.

4.3.1 Exact and numerical results for Poisson-Voronoi tessellations

Density of facets and other exact results:
Another extremely useful result with applications in stochastic geometry is the so-called *Blaschke-Petkanschin formula*; see the survey in [96, Chapter 4]. A simple proof of this integral geometric result based on invariant measure theory is established in [G] — most derivations of the Blaschke-Petkanschin formula have been based on manipulation of differential forms.

Using Blaschke-Petkanschin's formula it can be verified that for $0 \leq k \leq d$,

$$\lambda_{d-k} = \frac{\lambda^{k/d} 2^{k+1} \pi^{k/2} \Gamma(\frac{dk+d-k+1}{2}) \Gamma(\frac{d}{2}+1)^{k+((d-k)/d)} \Gamma(k + \frac{d-k}{d})}{(k+1)! \Gamma(\frac{dk+d-k}{2}) \Gamma(\frac{d+1}{2})^k \Gamma(\frac{d-k+1}{2}) d}. \tag{37}$$

Here and elsewhere in the following $\Gamma(\cdot)$ is the gamma function. This result for arbitrary dimensions d is due to Miles [136] who used an ergodic approach (see also [137]; for proofs based on (34), see [I, L, R]). Combining (37) with mean value relations as discussed in Section 4.2.2 various mean value characteristics for both the Poisson-Voronoi tessellation and its sectional tessellations can be calculated, see [I, Section 7; R, Theorem 3] and the review in [180, Section 10.6.1]. For example, if $d = 2$ we get $\lambda_1 = 2\sqrt{\lambda}$, which can be inserted into (33). Indeed many of the results when $d \leq 3$ were already established in [131].

Equation (34) has been used for deriving other exact results. In [V] various gamma-type results and conditional independence results between size and shape of different geometrical characteristics determined by a Poisson process are derived. One example is the fundamental region Δ of the typical Poisson-Voronoi cell as defined by the union of balls with centres at the vertices of $\mathcal{C}$ and containing 0 on their boundaries. Conditionally on $\mathcal{C}$

having $n\,(\geq d+1)$ faces, $|\Delta|_d$ is independent of the shape and orientation of Δ, and $|\Delta|_d$ follows a gamma distribution with shape parameter n and scale parameter $1/\lambda$; see also [L] and [138, 201] for other derivations of these results. Other examples of results of this type will be considered in Section 4.3.2.

Using (34) it can also be shown that the mean number of Gabriel (or full) neighbours to the typical Poisson-Voronoi cell is 2^d. The edges in a Gabriel graph are given by those line segments which join pairs of nuclei ξ, η so that the intersection between the line segment and the two Voronoi cells with nuclei ξ, η is non-void (the Gabriel graph is a connected subgraph of the Delaunay graph). If L is the length of the typical Gabriel edge, then L^d follows an exponential distribution with mean given by $\Gamma(1+d/2)/((\pi/4)^{d/2}\lambda)$ (see [L]).

Second-order characteristics and other numerical results:
Higher order moments and many other properties of the Poisson-Voronoi tessellation than those considered above are often more complicated to derive and require in general either numerical methods or simulation studies (see Section 4.4.1). Results obtained by numerical integration are briefly discussed below.

For $d \leq 3$, Gilbert [71] used (30) to obtain an integral expression for the variance of the size of the typical Poisson-Voronoi cell and typical sectional cells. Covariances and variances of other kind of characteristics can be calculated as well; in general the details are intricate. This is demonstrated in two unpublished papers by Brakke [26, 27]. Tables showing Gilbert's and Brakke's results can be found in [L, Section 4.2]. Also figures of Brakke's numerical results for the density functions of the length of the typical planar and spatial Poisson-Voronoi edge are shown; see also [146]. Table 10.4 in [180] (based on [138]) shows the distribution of the number of vertices in the typical planar Poisson-Voronoi cell.

Summary statistics as commonly used in stochastic geometry have also been considered. For the random process of Poisson-Voronoi faces, [147] contains results for spherical and linear contact distribution functions in various dimensions d, while the pair correlation function for the point process of vertices in planar and spatial Poisson-Voronoi tessellations are calculated in [87, 86].

4.3.2 Results for Poisson-Delaunay tessellations

Miles determined the distribution of the *typical Poisson-Delaunay cell* $\mathcal{D}$ (see [136]; a simple proof based on (34) is contained in [I, Theorem 7.5]). Let $RU_0, \ldots, RU_d$ denote the $d+1$ vertices of the typical Delaunay cell. This is almost surely a d-dimensional simplex centered at 0, where $R > 0$ is the typical vertex-nucleus distance and $U_0, \ldots, U_d$ are unit vectors. Then R is statistically independent of the directions $(U_0, \ldots, U_d)$, and R^d is gamma distributed with shape parameter d and scale parameter $\Gamma(1+d/2)/(\lambda\pi^{d/2})$. Further, the joint density function of $(U_0, \ldots, U_d)$ is proportional to the d-content of the $(d+1)$-simplex determined by these unit vectors (here the density is with respect to the uniform distribution on the product space of $d+1$ unit spheres in $\mathbf{R}^d$; the constant of proportionality is also known).

Thereby the moments of $|\mathcal{D}|_d$ can be derived:

$$E|\mathcal{D}|_d^k = \frac{(d+k-1)!\Gamma(\frac{d^2}{2})\Gamma(\frac{d^2+dk+k+1}{2})\Gamma(\frac{d+1}{2})^{d-k+1}\prod_{i=2}^{d+1}\Gamma(\frac{k+i}{2})/\Gamma(\frac{i}{2})}{(d-1)!\Gamma(\frac{d^2+1}{2})\Gamma(\frac{d^2+dk}{2})\Gamma(\frac{d+k+1}{2})^{d+1}(2^d\pi^{(d-1)/2}\lambda)^k}$$

for $k = 1, 2, \ldots$ The fact that R^d is gamma distributed and independent of $(U_0, \ldots, U_d)$ follows also from the more general results in [V]. A related result verified in that paper is the following. Let now $RU_0, \ldots, RU_d$ be the $d+1$ vertices of the (almost surely well-defined) point sampled Delaunay cell D^* from $\mathcal{T}(\Phi)$ which contains 0. Then R is gamma distributed with the same scale parameter as before, but with shape parameter $d+1$ (i.e. one larger than before), and R is statistically independent of $(U_0, \ldots, U_d)$.

Miles' result plays a fundamental role in the statistical theory of shape (see [106]). It has further been used in [167] for deriving the distribution of the area when $d = 2$ and the volume when $d = 3$ of the typical Poisson-Delaunay cell. The density for the planar case $d = 2$ involves a modified Bessel function; the expression for $d = 3$ becomes more complicated. Also the distributions of the typical angle at a vertex of the planar

Poisson-Delaunay tessellation and the typical edge of the Poisson-Delaunay tessellation when $d = 2, 3$ have been determined; see [L] and the references therein.

4.4 Simulation and more general models

In the previous sections mathematically tractable properties of Poisson-Voronoi tessellations were outlined. Further analysis of random Voronoi tessellations require Monte Carlo studies as briefly discussed in Section 4.4.1. Finally, Section 4.4.2 contains a few remarks on some generalisations: one possibility is to replace the Poisson process of nuclei by a Gibbs point process; another is to generalise the construction of the Voronoi tessellation.

4.4.1 Simulated results

Most simulation studies of Voronoi tessellations have so far been concerned with the Poisson case. Relying on the ergodic approach, [90] reports on a large scale study for approximating the polygonal characteristics of a typical planar Poisson-Voronoi cell by Monte Carlo methods. A more efficient method is based on combining (35) with the *radial simulation algorithm* [165] for a stationary Poisson process in $\mathbf{R}^d$. The radial simulation algorithm is also tailor-made for simulating a Voronoi tessellation within a ball when the nuclei come from a stationary Poisson process or a related point process obtained by thinning or clustering such as Matérn hard-core processes, Matérn cluster processes and other kinds of Poisson cluster processes with uniformly bounded clusters — thereby, the problem with edge effects can be avoided. Several simulated results are shown in [L, M] and [90, 165, 89, 141, 114, 115, 194].

4.4.2 Generalisations of Voronoi tessellations

As mentioned it is the Slivnyak-Mecke formula (34) which mainly makes the Poisson-Voronoi model and related models relatively tractable for mathematical analysis. This formula can be extended to characterise a Gibbs point process Φ by

$$(k+1)!E \sum_{x\subset\Phi:\, \Phi(x)=k+1} f(\Phi,x) = \tag{38}$$

$$\int\cdots\int E[\lambda(\{x_0,\dots,x_k\}|\Phi)f(\Phi\cup\{x_0,\dots,x_k\},\{x_0,\dots,x_k\})]$$
$$\mathrm{d}x_0\cdots\mathrm{d}x_k$$

for any integer $k \geq 0$ and non-negative measurable functions f; in fact (38) follows from the Nguyen-Zessin characterisation result (22) using induction on k, where $\lambda(\{x_0,\dots,x_k\}|\Phi)$ is the so-called conditional intensity defined by the exponential of minus 'the energy required to add $\{x_0,\dots,x_k\}$ to Φ'. Hence, (36) is easily modified to cover stationary Gibbs point processes as well.

For example, consider a hard-core Gibbs point process with

$$\lambda(\{x_0,\dots,x_k\}|\Phi) =$$
$$\beta^{k+1}\mathbf{1}[||\xi-\eta|| \geq \delta \text{ for distinct points } \{\xi,\eta\}\subset\Phi\cup\{x_0,\dots,x_k\}]$$

where $\beta > 0$ and $\delta > 0$ are parameters. Then each Voronoi cell contains the ball of diameter δ centered at its nucleus, and so as β increases we get more and more regular Voronoi cells. Another possibly even more interesting model is obtained by replacing the hard-core condition $||\xi-\eta|| \geq \delta$ by a hard-core condition on the size of the Voronoi cells (*Ord's process* studied in [A]). Though such models may be more realistic for applications than Poisson models, and (38) makes it possible to obtain various estimation equations for the model parameters as well as the characteristics of the accompanying Voronoi tessellation, it still remains to study these promising aspects in detail.

