19 August 2002

BAYESIAN ANALYSIS OF DEFORMED TESSELLATION MODELS

PAUL G. BLACKWELL,* *University of Sheffield*

JESPER MØLLER,** *Aalborg University*

Abstract

We define a class of tessellation models based on perturbing or deforming standard tessellations such as the Voronoi tessellation. We show how distributions over this class of 'deformed' tessellations can be used to define prior distributions for models based on tessellations, and how inference for such models can be carried out using Markov chain Monte Carlo methods; stability properties of the algorithms are investigated. Our approach applies not only to fixed dimension problems, but also to variable dimension problems, in which the number of cells in the tessellation is unknown. We illustrate our methods with two real examples. The first relates to reconstructing animal territories, represented by the individual cells of a tessellation, from observation of an inhomogeneous Poisson point process. The second example involves the analysis of an image of a cross-section through a sample of metal, with the tessellation modelling the micro-crystalline structure of the metal.

Keywords: random tessellation; Voronoi tessellation; deformed template; Bayesian inference; Markov chain Monte Carlo; reversible jump; Harris recurrence; uniform ergodicity; inhomogeneous Poisson process; image analysis.

AMS 2000 Subject Classification: Primary 60D05

Secondary 62M30;60G55

* Postal address: Department of Probability and Statistics, University of Sheffield, Hicks Building, Sheffield S3 7RH, UK. Email address: p.blackwell@sheffield.ac.uk

** Postal address: Department of Mathematical Sciences, Aalborg University Fredrik Bajers Vej 7G, DK-9220 Aalborg, Denmark. Email address: jm@math.auc.dk

1. Introduction

A major topic in stochastic geometry is random tessellations. Briefly, a tessellation of a d-dimensional Euclidean space is a subdivision of the space into non-overlapping d-dimensional sets called cells. Joseph Mecke is one of the main contributors to the development of a general theory for random tessellations, cf. Chapter 10 in [27] and the references therein. In particular his approach of Palm measure theory has proven to be very successful, and has indeed been of great value for one of us, see e.g. [15, 16, 17, 18, 19].

So far, research on random tessellations has focused rather on mathematical modelling and analysis than statistical aspects. Some recent exceptions are the work in [21, 1, 2, 20] which are all based on a Bayesian Markov chain Monte Carlo (MCMC) approach. We find that such an approach is both natural and very useful for many statistical applications of random tessellations, partly because of the complicated structures and models used and partly because some prior knowledge is often available. The abovementioned papers are all related to particular applications, and there is indeed scope for a further development of Bayesian MCMC methods for random tessellations. In this paper we present a way of modelling flexible priors for random deformations of tessellation models, discuss how a Bayesian MCMC analysis can be performed, and illustrate the methodology on some real datasets. The idea of using deformed templates for Bayesian image problems goes at least back to the seminal work by Ulf Grenander, see e.g. [11, 12].

In Sections 2 and 3, for simplicity and specificity we restrict attention to a template given by a planar Voronoi tessellation [17, 27, 22]. Specifically, we consider first a Voronoi tessellation defined on a convex compact set $S \subset \mathbb{R}^2$, with non-empty interior $\mathrm{int}S$ and boundary ∂S. This is generated by a finite point configuration $x = \{x_1, \ldots, x_n\} \subset \mathbb{R}^2$ of "nuclei" so that the Voronoi cell with nucleus x_i is given by

$$\mathcal{C}(x_i|x) = \{s \in S : \|s - x_i\| \leq \|s - x_j\| \text{ for all } j \neq i\}, \quad i = 1, \ldots, n.$$

An example is shown in Figure 1. A nonempty intersection between two or more cells is either empty, a single point called a vertex of $V(x)$, or a bounded line segment (of positive length) called an edge of $V(x)$. Let $\mathcal{V}(x)$ be the set of vertices of $V(x)$, and assume that

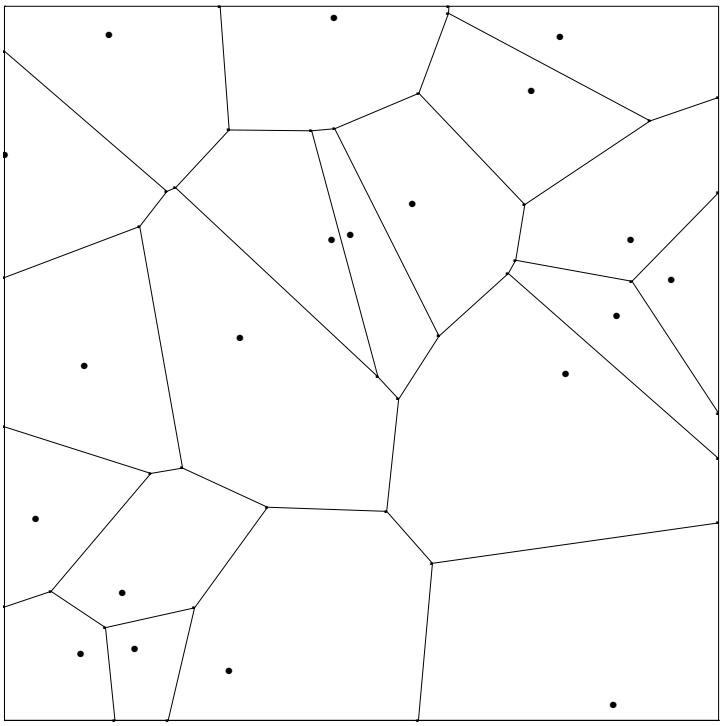

FIGURE 1: A Voronoi tessellation

(C1) $\mathcal{V}(x) \subset \text{int}S$ and no edge of $V(x)$ has more than two intersection points with ∂S.

Such intersection points are called boundary vertices of $V(x)$.