The construction of a Voronoi tessellation has been generalised in various ways. (i) Some probabilistic results for the generalised Voronoi tessellation $\mathcal{V}_n(\Phi)$, $n=1,2,\dots,$ — also called nearest order n diagrams as they extend the definition of $\mathcal{V}_1(\Phi)=\mathcal{V}(\Phi)$ by considering points closest to pairs of nuclei, triplets of nuclei, etc. — have been established in [134, 138] when

Φ is a stationary Poisson process. (ii) Considering the Voronoi tessellation as the result of growing nuclei with same speed and start of growth, one can generalise this growth process to obtain a Johnson and Mehl tessellation, see Section 4.5. (iii) Yet another generalisation is to replace the Euclidean distance used in the definition (27) of Voronoi cells with another Euclidean metric so that the growth is anisotropic (see [178]). (iv) Indeed many other related tessellations to the Voronoi tessellation have been considered in the literature (see, e.g., [155]), but present knowledge on random models for many of these constructions is rather limited and fragmentary.

4.5 Johnson-Mehl tessellations

Probabilistic models of tessellations with non-convex cells are more scarce and complex. Johnson-Mehl tessellations form perhaps the most tractable and flexible class of such models. A Johnson-Mehl tessellation results from a spatial birth process $\Phi = \{a_i\}$ of arrivals $a_i = (x_i, t_i) \in \mathbf{R}^d \times [0, \infty)$, where the nucleus x_i starts to grow at time t_i with a constant speed $v > 0$ in all directions so that we may imagine that a point $y \in \mathbf{R}^d$ is reached by x_i at time $T_i(y) = T(y, a_i) = t_i + \|x_i - y\|/v$. If the set

$$C(a_i \mid \Phi) = \{y \in \mathbf{R}^d \mid \forall a_j \in \Phi : T_i(y) \leq T_j(y)\} \tag{39}$$

of points in $\mathbf{R}^d$ first reached by a_i is non-empty, then $C(a_i \mid \Phi)$ is said to be the crystal associated with a_i. Under certain consistency conditions (see [J, R]) these crystals constitute a random tessellation, the *Johnson-Mehl tessellation generated by* Φ. Illuminating figures of this construction can be found in [117, 58]. A part of a planar Johnson-Mehl tessellation is shown in Figure 4. Examples of 3-dimensional crystals can be found in [L, M, R]. Voronoi tessellations appear as the special case with all crystals being convex, i.e. when all birth times t_i are equal.

Johnson-Mehl tessellations have been introduced as simplified models for crystal growth as discussed in papers by Kolmogorov, Johnson and Mehl and others [112, 101, 4, 5, 6, 51]. The so-called ASTM model, which is frequently used in metallurgy, can be described by a Johnson-Mehl tessellation (see [92]). Notice that (39) agrees with the 'physical definition' of

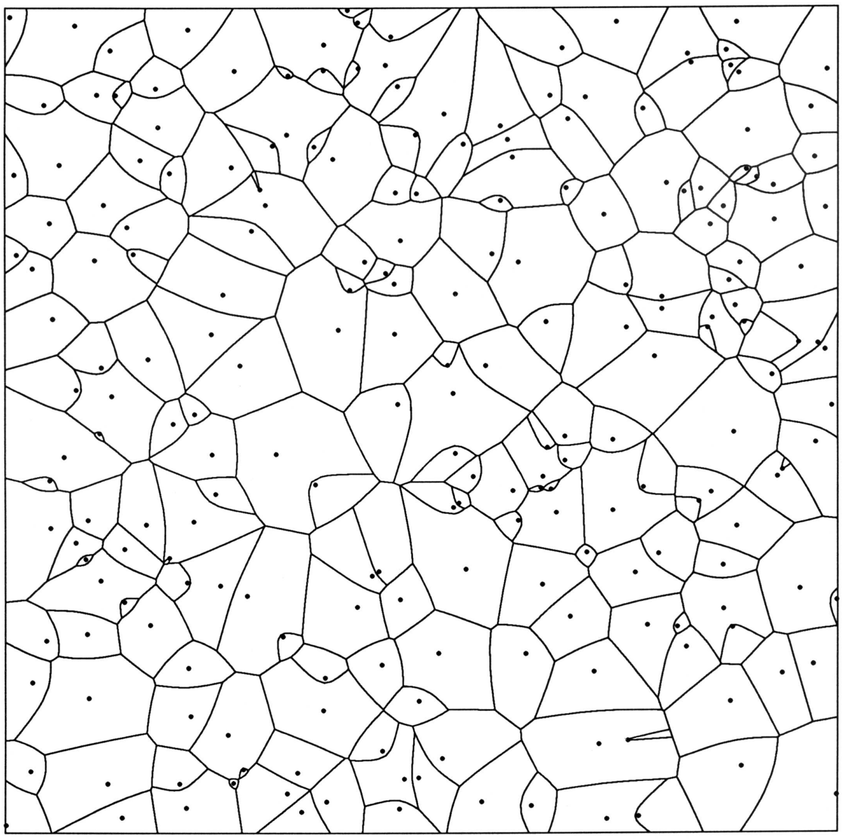

Figure 4: Johnson-Mehl tessellation generated by 300 nuclei within the unit square. To each nucleus is attached an exponentially distributed waiting time for the start of growth (mean waiting time = 100). These waiting times are mutually independent and independent of the positions of the nuclei which are generated by a binomial process. Only those nuclei which generate non-empty crystals are shown (total=222).

crystals where the growth of a crystal along a ray from its nucleus stops when it meets another growing crystal. The assumption of spherical growth at a constant speed is essential at least in the non-convex case — generalizations with non-spherical growth or a time dependent velocity seem not to be mathematically tractable when the physical definition is used. If the 'mathematical definition' (39) is used instead, then rather strange phenomena are allowed to occur: in the non-spherical case when, for example, crystals grow as boxes, crystals given by (39) may be disconnected; and if the growth rate decreases as a function of the distance to the nucleus, then nuclei may lie outside the crystals (39) which they generate; see also [51, 71, 135]. Similar problems occur if nuclei grow with constant but different velocities.

The geometric structure of Johnson-Mehl tessellations is more complicated than for Voronoi tessellations as explained in [J, R]. In the planar case, the edges of Johnson-Mehl crystals are hyperbolic arcs and three neighbouring crystals may share one or two vertices — while the edges of Voronoi cells are line segments and three neighbouring cells meet at one vertex only — compare Figures 3 and 4. In the spatial case and in higher dimensions, Johnson-Mehl crystals violate topological relations which are satisfied by Voronoi polytopes; e.g. a 'lense' has two faces, one edge, and no vertices. This implies deviations from the mean value relations which hold for Voronoi tessellations when $d \geq 3$ (see [J, Section 4.2]).

The case $d = 1$ has been studied in detail in [131], but as discussed in [71], the results do not lead to much insight for Johnson-Mehl tessellations in higher dimensions.

Most authors consider birth processes Φ which are Poissonian with a constant intensity, i.e. the intensity measure is both time-homogeneous and invariant under translations in $\mathbf{R}^d$ (stationarity), but the time-inhomogeneous case may be included as well as demonstrated in [135, 91, 92] and [J, M, R]. Specifically, suppose that Φ is a Poisson process on $\mathbf{R}^d \times [0, \infty)$ with an intensity measure which is a product measure of the Lebesgue measure on $\mathbf{R}^d$ and a locally finite measure Λ on $[0, \infty)$ with $\Lambda([0, \infty)) > 0$ (ensuring that Φ is almost surely non-empty) and so that

$$\lambda = \int p(t) \mathrm{d}\Lambda(t)$$

is finite. Here λ is the *intensity of crystals* and

$$p(t) = \exp\{-v^d \omega_d \int_0^t (t-s)^d \mathrm{d}\Lambda(s)\}$$

is the probability that an arbitrary fixed point in $\mathbf{R}^d$ is not reached by any growing nuclei at time t, where $\omega_d = \pi^{d/2}/\Gamma(\frac{1}{2}d+1)$ is the volume of the unit ball in $\mathbf{R}^d$. Various desirable geometrical and topological properties are satisfied for the crystals and facets of the Poisson-Johnson-Mehl tessellation (see [J, Proposition 3.1]). Moreover, if we consider the sectional tessellation obtained by intersecting the Poisson-Johnson-Mehl tessellation with a fixed k-dimensional affine subspace L_k of $\mathbf{R}^d$ with $0 < k \leq d$, the *density* $\mu_{k,p}$ *of* p*-facets* in this sectional tessellation (i.e. the mean p-content of p-facets per unit volume in L_k) is of a particularly nice form:

$$\mu_{k,p} = v^{md+n} c_{k,p} \int_0^\infty \{\int_0^t (t-s)^{d-1} \mathrm{d}\Lambda(s)\}^{m+1} p(t) \mathrm{d}t \tag{40}$$

with $m = k - p$, $0 \leq p \leq k$, and

$$c_{k,p} = \frac{2^{m+1}\pi^{(m+1)d/2}\Gamma\left(\frac{dm+n+1}{2}\right)\Gamma\left(\frac{k+1}{2}\right)}{(m+1)!\Gamma\left(\frac{dm+n}{2}\right)\Gamma\left(\frac{d+1}{2}\right)^{m+1}\Gamma\left(\frac{p+1}{2}\right)}. \tag{41}$$

A general proof of (40)-(41) was established in [R]. The formulae have previously been verified in certain special cases. Meijering [131] considered the time-homogeneous case $\Lambda(dt) = \alpha dt$ and with $k = d \leq 3$, treating each of the cases $p = 0, 1, 2$ separately. Miles [135] stated the results for the inhomogeneous case when $k = d = 3$ but without including a proof. Furthermore, it was shown in [J] that (40)-(41) hold for each of the cases (i) $p = 0$ and $0 < k \leq d$, (ii) $p = 1$ and $0 < k \leq d$, (iii) $k = d$ and $p = 2$; this covers all cases of p and k when $d \leq 3$.

Notice that (40) can be written as

$$\mu_{k,p} = \text{ constant } \int_0^\infty \text{ haz}(t)^{m+1} p(t) dt \tag{42}$$

where

$$\text{haz}(t) = -\frac{p'(t)}{p(t)} = v^d \sigma_d \int_0^t (t-s)^{d-1} \Lambda(ds)$$

is the 'hazard function' of the time at which an arbitrary point is crystallized and $\sigma_d = 2\pi^{d/2}/\Gamma(d/2)$ is the surface area of S^{d-1}. Formula (42) may be considered as a consequence of the strong independence properties of the Poisson process (or more precisely, the Slivnyak-Mecke formula (34)), but the perhaps surprising fact is that the constant in (42) depends only on the dimensions d, k, p and not on the measure Λ. If in (40) and (42) we replace ∞ by $T > 0$, then we get the density of those parts of p-interfaces in L_k which have been generated before time T.

Further first- and second-order characteristics of stationary Johnson-Mehl tessellations are studied in [J]. Most results for Poisson-Johnson-Mehl tessellations involve integration with respect to the intensity measure Λ. In order to evaluate these integrals exactly, Λ might be a linear combination of the following cases: (M1) a superposition of countably many stationary Poisson processes in $\mathbf{R}^d$ which arrive at different times; (M2) $\mathrm{d}\Lambda(t) = \alpha t^{\beta-1}\mathrm{d}t$, where $\alpha, \beta > 0$ are parameters. The case (M2) is extensively studied in [J, Section 5]; see also [R, Section 5], where among other things it is discussed how the results may be used for parameter estimation. Note that the distribution of the Poisson-Johnson-Mehl tessellation under (M2) is parameterized by (α, β) as we can set $v = 1$. Under (M2),

$$p(t) = \exp\left(-\alpha\omega_d B(d+1, \beta) t^{d+\beta}\right)$$

and

$$\lambda = \frac{\alpha^{d/(d+\beta)}}{\beta} \Gamma\left(1 + \frac{\beta}{d+\beta}\right) \left(\omega_d B(d+1, \beta)\right)^{-\beta/(d+\beta)}$$

where $B(\cdot,\cdot)$ is the beta function. Thus another parameterization is provided by the scale parameter $\lambda^{-1/d} > 0$ and the shape parameter $\beta > 0$. Explicit expressions of the densities $\mu_{k,p}$ can be found in [J, R].

The density of the birth time τ of the *typical crystal* $\mathcal{C}$ of a Poisson-Johnson-Mehl tessellation is given by $p(t)\mathrm{d}\Lambda(t)/\lambda$. Further, $\mathcal{C}$ is stochastically equivalent to $C((0,\tau)|\Phi_\tau \cup \{(0,\tau)\})$, where $\Phi_\tau = \{a \in \Phi | T(0,a) > \tau\}$. The distribution of $(\tau, \mathcal{C})$ under the models (M1) and (M2) is described in detail in [M, R]. For instance, in the case of (M2) with $v = 1$, we have that $\tau^{d+\beta} \sim \Gamma(\beta/(d+\beta), \alpha)$ is gamma-distributed. In particular, it is shown in [M] how arrivals can be generated by extending a main result in [165], whereby typical Poisson-Johnson-Mehl crystals can be simulated. The empirical results exhibited in [M] under (M2) show remarkably different crystals in two and three dimensions. Other simulated results including some non-Poissonian cases and modifications of the definition of the Johnson-Mehl crystals are reported in [58, 117, 88].

4.6 Statistical analysis

So far, research on random tessellations has focused rather on mathematical modelling and analysis than statistics. Statistical aspects of tessellations are discussed below.

Mostly non-parametric statistics and moment based methods have been considered. Viewing a stationary tessellation as a germ-grain model, non-parametric statistics based on general random set methods apply. In particular, it is useful to consider the edges of a planar tessellation as a fiber process, while stereological methods become useful for spatial tessellations (see [I, Section 6] and [180, Sections 10.8 and 11.5.4]). Due to the mean value relations for the geometrical and topological characteristics of stationary random tessellations (see, e.g., (32)–(33)), only a few number of mean values need to be estimated. The number of mean values needed is further reduced in the case of a normal tessellation, or if a parametric class of statistical models is considered. In particular, the mean values may be compared to the special relationships which hold for Poisson-Voronoi and Poisson-Johnson-Mehl tessellations. For Poisson-Voronoi tessellations, the intensity of vertices is the only parameter that needs to be estimated. As mentioned at the end of Section 4.3.1 various summary statistics have been investigated in detail for Poisson-Voronoi tessellations.

As Voronoi tessellations can be updated locally, they play an important role in many statistical applications as a convenient tool for speeding up computations (see, e.g., [170]). Moreover, in a Bayesian context, Voronoi tessellations in combination with conditional auto-regressions have been used in [3, 83, 82] for generating random and smoothed step functions which approximate the intensity function of an inhomogeneous Poisson point process.

Voronoi tessellations have occasionally been proposed for applications in image analysis, see for example [1]. In a current research project [144] we introduce a simple and flexible prior model for Bayesian image analysis and reservoir modelling by using interacting coloured Voronoi cells. This prior model can be viewed as a NNMPP (see Section 2.2) and it provides a simpler parameterization than the triangulation models used in [150], particularly if $d = 3$. It is clearly important to use simple and yet realistic prior models when Bayesian methods are used in reservoir applications to forecast the possibility of future production of oil and gas.

Yet another ongoing research project concerns inference problems in the situation where a spatial point process is observed on the edges (or faces) of an unobserved random tessellation (three-dimensional configurations of pores in translucent alumina provide examples of such situations, see [103]). In [191] we carry out a Bayesian analysis for estimating the parameters of the entire process and the underlying random tessellation.

References

[1] N. Ahuja and B.J. Schachter. *Pattern Models.* Wiley, New York, 1983.

[2] T. Arak, P. Clifford, and D. Surgailis. Point-based polygonal models for random graphs. *Advances of Applied Probability*, 25:348–372, 1993.

[3] E. Arjas and J. Heikkinen. An algorithm for nonparametric Bayesian estimation of a Poisson intensity. *Computational Statistics*, 12:385–402, 1997.

[4] M. Avrami. Kinetics of phase change I. *J. Chem. Phys.*, 7:1103–1112, 1939.

[5] M. Avrami. Kinetics of phase change II. *J. Chem. Phys.*, 8:212–224, 1940.

[6] M. Avrami. Kinetics of phase change III. *J. Chem. Phys.*, 9:117–184, 1941.

[7] A. Baddeley and R. Turner. Practical maximum pseudolikelihood for spatial point patterns. *Australian and new Zealand Journal of Statistics*, 2000. (To appear).

[8] A.J. Baddeley, W.S. Kendall, and M.N.M. van Lieshout. Quermass-interaction processes: Stability properties. *Advances of Applied Probability (SGSA)*, 1999. (To appear).

[9] A.J. Baddeley, J. Møller, and R. Waagepetersen. Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. Technical Report R-98-2019, Department of Mathematics, Aalborg University, 1998. *Statistica Neerlandica.* (To appear).

[10] A.J. Baddeley and M.N.M. van Lieshout. Stochastic geometry models in high-level vision. In K.V. Mardia and G.K. Kanji, editors, *Statistics and Images, Advances in Applied Statistics, a supplement to the Journal of Applied Statistics*, volume 20, chapter 11, pages 235–256. Carfax publishing, Abingdon, 1993.

[11] A.J. Baddeley and M.N.M. van Lieshout. Area-interaction point processes. *Annals of the Institute of Statistical Mathematics*, 46:601–619, 1995.

[12] O.E. Barndorff-Nielsen, M.M.N. van Lieshout, and W.S. Kendall, editors. *Stochastic Geometry: Likelihood and Computation*, Monographs on Statistics and Applied Probability, Boca Raton, 1999. Chapman and Hall/CRC, London.

[13] M. Berman and T.R. Turner. Approximating point process likelihoods with GLIM. *Applied Statistics*, 41:31–38, 1992.

[14] E. Bertin, J.-M. Billiot, and R. Drouilhet. Spatial Delaunay Gibbs point processes. Manuscript, 1997.

[15] E. Bertin, J.-M. Billiot, and R.D. Labsad. Existence of "nearest-neighbour" type spatial Gibbs models. Manuscript, 1997.

[16] J. Besag. Some methods of statistical analysis for spatial data. *Bulletin of the Institute of International Statistics*, 47:77–92, 1977.

[17] J. Besag, P. Green, D. Higdon, and K. Mengersen. Bayesian computation and stochastic systems. *Statistical Science*, 10:3–66, 1995.

[18] J. Besag, R. Milne, and S. Zachary. Point process limits of lattice processes. *Journal of Applied Probability*, 19:210–216, 1982.

[19] J.E. Besag. Spatial interaction and the statistical analysis of lattice systems (with discussion). *Journal of the Royal Statistical Society*, B 36:192–236, 1974.

[20] J.E. Besag. Statistical analysis of non-lattice data. *The Statistician*, 24:179–195, 1975.

[21] J.E. Besag. On the statistical analysis of dirty pictures (with discussion). *Journal of the Royal Statistical Society*, B 48:259–302, 1986.

[22] J.E. Besag. Discussion on the paper by Grenander and Miller. *Journal of the Royal Statistical Society*, B 56:591–592, 1994.

[23] J.E. Besag and P.J. Green. Spatial statistics and Bayesian computation (with discussion). *Journal of the Royal Statistical Society*, B 55:25–37, 1993.

[24] A.A. Borokov. *Asymptotic Methods in Queueing Theory.* Wiley, Chichester, 1984.

[25] A.A. Borokov and S.G. Foss. Two ergodicity criteria for stochastically recursive sequences. *Acta Applic. Math.*, 34:125–134, 1994.

[26] K.A. Brakke. Statistics of random plane Voronoi tessellations. Department of Mathematical Sciences, Susquehanna University, Manuscript, 1987.

[27] K.A. Brakke. Statistics of three dimensional random Voronoi tessellations. Department of Mathematical Sciences, Susquehanna University, Manuscript, 1987.

[28] L. Breyer, M.B. Hansen, and J. Møller. Likelihood analysis of random sets models with a view to Diggle's heather-incidence data. (In preparation).

[29] A. Brix and J. Møller. Space-time multitype log Gaussian Cox processes with a view to modelling weed data. Technical Report R-98-2012, Department of Mathematical Sciences, Aalborg University, 1998.

[30] S. Brooks and G. Roberts. Diagnosing convergence of Markov chain Monte Carlo algorithms. Technical report, Statistical Laboratory, University of Cambridge, 1996.

[31] G. Buffon. Essai d'arithmétique morale. *Supplemént à l'Historie naturelle*, 4, 1777.

[32] D.S. Carter and P.M. Prenter. Exponential spaces and counting processes. *Zeitschrift für Wahrsheinlichkeitstheorie und verwandte Gebiete*, 21:1–19, 1972.

[33] J.T. Chayes, L. Chayes, and R. Kotecky. The analysis of the Widom-Rowlinson model by stochastic geometric methods. *Communications in Mathematical Physics*, 172:551–569, 1995.

[34] J.T. Chayes and J. Machta. Graphical representations and cluster algorithms. Part II. (In preparation).