Next each vertex $v_j \in \mathcal{V}(x)$, is translated or perturbed by a vector $z_{v_j} \in \mathbb{R}^2$; in the sequel we abuse notation and write z_j for z_{v_j}. We assume that this creates vertices $v'_j = v_j + z_j,\ v_j \in \mathcal{V}(x)$, of a new tessellation of S with the same connectivity structure as in $V(x)$. Figure 2 shows a Voronoi tessellation and its perturbation. More precisely we assume that

(C2) the perturbed vertices are pairwise different and contained in int$S$;

(C3) the perturbed edges are disjoint except possibly at their endpoints.

Clearly, (C2) and (C3) ensure that the perturbed edges define a new tessellation of S with cells, edges and vertices which are in one-to-one correspondence with those of

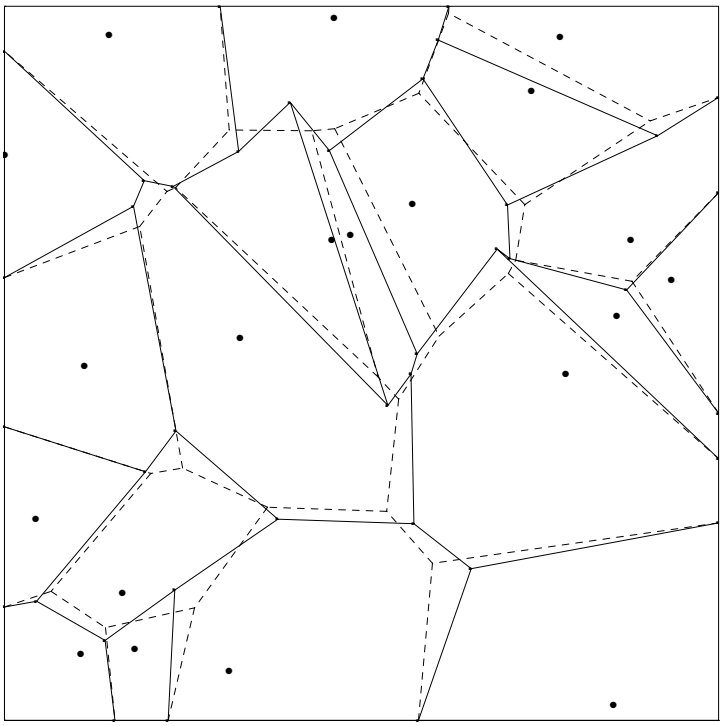

FIGURE 2: A perturbation of a Voronoi tessellation

$V(x)$. This perturbed (or deformed) tessellation is denoted $V(x,z)$ where $z = \{z_j : v_j \in \mathcal{V}(x)\} \subset \mathbb{R}^2$ specifies the associated "disturbances". The Voronoi cells are convex, while the cells of $V(x,z)$ are possibly nonconvex. However both types of cells are polygonal except possibly at the boundary edges.

Our approach is easiest to understand, and to implement, in the case where the point configuration of nuclei x is fixed, since the connectivity structure of $V(x)$, and hence the dimension of the problem, is then fixed and known. Section 2 considers this case, while Section 3 concerns the case where x is random. Note that we do not perturb the boundary vertices of $V(x)$. Our approach may easily be extended to this case, cf. Section 4, but we imagine that S is so large that perturbing the boundary vertices of $V(x)$ is not needed or may even be unnatural. For the same reason we henceforth let x be contained in S. For example, in Section 3.3, the disturbed tessellation is observed within an observation window $W \subset S$, where S is chosen to be sufficiently larger than W in order to account for edge effects.

Given (x, z) and various hyperparameters introduced in Sections 2 and 3, different types of data and observation models are considered in the application examples in Sections 2.3, 3.3 and 3.4. In Sections 2.1 and 3.1 we combine the observation model with a prior for (x, z) and the hyperparameters to obtain a posterior distribution for nuclei, perturbations and hyperparameters. Then in Sections 2.2 and 3.2 we construct various algorithms for making MCMC simulations from the posterior, and demonstrate their performance for making Bayesian inference in Sections 2.3, 3.3 and 3.4. Furthermore, in Sections 2.3, 3.3 and 3.4, we investigate the stability properties of our MCMC algorithms.

The data in Sections 2.3 and 3.3 is a point pattern of badger latrines, and the observation model is the same as in [2], namely an inhomogeneous Poisson process with an intensity that is related to the deformed tessellation. The data in Section 3.4 is a noisy blurred grey-scale image of the grain structure in a two-dimensional cross-section through a sample of metal. In both case the aim is to reconstruct the unobserved tessellation and study the uncertainty of various tessellation characteristics.

Finally, Section 4 contains some concluding remarks. In particular, we discuss how our approach easily extends to other types of template models for polygonal tessellations than just Voronoi tessellations.

Since MCMC methods are not very widely used by stochastic geometers, we presume only a limited knowledge, but the reader should at least be familiar with the Metropolis-Hastings algorithm in a fairly general setting — for background material on MCMC, see e.g. [9] or [23].

Besides the notation introduced above, we use the following notation throughout this paper. Let $\mathbf{1}_{(x,z)}$ denote the indicator function of the event that (C1)–(C3) are satisfied. Let $\mathrm{N}(\mu, \Sigma)$ denote the bivariate normal distribution with mean μ and covariance matrix Σ, and let $\varphi(s; \sigma^2) = \exp(-\|s\|^2/(2\sigma^2))/(2\pi\sigma^2)$ be the density for $\mathrm{N}(0, \sigma^2 I)$ where $\sigma^2 > 0$ and I is the identity matrix. Finally, let $n(A)$ denote the cardinality of any finite set A.

2. Deformation of Voronoi tessellations with fixed nuclei

Throughout this section we let the nuclei x be fixed and assume that (C1) is satisfied. We often suppress in the notation the dependence on x.