[35] S.N. Chiu, R. van de Weygaert, and D. Stoyan. The sectional Poisson-Voronoi tessellation is not a Voronoi tessellation. *Advances of Applied Probability*, 28:356–376, 1996.

[36] P. Clifford. Markov random fields in statistics. In G.R. Grimmett and D.J.A. Welsh, editors, *Disorder in Physical Systems. A volume in honour of J.M. Hammersley.* Clarendon Press, Oxford, 1990.

[37] P. Clifford and G. Nicholls. Comparison of birth-and-death and Metropolis-Hastings Markov chain Monte Carlo for the Strauss process. Department of Statistics, Oxford University. Manuscript, 1994.

[38] F. Comets. On consistency of a class of estimators for exponential families of Markov random fields on the lattice. *Annals of Statistics*, 20:455–468, 1992.

[39] F. Comets and M. Janžura. A central limit theorem for conditionally centered random fields with an application to Markov fields. Manuscript, 1997.

[40] T.H. Cormen, C.E. Leiserson, and R.L. Rivest. *Introduction to Algorithms.* McGraw-Hill, New York, 1989.

[41] M.K. Cowles and B.P. Carlin. Markov chain Monte Carlo convergence diagnostics: A comparative review. Division of Biostatistics, University of Minnesota. Manuscript, 1995.

[42] D.R. Cox. Partial likelihood. *Biometrika*, 62:269–276, 1975.

[43] N.A.C. Cressie. *Statistics for Spatial Data.* Wiley, New York, revised edition, 1993.

[44] M. Creutz. Confinement and the critical dimensionality of space-time. *Physical Review Letters*, 43:553–556, 1979.

[45] D.J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes*. Springer-Verlag, New York, 1988.

[46] P.J. Diggle. *Statistical Analysis of Spatial Point Patterns.* Academic Press, London, 1983.

[47] P.J. Diggle. A kernel method for smoothing point process data. *Applied Statistics*, 34:138–147, 1985.

[48] P.J. Diggle, T. Fiksel, P. Grabarnik, Y. Ogata, D. Stoyan, and M. Tanemura. On parameter estimation for pairwise interaction point processes. *International Statistical Review*, 62:99–117, 1994.

[49] G.L. Dirichlet. Über die Reduktion der positiven quadratischen Formen mit drei unbestimmten ganzen Zahlen. *J. Reine und Angew. Math.*, 40:209–227, 1850.

[50] R.G. Edwards and A.D. Sokal. Generalization of the Fortuin-Kasteleyn-Swendsen-Wang representation in Monte Carlo. *Physical Review*, 38:2009–2012, 1988.

[51] U.R. Evans. The laws of expanding circles and spheres in relation to the lateral growth of surface films and the grain size of metals. *Trans. Faraday Soc.*, 41:365–374, 1945.

[52] T. Fiksel. Estimation of parameterized pair potentials of marked and nonmarked Gibbsian point processes. *Elektronische Informationsverarbeitung und Kypernetik*, 20:270–278, 1984.

[53] J.A. Fill. An interruptible algorithm for perfect sampling via Markov chains. *Annals of Applied Probability*, 8:131–162, 1998.

[54] G.S. Fishman. An analysis of Swendsen-Wang and related sampling methods. Technical Report UNC/OR TR-96/04, Department of Operations Research, University of North Carolina, 1996.

[55] C.M. Fortuin and P.W. Kasteleyn. On the random cluster model. I. Introduction and relation to other models. *Physica*, 57:536–564, 1972.

[56] S.G. Foss. On ergodicity conditions in multi-server queues. *Siberian Math. J.*, 24:168–175, 1983.

[57] S.G. Foss and R.L. Tweedie. Perfect simulation and backward coupling. *Stochastic Models*, 1998. (To appear).

[58] H.J. Frost and C.V. Thompson. The effect of nucleation conditions on the topology and geometry of two-dimensional grain structures. *Acta Metall.*, 35:529–540, 1987.

[59] D.J. Gates and M. Westcott. Clustering estimates for spatial point distributions with unstable potentials. *Annals of the Institute of Statistical Mathematics*, A 38:123–135, 1986.

[60] A.E. Gelfand and A.F.M. Smith. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85:398–409, 1990.

[61] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.

[62] S. Geman and C. Graffigne. Markov random field models and their applications to computer vision. In A.M. Gleason, editor, *Proceedings of the International Congress of Mathematicians*, volume II, pages 1496–1517. American Mathematical Society, Providence, R.I., 1986.

[63] H.-O. Georgii. *Gibbs Measures and Phase Transition.* Walter de Gruyter, Berlin, 1988.

[64] H.-O. Georgii and O. Haggström. Phase transition in continuum Potts models. *Communications in Mathematical Physics*, 181:507–528, 1996.

[65] C.J. Geyer. Markov chain Monte Carlo maximum likelihood. In *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*, pages 156–163, 1991.

[66] C.J. Geyer. On the convergence of Monte Carlo maximum likelihood calculations. *Journal of Royal Society of Statistics*, B 56:261–274, 1994.

[67] C.J. Geyer. Likelihood inference for spatial point processes. In O.E. Barndorff-Nielsen, W.S. Kendall, and M.N.M. van Lieshout, editors, *Stochastic Geometry: Likelihood and Computation*, pages 79–140, Boca Raton, 1999. Chapman and Hall/CRC, London.

[68] C.J. Geyer and E.A. Thompson. Constrained Monte Carlo maximum likelihood for dependent data. *Journal of the Royal Society of Statistics*, B 54:657–699, 1992.

[69] C.J. Geyer and E.A. Thompson. Annealing Markov chain Monte Carlo with applications to pedigree analysis. *Journal of American Statistical Association*, 90:909–920, 1995.

[70] B. Gidas. Consistency of maximum likelihood and pseudo-likelihood estimators for Gibbs distributions. In W. Fleming and P.L. Lions, editors, *Stochastic Differential Systems, Stochastic Control Theory and Applications. IMA Volumes in Mathematics and Its Applications*, volume 10, pages 129–145. Springer, New York, 1988.

[71] E.N. Gilbert. Random subdivisions of space into crystals. *Annals of Mathematical Statistics*, 33:958–972, 1962.

[72] W.R. Gilks, S. Richardson, and D.J. Spiegelhalter. *Markov Chain Monte Carlo in Practice*. Chapman and Hall, London, 1996.

[73] M. Goulard, A. Särkkä, and P. Grabarnik. Parameter estimation for marked Gibbs point processes through the maximum pseudo-likelihood method. *Scandinavian Journal of Statistics*, 23:365–379, 1996.

[74] P. Green. Reversible jump MCMC computation and Bayesian model determination. *Biometrika*, 82:711–732, 1995.

[75] P.J. Green. MCMC in image analysis. In W.R. Gilks, S. Richardson, and D.J. Spiegelhalter, editors, *Markov chain Monte Carlo in Practice*, pages 381–399. Chapman and Hall, London, 1996.

[76] U. Grenander and M.I. Miller. Representation of knowledge in complex systems (with discussion). *Journal of the Royal Statistical Society*, B 56:549–603, 1994.

[77] G. Grimmett. The stochastic random-cluster process and the uniqueness of random-cluster measures. *Annals of Probability*, 23:1461–1510, 1995.

[78] X. Guyon. Estimation d'un champ par pseudo-vraisemblance conditionelle: Etude asymptotique et application au cas Markovian. In F. Droesbeke, editor, *Spatial Processes and Spatial Time Series Analysis (Proceedings of the Sixth Franco-Belgian Meeting of Statisticians, 1985)*, pages 15–62. Publications des Facultes Universitaires Saint-Louis, Brussels, 1987.

[79] X. Guyon. *Random Fields on a Network*. Springer-Verlag, New York, 1995.

[80] O. Häggstöm and K. Nelander. Exact sampling from anti-monotone systems. *Statistica Neerlandia*, 52:360–380, 1998.

[81] W. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57:97–109, 1970.

[82] J. Heikkinen and E. Arjas. Nonparametric Bayesian estimation of a spatial Poisson intensity using a partially observed concomitant variable. Manuscript, 1997. (Submitted for publication).

[83] J. Heikkinen and E. Arjas. Non-parametric Bayesian estimation of a spatial Poisson intensity. *Scandinavian Journal of Statistics*, 25:435–450, 1998.

[84] J. Heikkinen and H. Högmander. Fully Bayesian approach to image restoration with an application in biogeography. *Applied Statistics*, 43:569–582, 1994.

[85] J. Heikkinen and A. Penttinen. Bayesian smoothing in the estimation of the pair potential function of Gibbs point processes. Manuscript, 1997. (Submitted for publication).

[86] L. Heinrich, R. Körner, N. Mehlhorn, and L. Muche. Numerical and analytical computation of some second-order characteristics of spatial Poisson-Voronoi tessellations. *Statistics*, 1998. (To appear).

[87] L. Heinrich and L. Muche. On the pair-correlation function of planar Voronoi mosaics. Preprint 94-07, Fachbereich Mathematik, Technische Universität Bergakademie Freiberg, 1994.

[88] L. Heinrich and E. Schüle. Generation of the typical cell of a non-Poissonian Johnson-Mehl tessellation. *Stochastic Models*, 11:541–560, 1995.

[89] H. Hermann, H. Wendroch, and D. Stoyan. Cell-area distributions of planar Voronoi mosaics. *Metallography*, 23:189–200, 1989.

[90] A.L. Hinde and R.E. Miles. Monte Carlo estimates of the distribution of the random polygons of the Voronoi tessellation with respect to a Poisson process. *J. Statist. Comput. Simul.*, 10:205–223, 1980.

[91] V. Horálek. A note on the time-homogeneous Johnson-Mehl tessellation. *Advances of Applied Probability*, 20:684–685, 1988.

[92] V. Horálek. ASTM grain-size model and related random tessellation models. *Materials Characterization*, 25:263–284, 1990.

[93] F. Huang and Y. Ogata. Improvements of the maximum pseudolikelihood estimators in various spatial statistical models. Research Memorandum 645, The Institute of Statistical Mathematics, Tokyo, 1997.

[94] E. Ising. Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift für Physik*, 31:253–258, 1925.

[95] E.B. Jensen and J. Møller. Stereological versions of integral geometric formulae for n-dimensional ellipsoids. *Journal of Applied Probability*, 23:1031–1037, 1986.

[96] E.B. Vedel Jensen. *Local Stereology.* World Scientific, Singapore, 1998.

[97] J.L. Jensen. Efficiency of the pseudolikelihood estimate in a one-dimensional lattice gas. Research Report 352, Department of Theoretical statistics, University of Aarhus, 1996.

[98] J.L. Jensen. A note on asymptotic normality in the thermodynamic limit at low densities. *Advances of Applied Mathematics*, 12:387–399, 1991.

[99] J.L. Jensen. Asymptotic normality of estimates in spatial point processes. *Scandinavian Journal of Statistics*, 20:97–109, 1993.

[100] J.L. Jensen and H.R. Künsch. On asymptotic normality of pseudo likelihood estimates for pairwise interaction processes. *Annals of the Institute of Statistical Mathematics*, 46:475–486, 1994.

[101] W.A. Johnson and R.F. Mehl. Reaction kinetics in processes of nucleation and growth. *Trans. Amer. Inst. Min. Engrs.*, 135:416–458, 1939.

[102] O. Kallenberg. An informal guide to the theory of conditioning in point processes. *International Statistical Review*, 52:151–164, 1984.

[103] L.M. Karlsson and A. Liljeborg. Second-order stereology for pores in translucent alumina studied by confocal scanning laser microscopy. *Journal of Microscopy*, 175:186–194, 1994.

[104] R.L. Kashyap and R. Chellappa. Estimation and choice of neighbors in spatial-interaction models of images. *IEEE Transactions on Information Theory*, IT-29:60–72 (Corrigendum p. 629), 1983.

[105] F.P. Kelly and B.D. Ripley. A note on Strauss' model for clustering. *Biometrika*, 63:357–360, 1976.

[106] D.G. Kendall. A survey of the statistical theory of shape (with discussion). *Statistical Science*, 4:87–120, 1989.

[107] W.S. Kendall. A spatial Markov property for nearest-neighbour Markov point processes. *Journal of Applied Probability*, 28:767–778, 1990.

[108] W.S. Kendall. Perfect simulation for the area-interaction point process. In L. Accardi and C.C. Heyde, editors, *Probability Towards 2000*, pages 218–234. Springer, 1998.

[109] W.S. Kendall and J. Møller. Perfect Metropolis-Hastings simulation of locally stable point processes. Technical Report R-99-2001, Department of Mathematical Sciences, Aalborg University, 1999.

[110] R. Kindermann and J.L. Snell. *Markov Random Fields and their Applications*. American Mathematical Society, Providence, Rhode Island, 1980.

[111] W. Klein. Potts-model formulation of continuum percolation. *Physical Review*, B 26:2677–2678, 1982.

[112] A.N. Kolmogorov. Statistical theory of crystallization of metals. *Bull. Acad. Sci. USSR Mat. Ser.*, 1:355–359, 1937.

[113] H.R. Künsch, S. Geman, and A. Kehagias. Hidden Markov random fields. *Annals of Applied Probability*, 5:577–602, 1995.

[114] U. Lorz. Cell-area distributions of planar sections of spatial Voronoi tessellations. *Materials Char.*, 3:297–311, 1990.

[115] U. Lorz and U. Hahn. Geometric characteristics of spatial Voronoi tessellations and planar sections. Preprint 93-05, Fachbereich Mathematik, Technische Universität Bergakademie Freiberg, 1993.

[116] H.W. Lotwick and B.W. Silverman. Convergence of spatial birth-and-death processes. *Math. Proc. Cambridge Philos. Soc.*, 90:155–165, 1981.

[117] K.W. Mahin, K. Hanson, and J.W. Morris. Comparative analysis of the cellular and Johnson-Mehl microstructures through computer simulation. *Acta Metall.*, 28:443–453, 1980.

[118] E. Marinari and G. Parisi. Simulated tempering: A new Monte Carlo scheme. *Europhysics Letters*, 19:451–458, 1992.

[119] S. Mase. Marked Gibbs processes and asymptotic normality of maximum pseudo-likelihood estimators. Manuscript, 1997.

[120] S. Mase. Locally asymptotic normality of Gibbs models on a lattice. *Advances in Applied Probability*, 16:585–602, 1984.

[121] S. Mase. Asymptotic equivalence of grand canonical MLE and canonical MLE of pair potential functions of Gibbsian point process model. Technical Report 292, Statistical Research Group, Hiroshima University, 1991.

[122] S. Mase. Uniform LAN condition of planar Gibbsian point processes and optimality of maximum likelihood estimators of soft-core potential functions. *Probability Theory and Related Fields*, 92:51–67, 1992.

[123] S. Mase. Consistency of the maximum pseudo-likelihood estimator of continuous state space Gibbs processes. *Annals of Applied Probability*, 5:603–612, 1995.

[124] S. Mase, J. Møller, D. Stoyan, R.P. Waagepetersen, and G. Döge. Packing densities and simulated tempering for hard core Gibbs point processes. Technical Report R-99-2002, Department of Mathematical Sciences, Aalborg University, 1999.

[125] B. Matérn. *Spatial Variation.* Lecture Notes in Statistics. Springer-Verlag, Berlin, 1986.

[126] G. Matheron. *Random Sets and Integral Geometry.* Wiley, New York, 1975.

[127] J. Mecke. Stationäre zufällige Maße auf lokalkompakten Abelschen Gruppen. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 9:36–58, 1967.

[128] J. Mecke. Palm methods for stationary random mosaics. In R.V. Ambartzumian, editor, *Combinatorial Principles in Stochastic Geometry*, pages 124–132. Armenian Academy of Science Publishing House, Erevan, 1980.

[129] J. Mecke. Parametric representations of mean values for stationary random mosaics. *Math. Operationsf. Statist., Ser. Statist.*, 15:437–442, 1984.

[130] J. Mecke, R.G. Schneider, D. Stoyan, and W.R.R. Weil. *Stochastische Geometrie.* Birkhäuser Verlag, Basel, 1990.

[131] J.L. Meijering. Interface area, edge length, and number of vertices in crystal aggregates with random nucleation. *Philips Research Reports*, 8:270–290, 1953.

[132] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, 21:1087–1092, 1953.

[133] S.P. Meyn and R.L. Tweedie. *Markov Chains and Stochastic Stability.* Springer-Verlag, London, 1993.

[134] R.E. Miles. On the homogeneous planar Poisson point process. *Mathematical Biosciences*, 6:85–127, 1970.

[135] R.E. Miles. The random subdivision of space. *Supplement to Advances of Applied Probability*, pages 243–266, 1972.

[136] R.E. Miles. A synopsis of "Poisson Flats in Euclidean Spaces". In E.F. Harding and D.G. Kendall, editors, *Stochastic Geometry*, pages 202–227. Wiley, London, 1974.

[137] R.E. Miles. Sectional Voronoi tessellations. *Rev. Unión Math. Argentina*, 29:310–327, 1984.

[138] R.E. Miles and R.L. Maillardet. The basic structure of Voronoi and generalized Voronoi polygons. *Journal of Applied Probability*, 19A:97–112, 1982.

[139] A. Mira, J. Møller, and G.O. Roberts. Perfect slice samplers. (In preparation).

[140] J. Møller. Stereological analysis of particles of varying ellipsoidal shape. *Journal of Applied Probability*, 25:322–335, 1988.

[141] J. Møller, E.B. Jensen, J.S. Petersen, and H.J.G. Gundersen. Modelling an aggregate of space-filling cells from sectional data. Technical Report 182, Department of Theoretical Statistics, Aarhus University, 1989.

[142] J. Møller and G. Nicholls. Perfect simulation for sample-based inference. Technical Report R-99-2011, Department of Mathematical Sciences, Aalborg University, 1999.

[143] J. Møller and K. Schladitz. Extensions of Fill's algorithm for perfect simulation. *Journal of the Royal Statistical Society*, B 61:955–969, 1999.

[144] J. Møller and Ø. Skare. Bayesian image analysis with coloured voronoi tessellations and a view to applications in reservoir modelling. (In preparation).

[145] R.A. Moyeed and A.J. Baddeley. Stochastic approximation of the MLE for a spatial point pattern. *Scandinavian Journal of Statistics*, 18:39–50, 1991.

[146] L. Muche. Distributional properties of the Poisson-Voronoi tessellation. *Acta Stereologica*, 12:125–130, 1993.

[147] L. Muche and D. Stoyan. Contact and chord length distributions of the Poisson Voronoi tessellation. *Advances of Applied Probability*, 29:467–471, 1992.

[148] D.J. Murdoch and P.J. Green. Exact sampling from a continuous state space. *Scandinavian Journal of Statistics*, 25:483–502, 1998.

[149] X.X. Nguyen and H. Zessin. Integral and differential characterizations of Gibbs processes. *Mathematische Nachrichten*, 88:105–115, 1979.

[150] G. Nicholls. Bayesian image analysis with Markov chain Monte Carlo and colored continuum triangulation mosaics. *Journal of the Royal Statistical Society*, B 60:643–659, 1998.

[151] Y. Ogata and M. Tanemura. Estimation of interaction potentials of spatial point patterns through the maximum likelihood procedure. *Annals of the Institute of Statistical Mathematics*, B 33:315–338, 1981.

[152] Y. Ogata and M. Tanemura. Likelihood analysis of spatial point patterns. *Journal of the Royal Statistical Society*, B 46:496–518, 1984.

[153] Y. Ogata and M. Tanemura. Estimation of interaction potentials of marked spatial point patterns through the maximum likelihood method. *Biometrics*, 41:421–433, 1985.

[154] Y. Ogata and M. Tanemura. Likelihood estimation of soft-core interaction potentials for Gibbsian point patterns. *Annals of the Institute of Statistical Mathematics*, 41:583–600, 1989.

[155] A. Okabe, B. Boots, and K. Sugihara. *Spatial Tessellations. Concepts and Applications of Voronoi Diagrams.* Wiley, Chichester, 1992.

[156] A. Penttinen. *Modelling Interaction in Spatial Point Patterns: Parameter Estimation by the Maximum Likelihood Method*. Number 7 in Jyväskylä Studies in Computer Science, Economics, and Statistics, 1984.

[157] A. Penttinen, D. Stoyan, and H.M. Henttonen. Marked point processes in forest statistics. *Forest Science*, 38:806–824, 1992.

[158] D.K. Pickard. Inference for discrete Markov fields: The simplest nontrivial case. *Journal of the American Statistical Association*, 82:90–96, 1987.