2.1. Specification of prior and posterior distributions

To begin with the disturbances z are taken to be independent with density $\varphi(s;\sigma^2)$ subject to (C2) and (C3) holding. Thus the prior for the disturbances z given σ^2 is

$$\pi(z|\sigma^2) = c(\sigma^2) \prod_{\mathcal{V}(x)} \varphi(z_j;\sigma^2)\mathbf{1}_{(x,z)} \tag{1}$$

where $c(\sigma^2)$ is a normalising constant which depends on σ^2. Note that a closed form expression for $c(\sigma^2)$ is unknown. An obvious way to complete the prior for the perturbed tessellation would be to take a conjugate prior $\pi(\sigma^2)$ for σ^2, say $\sigma^2 \sim$ InverseGamma(a_1, a_2) where $a_1, a_2 > 0$ are user-specified parameters (see Section 2.3), and then let

$$\begin{aligned} \pi(z,\sigma^2) &= \pi(z|\sigma^2)\pi(\sigma^2) \\ &\propto c(\sigma^2) \prod_{\mathcal{V}(x)} \varphi(z_j;\sigma^2)\mathbf{1}_{(x,z)}(\sigma^2)^{-(a_1+1)} \exp(-a_2/\sigma^2). \end{aligned} \tag{2}$$

This conditional formulation, although natural, would however lead to difficulties in any Metropolis-Hastings implementation because $c(\sigma^2)$ is unknown. Specifically, the Hastings ratio for a proposal $\sigma^2 \to \tilde{\sigma}^2$ would depend on $c(\sigma^2)/c(\tilde{\sigma}^2)$. In principle this ratio of normalising constants may be estimated by e.g. path sampling [6], but this would be computationally demanding.

Instead, we define a joint prior for z and σ^2 directly by

$$\pi(z,\sigma^2) \propto \prod_{\mathcal{V}(x)} \varphi(z_j;\sigma^2)\mathbf{1}_{(x,z)}(\sigma^2)^{-(a_1+1)} \exp(-a_2/\sigma^2). \tag{3}$$

This formulation is not identical to that in (2), since $c(\sigma^2) \neq 1$ and the marginal prior for σ^2 is no longer exactly InverseGamma(a_1, a_2), but the conditional prior distribution (1) is the same in the two models. Furthermore, the marginal priors for σ^2 may often be expected to be rather close, as the case of small perturbations is likely to be the most relevant in practical applications. More specifically $c(\sigma^2)$ is close to 1, when σ^2

is sufficiently small that (C2) and (C3) almost always hold under the model (2). For the rest of this section, we consider the model defined in (3).

As well as the prior for the tessellation, we need to consider the model generating observations conditional on the tessellation, and the prior for its parameters. For the present, we write θ for the (possibly vector) parameter of that model, assume θ to be independent of (z, σ^2) with prior $\pi(\theta)$, and write $L(y|x, z, \theta)$ for the likelihood for observations y (this is what we called the observation model in Section 1). Examples of likelihoods are given in Sections 2.3 and 3.3.

Finally, combining the prior and the likelihood terms we obtain the posterior density

$$\pi(z, \sigma^2, \theta|y) \propto \prod_{\mathcal{V}(x)} \varphi(z_j; \sigma^2) \mathbf{1}_{(x,z)} (\sigma^2)^{-(a_1+1)} \exp(-a_2/\sigma^2) L(y|x, z, \theta)\pi(\theta).$$

2.2. MCMC implementation

We can now specify a MCMC algorithm generating a Markov chain with the posterior 2.1 as its equilibrium distribution. We use a so-called hybrid algorithm (sometimes also called a Metropolis-within-Gibbs algorithm) with separate updates for z, σ^2 and θ; specific updating schemes are given in Sections 2.3, 3.3 and 3.4. We denote the generic current state of the chain by (z, σ^2, θ), and assume that it is in the support of the posterior, i.e. $\mathbf{1}_{(x,z)} = 1$ and $\sigma^2 > 0$. Updating for θ depends on the form of the likelihood, cf. Section 2.3, so in the present section we just describe the updates of z and σ^2.

To update z, we first choose some random set of vertices $\mathcal{V} \subset \mathcal{V}(x)$; a single random vertex, the vertices associated with a random nucleus or the whole of $\mathcal{V}(x)$ are obvious choices, but any scheme which picks each vertex with positive probability and which does not depend on z could be used. We then propose disturbances z', with

$$z'_j \sim \mathrm{N}(z_j, \tau_z^2 I)$$

independently for each z_j corresponding to a vertex in $\mathcal{V}$, and $z'_j = z_j$ otherwise. The Hastings ratio for such a proposal is

$$\begin{aligned} r(z, z'|\sigma^2, \theta) &= \frac{\prod_{\mathcal{V}} \varphi(z'_j; \sigma^2) \mathbf{1}_{(x,z')} L(y|x, z', \theta)}{\prod_{\mathcal{V}} \varphi(z_j; \sigma^2) L(y|x, z, \theta)} \\ &= \exp\Big(\sum_{\mathcal{V}} \big(||z_j||^2 - ||z'_j||^2 \big) / (2\pi\sigma^2) \Big) \mathbf{1}_{(x,z')} L(y|x, z', \theta) / L(y|x, z, \theta) \end{aligned}$$

and the acceptance probability is $1 \wedge r(z, z'|\sigma^2, \theta)$.

To update σ^2, we use Gibbs sampling, exploiting the fact that

$$\sigma^2|z, y \sim \text{InverseGamma}(a_1 + n(\mathcal{V}(x)), a_2 + \frac{1}{2}\sum_{\mathcal{V}(x)} ||z_j||^2).$$

2.3. Reconstruction of badger territories

In practice, this framework is rather limited in its applicability, because of the assumption that the nuclei are known. A case where this assumption is natural is in the reconstruction of animal territories where nest sites or the equivalent are known. [2] gives an example concerning the reconstruction of badger (*Meles meles*) territories when the locations of their setts (burrows) are known. That paper concentrates on the case where the tessellation is assumed to be Voronoi, but the nuclei are not precisely known; they are required to be consistent with the observed setts, in the sense that the Voronoi tessellation of the nuclei must contain one sett per cell. A natural alternative, however, is to fix the nuclei to coincide with the setts and allow the tessellation to be perturbed away from the Voronoi. The approach of [1] allows this by considering a weighted tessellation, but below we instead apply the model of this section.

The observation in this case consists of the locations of badger latrines, which tend to occur close to the boundaries between territories. Figure 3 shows the data set used here; the smaller dots indicate locations of latrines, and the larger circles indicate the known locations of setts. We assume that the latrines can be modelled by an inhomogeneous Poisson process, observed within a window $W \subseteq S$, with an intensity that is related to the deformed tessellation, c.f. [2]. Specifically, let $E(x, z) \subset S$ be the union of all edges of $V(x, z)$, and for each $s \in S$, let $d(s, E(x, z)) = \bigwedge_{e \in E(x,z)} ||s - e||$ denote the minimum distance from s to a point on such an edge. Then the intensity of the process at s is

$$\rho(s; x, z, \alpha, \beta, \gamma) = \alpha + \beta g(d(s, E(x, z))/(2\gamma))$$

where $\alpha, \beta, \gamma > 0$, and $g(\cdot)$ is the decreasing sigmoid function

$$g(d) = \begin{cases} 1/\left(1 + (d/(1-d))^2\right), & 0 \le d \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

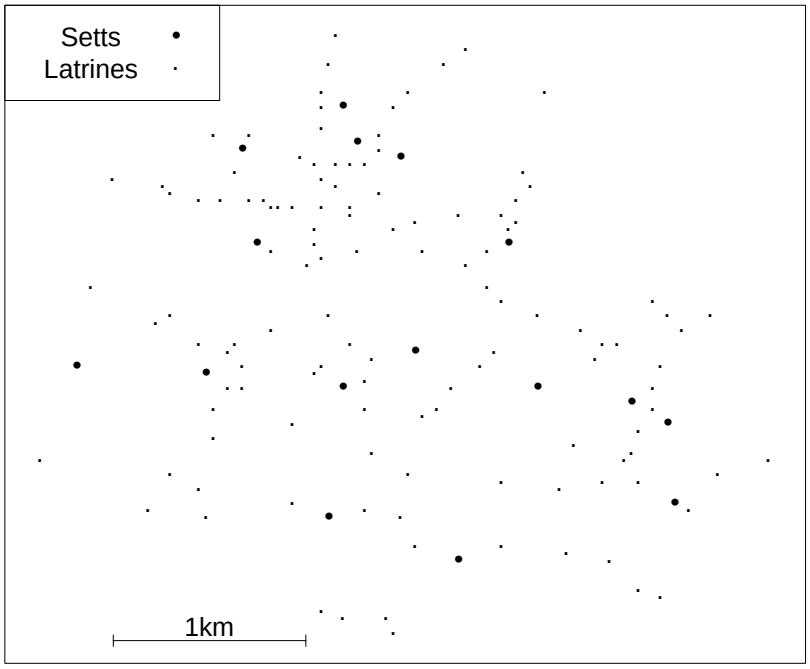

FIGURE 3: Badger setts and latrines

Thus α represents the background intensity of the process of latrines, β represents the extra intensity at a territory border, and γ represents the distance at which the extra intensity drops to half its value at a boundary; in the notation of Section 2.1, $\theta = (\alpha, \beta, \gamma)$.

We take the priors for these parameters to be independent Gamma distributions, with $\alpha \sim \mathrm{Ga}(b_1, b_2)$, $\beta \sim \mathrm{Ga}(c_1, c_2)$, $\gamma \sim \mathrm{Ga}(d_1, d_2)$.

For the purposes of this section, we take the observation window W to be the same as the space S on which the tessellation is defined. In Section 3.3 we extend this to the case where W is strictly contained in S and where the nuclei outside W are not fixed, as a way of dealing with edge effects.

Writing $y = \{y_1, \ldots, y_k\}$ for the realisation of the point process within W, the likelihood is given by

$$L(y|x, z, \theta) = \exp(-\int_W \rho(s; x, z, \theta)\mathrm{d}s) \prod_{y_i \in y} \rho(y_i; x, z, \theta).$$

Evaluation of this likelihood can be carried out exactly if all the cells are convex (see [2]

for details), but extending this calculation to non-convex polygons, which arise when the Voronoi tessellation is perturbed, is not straightforward using the same technique. Instead, we use the approach of [1] which uses a spatial discretisation; details are omitted here.

2.3.1. *Algorithmic details and stability properties* We now specify the θ-update and various hyperparameters, and study the stability properties of the hybrid algorithm.

The likelihood involves α, β, γ in a rather complex way; we therefore update these parameters simultaneously, using independent Gaussian random walk proposals. More precisely, if $\theta = (\alpha, \beta, \gamma)$ is the current state, we propose independent normal variates α', β', γ' with means α, β, γ and standard deviations $\tau_\alpha, \tau_\beta, \tau_\gamma$, respectively. The Hastings ratio for the proposal $\theta' = (\alpha', \beta', \gamma')$ is then

$$h(\theta, \theta'|z, \theta) = \frac{\pi(\theta')L(y|x, z, \theta')\mathbf{1}[\theta' \in (0, \infty)^3]}{\pi(\theta)L(y|x, z, \theta)} \tag{4}$$

and the acceptance probability is $1 \wedge h(\theta, \theta'|z, \theta)$.

To complete the implementation of the algorithm, we need to specify the values for $\tau_\alpha, \tau_\beta, \tau_\gamma$, the value of τ_z and choice of vertex set $\mathcal{V}$ in updating z, and the schedule of different types of step. For simplicity, we use a standard systematic-scan for updating z, σ, θ, and let $\mathcal{V}$ consist of all vertices in $V(x)$. The standard deviations are chosen in line with the results in [24], i.e. so that the acceptance probability is about 20-30 % for both z-updates and θ-updates.

The following proposition ensures that the Markov chain underlying the hybrid algorithm converges toward the posterior distribution whatever the initial state, and that the law of large numbers applies for estimating posterior means by ergodic averages (see e.g. [14]). That is, the initial state can be given by any $(z_0, \sigma_0^2, \theta_0)$ in the support of the posterior; e.g. $z_0 = 0$ and suitably chosen $\sigma_0^2 > 0$ and $\theta_0 \in (0, \infty)^3$.