[159] A. Pievatolo and P.J. Green. Boundary detection through dynamic polygons. *Journal of Royal Statistical Society*, B 60:609–626, 19.

[160] A. Possolo. Estimation of binary Markov random fields. Technical Report 77, Department of Statistics, University of Washington, 1986.

[161] C. Preston. *Random Fields*. Lecture Notes in Mathematics, 534. Springer-Verlag, Berlin-Heidelberg, 1976.

[162] C.J. Preston. Spatial birth-and-death processes. *Bulletin of the International Statistical Institute*, 46:371–391, 1977.

[163] J.G. Propp and D.B. Wilson. Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures and Algorithms*, 9:223–252, 1996.

[164] J.G. Propp and D.B. Wilson. How to get a perfectly random sample from a generic Markov chain and generate a random spanning tree to a directed graph. *Journal of Algorithms*, 27:170–217, 1998.

[165] M.P. Quine and D.F. Watson. Radial simulation of n-dimensional Poisson processes. *Journal of Applied Probability*, 21:548–557, 1984.

[166] W. Radecke. Some mean-value relations on stationary random mosaics in the space. *Mathematische Nachrichten*, 97:203–210, 1980.

[167] P.N. Rathie. On the volume of the typical Poisson-Delaunay cell. *Journal of Applied Probability*, 29:740–744, 1992.

[168] B.D. Ripley. Modelling spatial patterns (with discussion). *Journal of the Royal Statistical Society*, B 39:172–212, 1977.

[169] B.D. Ripley. Simulating spatial patterns: dependent samples from a multivariate density. Algorithm AS 137. *Applied Statistics*, 28:109–112, 1979.

[170] B.D. Ripley. *Spatial Statistics*. Wiley, New York, 1981.

[171] B.D. Ripley. *Statistical Inference for Spatial Processes*. Cambridge University Press, Cambridge, 1988.

[172] B.D. Ripley and F.P. Kelly. Markov point processes. *Journal of the London Mathematical Society*, 15:188–192, 1977.

[173] G.O. Roberts and J.S. Rosenthal. Geometric ergodicity and hybrid Markov chains. *Electronic Communications in Probability*, 2:13–25, 1997.

[174] G.O. Roberts and R.L. Tweedie. Exponential convergence of Langevin diffusions and their discrete approximations. *Bernoulli*, 2:341–363, 1996.

[175] H. Rue and M.A. Hurn. Bayesian object identification. *Biometrika*, 86:649–660, 1999.

[176] D. Ruelle. *Statistical Mechanics: Rigorous Results*. W.A. Benjamin, Reading, Massachusetts, 1969.

[177] D. Ruelle. Existence of a phase transition in a continuous classsical system. *Physical Review Letters*, 27:1040–1041, 1971.

[178] T.H. Scheike. Anisotropic growth of Voronoi cells. *Advances of Applied Probability*, 26:43–53, 1994.

[179] A.D. Sokal. Monte Carlo methods in statistical mechanics: foundations and new algorithms. In *Cours de Troisième Cycle de la Physique en Suisse Romande*, Lausanne, 1989.

[180] D. Stoyan, W.S. Kendall, and J. Mecke. *Stochastic Geometry and its Applications*. Wiley, Chichester, second edition, 1995.

[181] D. Stoyan and H. Stoyan. *Fractals, Random Shapes and point Fields*. Wiley, Chichester, 1994.

[182] D. Strauss. Analyzing binary lattice data with the nearest-neighbor property. *Journal of Applied Probability*, 12:702–712, 1975.

[183] D. Strauss. A general class of models for interaction. *SIAM Review*, 28:513–527, 1986.

[184] D.J. Strauss. A model for clustering. *Biometrika*, 63:467–475, 1975.

[185] D.J. Strauss and M. Ikeda. Pseudolikelihood estimation for social networks. *Journal of the American Statistical Association*, 85:204–212, 1990.

[186] Student. On the error of counting with a haemacytometer. *Biometrika*, 5:351–360, 1907.

[187] P. Suomela. *Construction of Nearest Neighbour Systems*. Disertation 10, Department of Mathematics, University of Helsinki, 1976.

[188] M. Sweeny. Monte Carlo study of weighted percolation clusters relevant to the Potts models. *Physical Review*, B 27:4445–4455, 1983.

[189] R.H. Swendsen and J.-S. Wang. Nonuniversal critical dynamics in Monte Carlo simulations. *Physical Review Letters*, 58:86–88, 1987.

[190] R. Tacacs. Estimator for the pair-potential of a Gibbsian point process. *Statistics*, 17:429–433, 1986.

[191] M. Tanemura, E.B. Vedel Jensen, and J. Møller. Random tessellation networks and some inference problems. (In preparation).

[192] E. Thönnes. Perfect simulation of some point processes for the impatient user. *Advances in Applied Probability*, 31:69–87, 1999.

[193] H. Tjelmeland and J. Besag. Markov random fields with higher-order interactions. *Scandinavian Journal of Statistics*, 25:415–433, 1998.

[194] R. van de Weygaert. Fragmenting the Universe III. The construction and statistics of 3-d Voronoi tessellations. *Astron. Astrophys.*, 283:361–406, 1994.

[195] G. Voronoi. Nouvelles applications des parametres continus a la theorie des formes quadratiques. *J. Reine und Angew. Math.*, 134:198–287, 1908.

[196] R. Waagepetersen. Phase transition and simulation for a penalized Ising model with applications in image analysis. Research Report 381, Department of Theoretical Statistics, University of Aarhus, 1997.

[197] B. Widom and J.S. Rowlinson. A new model for the study of liquid-vapor phase transitions. *Journal of Chemical Physics*, 52:1670–1684, 1970.

[198] A.T.A. Wood and G. Chan. Simulation of stationary Gaussian processes in $[0,1]^d$. *Journal of Computational and Graphical Statistics*, 3:409–432, 1994.

[199] L. Younes. Estimation and annealing for Gibbsian fields. *Annales de l'Institut Henri Poincaré. Probabilitiés et Statistiques*, 24:269–294, 1988.

[200] M. Zähle. Random cell complexes and generalized sets. *Annals of Probability*, 16:1742–1766, 1988.

[201] S.A. Zuyev. Estimates of the Voronoi polygon's geometric characteristics. *Random Structures and Algorithms*, 3:149–162, 1992.

Dansk resumé

Rumlige data, såsom digitale billeder, plane punktmønstre og mere generelle to- eller højere-dimensionale rumlige strukturer, hvor en eventuel tidsfaktor er inddraget, dukker op i mange forskellige sammenhænge. Eksempler herpå omfatter landbrugsforsøg, forstvæsen, fysik, geologi, geografi, meterologi, biologi, medicin og adskillige teknologiske anvendelser af billedanalyseapparatur. Rumlig statistik omhandler statistisk modellering og metoder til at drage inferens om rumlige data. Den statistiske analyse af rumlige data kompliceres af de enorme datamængder og komplekse statistiske modeller, der typisk finder anvendelse i praksis.

Stokastisk geometri omhandler — i traditionel forstand — stokastisk modellering og analyse af tilfældige delmængder i et euklidisk rum. Imidlertid har en udvikling inden for stokastisk geometri fundet sted, så det i dag er vanskeligt at skelne skarpt mellem rumlig statistik og stokastisk geometri; se eksempelvis [180].

Igennem de sidste godt tyve år har en rivende udvikling fundet sted med hensyn til statistiske anvendelser og mere grundforskningsorienterede aspekter af rumlig statistik og stokastisk geometri. Mange ideer med oprindelse i statistisk fysik, herunder rumlige Markov modeller og Markov kæde Monte Carlo metoder, samt adgangen til stadig hurtigere og billigere computere har haft og har stadig en afgørende betydning for denne udvikling.

Nærværende afhandling omhandler visse aspekter af den nyeste forskning inden for rumlig statistik, stokastisk geometri og Markov kæde Monte Carlo metoder. Den består af tre dele:

(i) *statistiske modeller for lattice data, punktmønstre og rumlige fødsels-dødsprocesser*;
(ii) *Markov kæde Monte Carlo metoder og statistisk inferens for rumlige modeller*;
(iii) *stokastiske tessellationer*, der udgør et hovedemne inden for stokastisk geometri (en tessellation er løst sagt en opsplitning af et rum i disjunkte områder).

Afhandlingen er baseret på en monografi [L] og tyve udvalgte artikler [A–K, M–V] (se nogle af de første sider i afhandlingen). Disse bidrag kan naturligt inddeles efter hvilken type af rumlig proces, der hovedsagligt studeres:

stokastiske felter/latticeprocesser:	[E, N, U];
punktprocesser:	[A, B, C, D, F, K, P, T];
rumlige fødsels-dødsprocesser:	[H, S];
stokastiske tessellationer:	[G, I, J, L, M, O, R, V].

(i) Rumlig variation i lattice- og punktprocesser

Der findes forskellige typer af modeller for rumlig variation. En type er Markov modeller, hvor betinget uafhængighed muliggør specifikation og analyse af komplicerede statistiske modeller ved hjælp af lokale beskrivelser i rum og (eventuelt) tid. En anden type er dobbeltstokastiske processer, hvor den rumlige proces fremkommer ved at betinge med en underliggende intensitetsproces, som beskriver den rumlige heterogenitet.

Anvendelser af Markov felter i statistik har især været forbundet med bayesiansk billedbehandling af 'støjfyldte' digitale billeder. I [E] diskuteres, hvorledes disse modeller og metoder finder anvendelse til estimation i atlasundersøgelser af fuglearters udbredelse.

I mange billedbehandlingsanvendelser er der en bestemt struktur i det 'sande' billede og en realistisk specifikation af priorfordelingen er påkrævet. I [U] introduceres en fleksibel klasse af modeller kaldet Markov sammenhængskomponent felter. Ved hjælp af disse nye modeller kan den globale struktur i et billede modelleres mere direkte end ved brug af de klassiske Markov felter. Det vises i [U], at Markov sammenhængskomponent felter besidder en række nyttige markovegenskaber, at de kan udvides til et uendeligt lattice og at de har mangfoldige anvendelsesmuligheder.

I [A] indføres en generaliseret markovegenskab for rumlige punktprocesser, således at interaktionen mellem punkterne kan afhænge af realisationen af processen. Dette leder til en ny klasse af Markov punktprocesser, hvor en sætning af Hammersley-Clifford type kan etableres. Modellerne omfatter mange interessante eksempler, herunder de klassiske Ripley-Kelly Markov punktprocesser, der som bekendt udgør en rig klasse af modeller for frastødning mellem punkter (eller mere generelt geometriske objekter). Som diskuteret i [P], hvad angår tiltrækning mellem punkter/objekter, så synes det modsatte mere at være tilfældet — faktisk viser de fleste studerede eksempler i litteraturen på Ripley-Kelly Markov punktprocesser med tiltrækning en ekstrem grad af klusterdannelse.