Proposition 1. *The hybrid algorithm as specified above is ergodic, i.e. aperiodic and Harris recurrent.*

Proof. Set $m = n(\mathcal{V}(x))$. The support of the posterior is given by $A = B \times (0, \infty)^4$ where $B = \{z \in \mathbb{R}^{2m} : \mathbf{1}_{(x,z)} = 1\}$ is bounded. The chain is obviously irreducible on A, and by construction of the algorithm it is easily verified that any d-cycle (see [14])

must have $d = 1$, i.e. the chain is aperiodic. By Theorem 1 in [3], the chain is Harris recurrent.

It is tempting to try to establish geometric ergodicity, since this implies a central limit theorem for ergodic averages (see e.g. [14]). There exist a few papers on conditions ensuring geometric ergodicity for hybrid algorithms, see [25, 26] and , but these conditions seem hard to use in the present setting. Suppose we restrict the priors for $\sigma^2, \alpha, \beta, \gamma$ to compact intervals $I_{\sigma^2}, I_\alpha, I_\beta, I_\gamma \subset (0, \infty)$, whereby the prior densities and the intensity $\rho(\cdot)$ are uniformly bounded from above and away from 0. These restrictions on the ranges of $\sigma^2, \alpha, \beta, \gamma$ can be chosen to have a negligible effect on the practical interpretation of the model. An upper bound on σ^2 and γ that is large compared with the dimensions of S will make no practical difference, nor will an upper bound on α and β that is large compared with some realistic upper limit on the density of latrines (based on their physical size). Similarly, a lower bound on σ^2 that is small compared with the likely accuracy of the sett and latrine locations in the data will have no practical effect, nor will a lower bound on α that is small compared with $|W|^{-1}$. Finally, if there is prior belief that $\beta \approx 0$ or $\gamma \approx 0$, this will be reflected in the choice of hyperparameters, with $c_1 \leq 1$ or $d_1 \leq 1$ respectively.

We adjust for these restrictions in the Gibbs update for σ^2 which now follows an inverse gamma distribution restricted to I_{σ^2}, and in the Hastings ratio (4) where we replace $(0, \infty)^3$ in the indicator function by $I_\alpha \times I_\beta \times I_\gamma$. Then the algorithm becomes not only geometrically but uniformly ergodic as shown in the next proposition.

Proposition 2. *The modified hybrid algorithm as specified above is uniformly ergodic.*

Proof. Let $P^m(\cdot, \cdot)$ denote the m-step transition kernel for the Markov chain associated with the modified hybrid algorithm, and A_{mod} the support of the modified posterior distribution. Then $A_{\text{mod}} = B \times I_{\sigma^2} \times I_\alpha \times I_\beta \times I_\gamma$ with B as in the proof of Proposition 1. By continuity of densities, boundedness of B, and the definitions of $I_{\sigma^2}, I_\alpha, I_\beta, I_\gamma$, the modified posterior density and the various proposal densities are uniformly bounded from above and away from 0 as long as the proposal is in A_{mod}, and hence the different Hastings ratios are uniformly bounded away from zero as long as the proposal is in A_{mod}. Now, uniform ergodicity is equivalent to the condition that for some $m \in \mathbb{N}$, some nontrivial measure Q, and all $(z, \sigma^2, \theta) \in A_{\text{mod}}$, we have that

$P^m((z,\sigma^2,\theta),\cdot) \geq Q(\cdot)$ (Theorem 16.0.2 in [14]). But if $m = n(\mathcal{V}(x))$ and $\epsilon > 0$ is a lower bound for both the two Hastings ratios (when the proposal is in A_{mod}) and the three proposal densities (including the inverse gamma distribution restricted to I_{σ^2}), then for any Borel set $F \subseteq A_{\mathrm{mod}}$, $P^m(\cdot,F) \geq (\epsilon^5 m!/m^m) \int_{(z,\sigma^2,\theta)\in F} \mathrm{d}z\mathrm{d}\sigma^2\mathrm{d}\theta$, where $m!/m^m$ is the probability that the m selected vertices of $V(x)$ are all different.

2.3.2. *Results* The results given here are based on the data in Figure 3, with all distances expressed in kilometres, and where $W = S$ is the rectangular region shown in Figure 3. Note that in this application, the underlying (undeformed) Voronoi tessellation is fixed, with nuclei equal to the setts, and the individual cells can be interpreted as the territories of particular badger groups.

The priors for α, β, γ are diffuse, i.e. they correspond to letting $b_1, b_2, c_1, c_2, d_1, d_2 \to 0$. The range of σ^2 is bounded above at 0.05, since sampling from the prior shows that large deviations from Voronoi tessellations are inconsistent with the biological meaning of the tessellations. Apart from this constraint, the prior for z and σ^2 is of the form in (3), with $a_1 = 20, a_2 = 0.01$, so the prior for σ^2 is approximately InverseGamma(a_1, a_2).

The posterior used here is based on 2 runs each of 70,000 iterations, with a burn-in of 10,000 iterations, starting with the zs generated independently given an initial value of $\sigma = 0.025$ (subject to the conditions (C1)–(C3) on the edges of the resulting tessellation).

As with any fully Bayesian analysis of a complex model, there are many aspects of the joint posterior that can readily be investigated; here we simply give some illustrative examples. Figure 4 shows the most likely single tessellation under this model. Displaying the uncertainty in the whole tessellation is difficult, but Figure 5 shows a sample (of size 10) from the posterior distribution for one particular cell, indicating the uncertainty in its boundaries. Finally, Figure 6 shows the posterior expected intensity of latrines at each location, under this model, in latrines per square kilometre. Again, the uncertainty is not shown, though the information is readily available from the MCMC output.

The efficiency of the MCMC algorithm can be controlled through the proposal variance, τ_z^2; the runs here have an acceptance rate of about 40%, somewhat higher than is optimal.

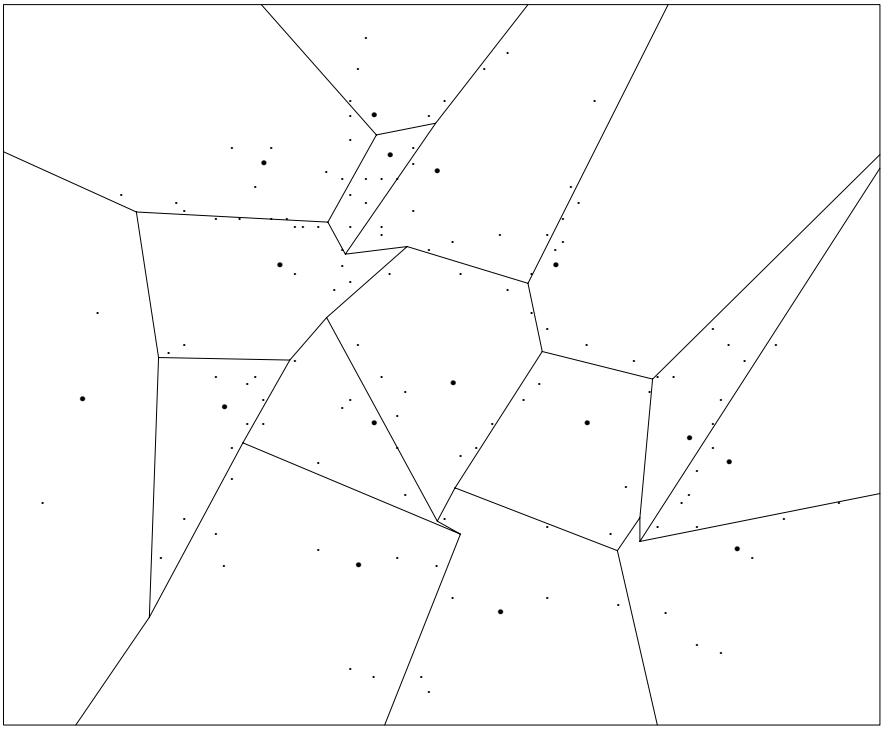

FIGURE 4: The most likely reconstruction of the territories

3. Unknown nuclei

The model in Section 2.1 can be extended to the case where the number and locations of the nuclei $x \subset S$ are unknown. In the sequel we let $T \subseteq S$ be a given Borel set with area $|T| > 0$, and assume that $x \setminus T$ is fixed but $x_T \equiv x \cap T$ is unknown. Section 3.3 contains an example of how T may be specified. We suppress in the notation the dependence on $x \setminus T$.

3.1. Specification of prior and posterior distributions

We need to specify a prior for x_T. The natural candidate is a homogeneous Poisson process on T with unknown intensity hyper-parameter $\lambda \sim \text{Gamma}(e_1, e_2)$, say, where $e_1, e_2 > 0$ are user-specified parameters, see Sections 3.3 and 3.4. If, as seems natural, our prior beliefs about the extent to which the tessellation is deformed are scale-invariant, then σ^2 should vary inversely with λ, so that $\sigma^2|\lambda \sim \text{InverseGamma}(a_1, a_2/\lambda)$, say.

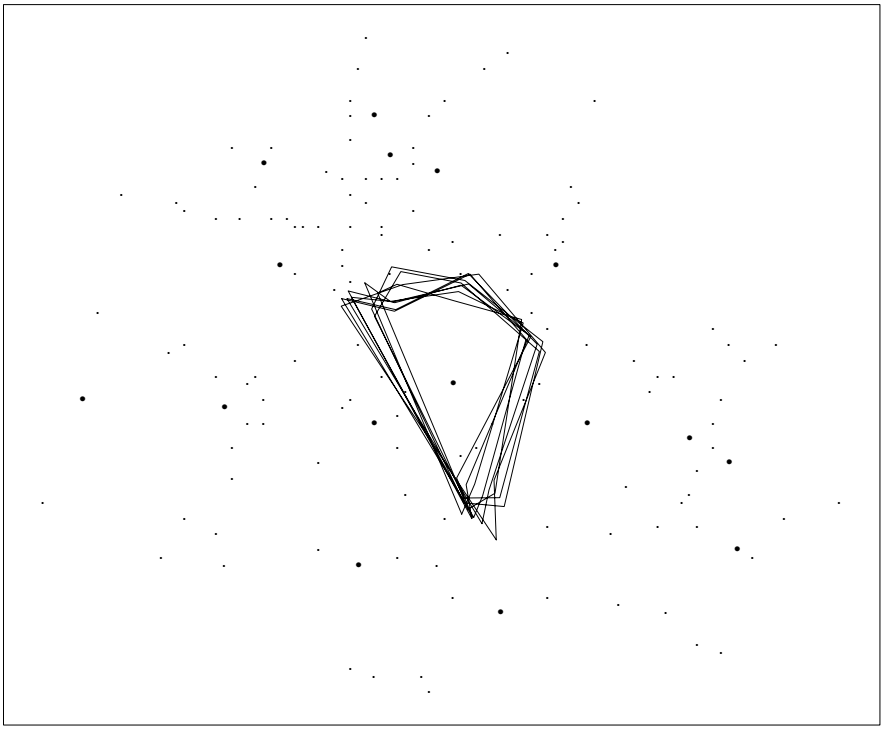

FIGURE 5: A sample from the posterior for a particular cell

However, as in Section 2.1, rather than the obvious conditional formulation of the prior for the tessellation, it is necessary to specify a joint prior for $x_T, z, \lambda, \sigma^2$, to avoid difficulties with unknown normalising constants. We therefore take

$$\pi(x_T, z, \lambda, \sigma^2) \propto \lambda^{e_1-1} e^{-e_2\lambda} \lambda^{n(x_T)} e^{-\lambda|T|} \prod_{\mathcal{V}(x)} \varphi(z_j; \sigma^2) \mathbf{1}_{(x,z)} (\sigma^2)^{-(a_1+1)} \exp(-a_2/(\lambda\sigma^2)). \quad (5)$$