I [B, K, P] diskuteres, hvordan realistiske Markov kluster punktprocesser kan konstrueres. Det vises i [B], at kluster punktprocesser med uniformt begrænset klusterdiameter i almindelighed vil være Baddeley-Møller Markov [A], men ikke Ripley-Kelly Markov. I [D, K, P] introduceres et karakteriseringsresultat for den såkaldte 'continuum random cluster process', der er endnu et interessant eksempel på en model, som kun er Baddeley-Møller Markov. Et tilsvarende karakteriseringsresultat vedrørende 'area interaction' punktprocessen etableres i [D]. Selvom denne model udgør en undtagelse som en Ripley-Kelly Markov punktproces med såvel frastødning som tiltrækning alt afhængig af, hvordan interaktionsparameteren specificeres, så er det i en vis forstand mere naturligt at betragte 'area interaction' punktprocessen som en Baddeley-Møller Markov punktproces.

En anden klasse af modeller for kluster punktprocesser er Cox processer, der fremkommer ved først at betinge med en stokastisk intensitetsproces og dernæst betragte en Poisson punktproces med intensitetsfunktion lig med intensitetsprocessen. Hvis intensitetsprocessen er en exponentieret Gaussproces, fås klassen af log-gaussiske Cox-processer som introduceret i [T]. Disse udgør en flexibel og anvendelig modelklasse med en række nydelige matematiske egenskaber, som muliggører en detaljeret statistisk analyse (herom senere). Eksempelvis er en log-gaussisk Cox-proces identificeret ved dens middelintensitet og parvise korrelationsfunktion.

Rumlige fødsels-dødsprocesser som introduceret af Preston i 1977 har især fundet anvendelse i forbindelse med simulation af rumlige punktprocesser. En rumlig fødsels-dødsproces er en tidshomogen Markov proces med kontinuert tid samt overgange, der enten består i tilføjelse af et nyt punkt eller fjernelse af et eksisterende punkt. Det demonstreres i [S], at rumlige fødsels-dødsprocesser også finder anvendelse til beskrivelse af visse dynamiske rumlige fænomener. Specielt præsenteres i [S] en flexibel parametrisk klasse af rumlige fødsels-dødsprocesmodeller på linien med henblik på at beskrive udviklingen af lineære sandklitter i Kalahari Ørknen.

(ii) Monte Carlo metoder og statistisk inferens for lattice- og punktprocesser

Langt tidligere end i andre delområder af matematisk statistik, faktisk siden midten af 70-erne, har Markov kæde Monte Carlo (MCMC) metoder fundet hyppig anvendelse i rumlig statistik. MCMC metoder er særlig relevant for lattice- og punktprocesser, hvor en eksplicit analyse sjældent er mulig. I det følgende vil jeg fortrinsvis diskutere stokastisk simulation og inferens for rumlige punktprocesser.

Den traditionelle simulationsteknik for Markov punktprocesser er baseret på generering af en rumlig fødsels-dødsproces, som konvergerer mod en given punktproces model. Det vises i [H], at konvergensraten er geometrisk under visse betingelser, som er lette at checke i de fleste konkrete situationer.

En generel simulationsteknik er baseret på Metropolis-Hastingsalgoritmen. I [C] udledes en række algoritmer af denne type for simulation af punktprocesser. Specielt studeres en af disse algoritmer i detaljer i forbindelse med Strauss processen, hvor geometrisk ergodicitet af Markov kæden er opfyldt (dette resultat er siden udvidet til at dække lokalt stabile punktprocesser). Diskussionsindlægget [K] giver en kort beskrivelse og vurdering af forskellige simulationsteknikker baseret på fødsels-dødsprocesser eller Metropolis-Hastingsalgoritmen. En mere fyldestgørende diskussion findes i [P].

Der findes forskellige statistisk baserede metoder til at sandsynliggøre om en Markov kæde praktisk set er i ligevægt, men der hersker stor usikkerhed om deres anvendelighed i de tilfælde, hvor ligevægtsfordelingen i særlig grad er koncentreret i to eller flere modus. Desuden findes der i visse relativt simple modeller kvantitative resultater, som angiver, hvor tæt på ligevægt en Markov kæde kan betragtes at være på et givent tidspunkt;

men som bemærket i [P] er disse resultater af meget begrænset anvendelighed i forbindelse med punktprocesser. Det er derfor af stor vigtighed, at det nu er både teoretisk og praktisk muligt at modificere Propp-Wilsons 'coupling from the past' (CFTP) procedure for direkte simulation til perfekt simulation af ganske mange rumlige processer. I Propp og Wilsons artikel fra 1996 antages tilstandsrummet at være endeligt og udstyret med en partiel ordning, således at der findes et entydigt minimum og maximum og den anvendte sampler er monoton. I [N] vises det, hvordan CFTP-proceduren også lader sig anvende, hvis tilstandsrummet er tælleligt eller udgør et kontinuum, og hvis Gibbssampleren er enten monoton eller anti-monoton med hensyn til en naturlig partiel ordning, hvor der ikke nødvendigvis findes et maximum. Dermed kan auto-binomial, auto-Poisson, auto-negativbinomial, auto-gamma og visse parvis-differens Gibbsfordelinger simuleres eksakt. I [D] studeres en to-komponent Gibbssampler for Widom og Rowlinsons bivariate punktproces, hvor tilstandsrummet er af varierende dimension og sampleren er monoton med hensyn til en bestemt partiel ordning, men hvor der hverken findes et minimalt eller maximalt element. Ved at udtynde fra uafhængige Poissonprocesser udvikles en hurtig perfekt simulationsalgoritme for Widom-Rowlinson modellen og visse andre beslægtede punktprocesmodeller.

Likelihoodanalyse af parametriske modeller for rumlige processer kompliceres af, at normeringskonstanten sædvanligvis ikke er eksplicit kendt. I [F] diskuteres, hvordan Besags pseudolikelihood for latticeprocesser naturligt lader sig modificere til en pseudolikelihood for rumlige punktprocesser. Eftersom pseudolikelihoodfunktionen udelukkende er specificeret ved de lokale afhængigheder i punktprocessen, er den lettere at håndtere beregningsmæssigt. I [F] etableres konsistens af maximum pseudolikelihood estimatoren for eksponentielle familier af rumlige punktprocesmodeller med endelig interaktionsradius.

Alternativt kan man approximere normeringskonstanten i likelihoodfunktionen ved hjælp af MCMC metoder. Dette studeres nærmere i [C] i forbindelse med en diskussion af approximative betingede og ubetingede maximum likelihood estimatorer for parametrene i en Strauss proces. Det bør bemærkes, at likelihoodfunktionen kun er lokalt approximeret, hvorfor metoden bør anvendes med forsigtighed. Til gengæld er det kendt, at maximum pseudolikelihood estimatoren ikke er så efficient som maximum likelihood estimatoren. Dette understøttes også af forskellige empiriske undersøgelser, hvor maximum likelihood estimatet er approximeret ved MCMC metoder. Desuden kan likelihoodfunktionen approximeres ved MCMC metoder i situationer med manglende data.

Likelihood inferens for rumlige fødsels-dødsprocesser er studeret i [S]. I tilfælde af fuldstændig observation på et begrænset tidsinterval er likelihoodfunktionen eksplicit kendt. I [S] udledes en approximativ likelihood funktion med henblik på at inddrage kanteffekter. Sammenlignet med lattice- og punktprocesser er det langt lettere at gennemføre en detaljeret likelihoodanalyse for rumlige fødsels-dødsprocesser.

Parametriske log-gaussiske Cox-procesmodeller er lettere at analysere end andre kendte parametriske modeller for klusterdannelse i punktprocesser. Det skyldes tildels de simple relationer i [T] mellem produkttæthederne for de reducerede moment mål af punktprocessen og middel- og kovariansfunktionerne i den underliggende gaussiske intensitetsproces. Dermed kan en minimum kontrast metode for parameterestimation og forskellige nye teknikker til modelkontrol udledes, som nærmere beskrevet i [T]. Men mindst ligeså afgørende er det, at der findes meget effektive simulationsteknikker. Intensitetsprocessen kan diskretiseres, hvorved diskrete Fourier transformationer muliggør en hurtig simulation af først den gaussiske proces og dernæst punktprocessen. I praksis er punktmønsteret observeret i et begrænset vindue, mens intensitetsprocessen er uobserveret. I [T] vises det, hvordan en Metropolis-justeret Langevin algoritme kan anvendes til betinget simulation af henholdsvis intensitetsprocessen og den uobserverede del af punktprocessen, når der betinges med det observerede punktmønster. Algoritmen er af Metroplis-Hastings type med en effektiv opdateringsmekanisme, der er konstrueret i overensstemmelse med en diskretisering af en Langevin diffusionsproces. Dermed kan intensitetsprocessen og den uobserverede del af punktprocessen let predikteres.

(iii) Stokastiske tessellationer

Stokastiske tessellationer er et hovedemne inden for stokastisk geometri. Min forskning af stokastiske tessellationer har koncentreret sig om stokastiske inddellinger af det d-dimensionale euklidiske rum i disjunkte d-dimensionale delområder kaldet celler i det følgende. Eksempler herpå er celleaggregater i biologi, krystalinske strukturer og geografiske regioner. Jeg har specielt studeret Poissonmodeller for Voronoi og Johnson-Mehl tessellationer, der udgør to af de vigtigste modelklasser.

Kort fortalt er en Voronoi tessellation frembragt af en rumlig punktproces, hvor punkterne kaldes kerner, og hver kerne frembringer en Voronoi celle, der består af de punkter i det d-dimesionale euklidiske rum, som ligger tættest på den pågældende kerne. Dette kan også anskues som resultatet af en vækstproces, hvor hver kerne vokser med samme hastighed i enhver retning indtil væksten stoppes af en anden voksende kerne. Voronoicellerne er derfor konvekse. Hvis kernerne derimod ankommer til forskellige tidspunkter, men ellers vokser efter samme mekanisme som i Voronoi tessellationen, så fremkommer en Johnson-Mehl tessellation, hvor cellerne kan være ikke konvekse.