We again take θ to be independent of the $(x_T, z, \lambda, \sigma^2)$ *a priori*, with prior $\pi(\theta)$ that depends on the application. Hence, combining the prior and the likelihood terms we obtain the posterior density

$$\pi(x_T, z, \lambda, \sigma^2, \theta|y) \propto \lambda^{e_1-1} e^{-e_2\lambda} \lambda^{n(x_T)} e^{-\lambda|T|} \times \prod_{\mathcal{V}(x)} \varphi(z_j; \sigma^2) \mathbf{1}_{(x,z)} (\sigma^2)^{a_1-1} \exp(-a_2/(\lambda\sigma^2)) L(y|x, z, \theta)\pi(\theta). \quad (6)$$

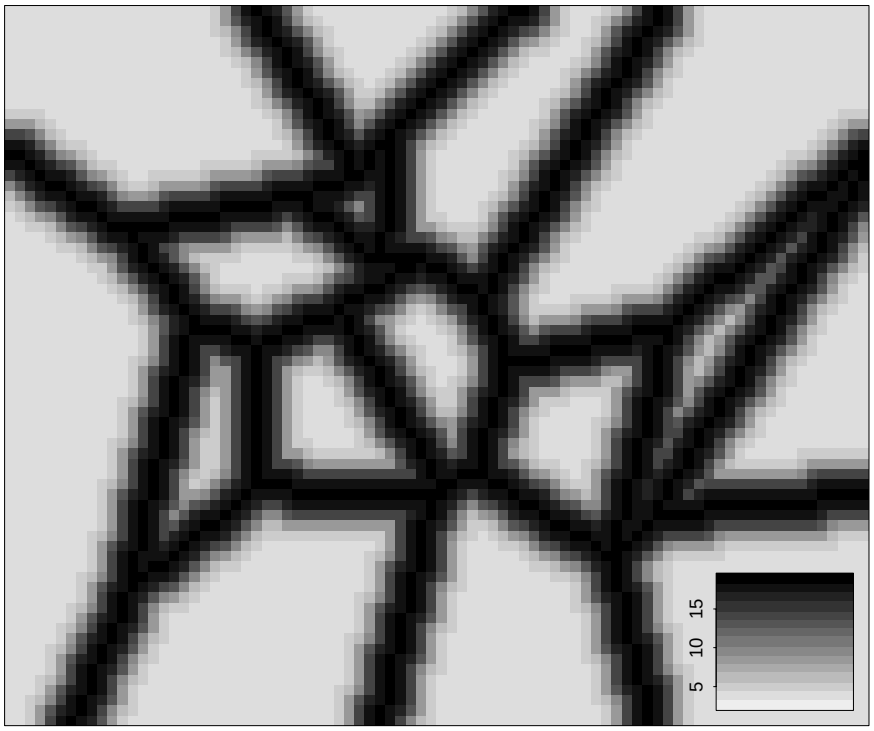

FIGURE 6: Posterior intensity surface

3.2. MCMC implementation

The MCMC algorithm for this posterior is a hybrid one, with separate Gibbs or Metropolis-Hastings steps for updating z, θ, σ^2, and λ, and for updating x_T (with necessary changes to z, as detailed below). Updating for θ is given in Section 3.3, and the updating for z and for σ^2 is as in Section 2 with a_2 replaced by a_2/λ. So in this section we just describe the updates for λ and x_T, assuming the current state is in the support of the posterior.

For λ, recalling its interdependence with σ^2, we have

$$\pi(\lambda|x_T, z, \lambda, \sigma^2, \theta, y) \propto \lambda^{e_1-1+n(x_T)-a_1} \exp(-(e_2+|T|)\lambda) \exp(-a_2/(\sigma^2\lambda))$$

which is the form of the Generalized Inverse Gaussian distribution (GIG) [13]. We have

$$\lambda|x_T, z, \lambda, \sigma^2, \theta, y \sim \text{GIG}(2(e_2+|T|), 2a_2/\sigma^2, e_1+n(x_T)-a_1),$$

and Gibbs sampling is possible using the algorithm of [5].

The main novelty lies in the updating of x_T through Metropolis-Hastings steps in which single nuclei are either born or killed; for convenience, these types of steps are taken to be equally likely. A step in which a nucleus is moved from one location to another is also possible; however, since this will in general lead to the deletion of some vertices, and the creation of others, it is convenient to regard it as a combination of a death step and a birth step. Since changing x_T changes $\mathcal{V}(x)$, these steps necessarily affect z too.

In a birth step, a new nucleus is proposed at a uniform random location ξ in T. The birth leads to the creation of new vertices,

$$\mathcal{V}_+(\xi|x\cup\{\xi\}) = \mathcal{V}(x\cup\{\xi\})\setminus\mathcal{V}(x),$$

and also to the deletion of the vertices

$$\mathcal{V}_-(\xi|x\cup\{\xi\}) = \mathcal{C}(\xi|x\cup\{\xi\})\cap\mathcal{V}(x).$$

Writing z' for the proposed displacements of the vertices, we retain existing displacements where possible, so that $z'_j = z_j$ for vertices in $\mathcal{V}(x)\setminus\mathcal{V}_-(\xi|x\cup\{\xi\})$ and propose new values, from the prior, for new vertices, so that $z'_j \sim \mathrm{N}(0,\sigma^2 I)$ for vertices in $\mathcal{V}_+(\xi|x\cup\{\xi\})$.

In a death step, the deletion of a uniformly randomly chosen nucleus η from x_T is proposed. Again, displacements of newly created vertices are proposed from the prior. In the notation already introduced for the birth step, the new vertices are simply those of $V(x)$ that lie in $\mathcal{C}(\eta|x)$, precisely $\mathcal{V}_-(\eta|x)$, and the vertices removed are just those of the cell of η, that is $\mathcal{V}_+(\eta|x)$.