De fleste resultater for stokastiske tessellationer findes spredt i litteraturen, og for det meste etableres resultaterne for specielle modeller såsom Poisson-Voronoi og Poisson-hyperplan tessellationer. Det omfangsrige oversigtsarbejde [I] giver en samlet fremstilling af teorien for generelle stationære stokastiske tessellationer med konvekse celler. Ved hjælp af teorien for Palm mål etableres i [I] adskillige middelværdirelationer mellem forskellige geometriske karakteristika ved en stokastisk tessellation og det diskuteres, hvorledes relationerne kan benyttes til estimation af bestemte karakteristika udfra målinger af en tessellation observeret i et begrænset vindue. Desuden anskues stokastiske tessellationer udfra en stereologisk synsvinkel, idet der etableres en forbindelse mellem karakteristika for en stokastisk tessellation i det d-dimesionale rum og karakteristika ved tessellationen snittet med et lavere-dimensionalt affint underrum. Endelig gives der en oversigt over kendte resultater for Poisson-Voronoi tessellationen og dens duale Poisson-Delaunay tessellation. Blandt andet gives et alternativt bevis for nogle af

Miles' berømte resultater, som bygger på en vigtig klassisk formel fra integralgeometrien kaldet Blaschke-Petkantschins formel. En simpelt bevis for denne formel findes i [G]; beviset er baseret på invariant målteori, mens tidligere beviser involverer manipulation med differentialformer.

En endnu grundigere fremstilling af stokastiske Voronoi og Delaunay tessellationer findes i monografien [L], som jeg udarbejdede i forbindelse med afholdelsen af en sommerskole i 1992 i Seattle. For at forenkle fremstillingen i [L] betragtes kun tilfældet, hvor $d \leq 3$, men de fleste resultater lader sig umiddelbart generalisere til højere dimensioner. I [L] gives en detaljeret og systematisk beskrivelse af stokastiske Voronoi og Delaunay tessellationer vedrørende den historiske udvikling, anvendelsesområder, topologiske og geometriske egenskaber, sandsynlighedsteoretiske resultater, stokastisk simulation og specifikke statistiske modeller.

Nogle nye resultater om Poisson-Voronoi tessellationer er inkluderet i oversigtsarbejdet [O]. Det uddybes i [O], hvorfor Poissonmodeller spiller en central rolle i teorien for stokastiske tessellationer på grund af den såkaldte Slivnyak-Mecke formel, som giver en entydig karakterisering af Poissonprocessen. Følgelig har der nok været en vis tilbøjelighed til at betragte stokastiske tessellationer baseret på Poissonmodeller som værende de eneste interessante og matematisk håndterlige modeller. Som diskuteret i [O] udgør Gibbs punktprocesser som karakteriseret ved Nguyen-Zessins integralligninger — der kan betragtes som en generalisering af Slivnyak-Meckes formel — et interessant alternativt.

En del af resultaterne i [V] finder anvendelse inden for teorien for Poisson-Voronoi, Poisson-Delaunay og Poisson-hyperplan tessellationer. I [V] studeres Poissonprocesser på generelle tilstandsrum og med intensitetsmålet skalaret med en positiv parameter, og det vises at forskellige geometriske karakteristika er beskrevet ved bestemte betingede uafhængighedsresultater og gammafordelinger. Desuden gives et alternativt bevis af Miles' komplementaritetssætning ved hjælp af Palm målteori.

I de fleste specifikke modeller for tessellationer er cellerne konvekse. Johnson-Mehl tessellationerne udgør en fleksibel klasse af tessellationer, som tillader celler, der ikke er konvekse. De er vanskeligere at analysere end Voronoi tessellationer på grund af mere komplicerede geometriske og topologiske egenskaber. [J] indeholder en dybtgående og omfangsrig analyse af Johnson-Mehl tessellationer: deres geometriske og topologiske egenskaber studeres; Palm målteori og middelværdirelationer for stationære Johnson-Mehl tessellationer og deres snit med lavere-dimensionale affine underrum udvikles; nye resultater for Poisson-Johnson-Mehl tessellationer etableres; og bestemte parametriske modelklasser undersøges i detaljer.

I [M] beskrives, hvordan typiske Poisson-Johnson-Mehl celler kan simuleres og forskellige empiriske resultater præsenteres og diskuteres. Specielt generaliseres Quine og Watsons radiale simulationsalgoritme for stationære Poissonprocesser i det d-dimesionale euklidiske rum til de Poisson-fødselsprocesser, der omtales i [J].

En aktuel oversigt over resultater for stokastiske Johnson-Mehl tessellationer findes i [R]. Desuden bevises et hovedresultat i [R] vedrørende de geometriske karakteristika for Poisson-Johnson-Mehl tessellationer, som tidligere kun har været bevist i nogle specialtilfælde i [J].

(iv) Perspektiver — fremtidigt arbejde og uløste problemer

Der er fortsat et stort behov for udvikling af nye statistiske modeller og metoder inden for rumlig statistik og stokastisk geometri. I betragtning af de fremskridt der finder sted inden for MCMC — perfekt simulation er et godt eksempel herpå — og hvor hurtigt kendskabet til MCMC metoder spredes, må det formodes, at samspillet mellem rumlig statistik, stokastisk geometri og MCMC også i årene fremover vil være et aktivt forskningsområde. I det følgende beskrives kort nogle igangværende eller nylig afsluttede forskningsprojekter, der er relateret til denne afhandling, og som peger på nogle mulige fremtidige indsatsområder.

Udvekslingen af frugtbare ideer mellem statistisk fysik og rumlig statistik/stokastisk geometri må forventes at fortsætte fremover. For eksempel er Gibbs hard-core punktprocesser og faseovergangsfænomener et kontroversielt emne i statistisk fysik; et matematisk bevis for eksistens af en faseovergang er endnu ikke etableret, selvom forskellige teoretiske og simulerede resultater peger i den retning. Gibbs hard-core punktprocesser er også interessante i rumlig statistik og stokastisk geometri; fx kan de modellere en langt højere grad af regularitet end Máterns hard-core modeller og den såkaldte SSI-proces. I [124] studereres visse monotoni- og konvergensegenskaber ved pakningstætheden for en Gibbs hard-core punktproces, og 'simulated tempering' vises at være et effektivt alternativ til de simulationsteknikker, som fysikerne har anvendt.

Likelihood inferens og MCMC metoder har endnu ikke været anvendt i samme omfang inden for stokastisk geometri som i rumlig statistik. Modeller for tilfældige mængder som studeret i stokastisk geometri konstrueres ofte ved at specificere en mærket punktprocesmodel af geometriske objekter, så den tilfældige mængde er givet ved foreningsmængden af de geometriske objekter. I [28] diskuteres, hvorledes likelihood analyse baseret på MCMC metoder for 'missing data' problemer dermed finder anvendelse i forbindelse med den boolske model og beslægtede modeller fra stokastisk geometri.

Statistisk analyse af rumlig inhomogenitet i punktprocesser med lokal interaktion er et relativt uopdyrket emne til trods for dets vigtighed i en række anvendelser. Hvis inhomogeniteten er den dominerende faktor, er log-gaussiske Cox-processer en relevant modelklasse; resultater og metoder for det homogene tilfælde som betragtet i [T] kan udvides til det inhomogene tilfælde. En anden interessant modelklasse fremkommer ved at udtynde i en homogen Markov punktproces, hvor variationen i udtyndingssandsynlighederne bestemmer inhomogeniteten i den udtyndende punktproces. I [9] udvikles forskellige parametriske og semiparametriske inferensprocedurer for denne type af modeller.

Et andet vigtigt indsatsområde er dynamiske rumlige processer — hovedparten af den hidtil udviklede teori i rumlig statistik og stokastisk geometri er relateret til rumlige datasæt observeret på et givent tidspunkt. I [29] udvides multitype log-gaussiske Cox-processer til rumlige processer med kontinuert tid, bl.a. med henblik på modellering af ukrudtsdata i form af en tidsrække af forskellige punktmønstre svarende til forskellige arter.

Asymptotisk estimations- og testteori for rumlige processer er et vanskeligt emne af flere grunde — ikke mindst i forbindelse med Markov sammenhængskomponent felter og Baddeley-Møller Markov punktprocesser, hvor lokale interaktioner så at sige kan være uendelige (i den eksisterende asymptotiske teori for klassiske Markov felter og Markov punktprocesser antages ofte, at der er en endelig interaktionsradius). Hvis der ikke sker en udvikling, hvor flere asymptotiske resultater gøres lettere tilgængelige og anvendelige i forbindelse med rumlige processer, så vil metoder baseret på MCMC fortsat være af større vigtighed i praksis.

Det er nok for tidligt at forudsige, hvilken retning forskningen og anvendelser af perfekt simulation tager fremover. I [109] udvikles perfekte simulationsalgoritmer for lokalt stabile punktprocesser, som udgør en stor klasse af punktprocesmodeller. Udvidelser af Fills perfekte simulationsalgoritme studeres i [143]. Og i [142] introduceres en perfekt 'simulated tempering' algoritme, mens [139] omhandler perfekt slice sampling. Adskellige eksempler på konkrete 'rumlige' anvendelser af disse perfekte simulationsalgoritmer findes i [109, 143, 142, 139]. Men mange andre interessante emner bør også studeres. Hvordan anvendes perfekt simulation bedst i forbindelse med Monte Carlo metoder? Hvor gode er diverse konventionelle MCMC algoritmer og metoder, når de testes ved hjælp af perfekte simulationer?

Hidtidige studier af stokastiske tessellationer har fokuseret mere på de matematiske end de statistiske aspekter. Dette forhold vil muligvis ændre sig fremover — der er en udbredt opfattelse af, at de 'pæne' egenskaber ved analytisk håndterlige modeller for stokastiske tessellationer nu stort set er etableret. MCMC metoder har endnu ikke været anvendt i synderligt omfang i forbindelse med statistisk analyse af stokastiske tessellationer. I [144] undersøges, hvorvidt en ny prior model for to- eller tre-dimensionelle Voronoi tessellationer og MCMC metoder finder anvendelse i bayesiansk billedbehandling, herunder til prediktion i forbindelse med olie og gas forekomster — prior modellen er af samme type som modellerne i [A]. Endelig omhandler [191] en bayesiansk analyse af stokastiske tessellationer, når der kun er en tilknyttet punktproces — defineret på de lavere-dimensionale facetter af tessellationen — som kan observeres.