The birth and death steps can be considered as a version of Peter Green's reversible jump MCMC algorithm [10], and the Hastings ratio can be derived along similar lines as in [8]. The Hastings ratio for a birth proposal $x_T \to x_T\cup\{\xi\}$ is given by

$$r(x_T,z,\xi,z'|\lambda,\theta,y) = \frac{\pi(x_T\cup\{\xi\},z',\lambda,\sigma^2,\theta|y)}{\pi(x_T,z,\lambda,\sigma^2,\theta|y)} \times \frac{\frac{1}{n(x_T)+1}\prod_{\mathcal{V}_-(\xi|x\cup\{\xi\})}\varphi(z_j;\sigma^2)}{\frac{1}{|T|}\prod_{\mathcal{V}_+(\xi|x\cup\{\xi\})}\varphi(z'_j;\sigma^2)}.$$

This simplifies to

$$r(x_T,z,\xi,z'|\lambda,\theta,y) = \frac{|T|\lambda\mathbf{1}_{(x\cup\{\xi\},z')}L(y|x\cup\{\xi\},z',\theta)}{(n(x_T)+1)L(y|x,z,\theta)} \tag{7}$$

cancelling terms and using (6) and the facts that $\mathbf{1}_{(x,z)} = 1$ for an existing tessellation,

$$\mathcal{V}(x \cup \{\xi\}) = [\mathcal{V}(x) \setminus \mathcal{V}_-(\xi|x \cup \{\xi\})] \cup \mathcal{V}_+(\xi|x \cup \{\xi\}) \tag{8}$$

and

$$z \text{ and } z' \text{ agree on } \mathcal{V}(x) \setminus \mathcal{V}_-(\xi|x \cup \{\xi\}). \tag{9}$$

The acceptance probability for a birth proposal $(x_T, z) \to (x_T \cup \{\xi\}, z')$ is then $1 \wedge r(x_T, z, \xi, z'|\ \lambda, \theta, y)$, and for a death proposal $(x_T \cup \{\xi\}, z') \to (x_T, z)$ is $1 \wedge r(x_T, z, \xi, z'|\lambda, \theta, y)^{-1}$. In our code for birth and death updates we exploit the properties (8) and (9), i.e. that these types of updates depend only on local information.

Finally, we consider a step in which a nucleus is moved, assuming $x_T \neq \emptyset$. A nucleus η is uniformly randomly chosen from x_T, and it is proposed to move η to a uniformly randomly chosen point ξ in $b(\eta, r) \cap T$, where $b(\eta, r)$ denotes the ball in $\mathbb{R}^2$ with centre η and radius $r > 0$. We can split this into a death step followed by a birth step: first, $x_T \to x_T \setminus \{\eta\}$ leads to the deletion of $\mathcal{V}_+(\eta|x)$ and the creation of $\mathcal{V}_-(\eta|x)$; next, $x_T \setminus \{\eta\} \to (x \setminus \{\eta\}) \cup \{\xi\}$ leads to the deletion of $\mathcal{V}_-(\xi|x \setminus \{\eta\}) \cup \{\xi\})$ and the creation of $\mathcal{V}_+(\xi|x \setminus \{\eta\}) \cup \{\xi\})$. We retain all other displacements in z, and newly created vertices are proposed from the prior as before. Write z' for the proposed displacements. By similar arguments as those leading to (7), we obtain the Hastings ratio for the move proposal $(x_T, z) \to ((x_T \setminus \{\eta\}) \cup \{\xi\}, z')$,

$$r(x_T, z, \eta, \xi, z'|\lambda, \theta, y) = \frac{|b(\eta, r) \cap T|\mathbf{1}_{\{(x \setminus \{\eta\}) \cup \{\xi\}, z'\}} L(y|(x \setminus \{\eta\}) \cup \{\xi\}, z', \theta)}{|b(\xi, r) \cap T| L(y|x, z, \theta)}$$

and the acceptance probability is $1 \wedge r(x_T, z, \eta, \xi, z'|\lambda, \theta, y)$.

3.3. Reconstruction of badger territories taking boundary effects into account

For this section, we revisit the example of Section 2.3, but we take the space S, on which the tessellation is defined, to be strictly larger than the window W over which the point process of latrines is observed (further details are given in Section 3.3.2). The nuclei in the observation window, $x \cap W$, are assumed known as before; on $T = S \setminus W$, the process of nuclei is unobserved.

We have the same likelihood as in Section 2.3. All latrines contributing to the

likelihood are in W, and the integral of the intensity is again over W, but nuclei in T can affect the likelihood, since their existence can generate edges which intersect W.

3.3.1. *Algorithmic details and stability properties* The updating scheme of Section 2.3 is extended with systematic updates for λ by Gibbs sampling and for x_T by moving, creating and destroying nuclei, as described in Section 3.2

Note that, for specificity, we use a systematic updating scheme; however Propositions 3 and 4 below are true for a random updating scheme also.

Proposition 3. *The hybrid algorithm as specified above is aperiodic and irreducible.*

Proof. The support of the posterior is $D = \{(x_T, z) : \mathbf{1}_{(x,z)} = 1\} \times (0,\infty)^5$, and irreducibility on D and aperiodic is straightforwardly verified.

Because of the complicated x_T-updates, it seems hard to establish Harris recurrence by the result in [3] used in the proof of Proposition 1. But note that the Hastings ratio in (7) for a birth proposal is smaller than

$$\frac{|T|\lambda}{n(x_T)+1}\left(\frac{\alpha+\beta}{\beta}\right)^k e^{\beta|W|}.$$

Suppose we restrict the priors for $\lambda, \sigma^2, \alpha, \beta, \gamma$ to compact intervals $I_\lambda, I_{\sigma^2}, I_\alpha, I_\beta, I_\gamma \subset (0,\infty)$. These restrictions can, as with Proposition 2, be chosen to have negligible effect on practical results. An upper bound on λ which is large compared with a realistic upper limit on the number of territories in S will have no effect; the other constraints are discussed before Proposition 2.

These restrictions imply that the Hastings ratio for a birth proposal is always strictly smaller than $c/(n(x_T)+1)$ for some constant c (i.e. independent of $(x_T, z, \lambda, \sigma^2, \theta) \in D_{\mathrm{mod}}$), and we obtain not only Harris recurrency but geometric ergodicity for the modified hybrid algorithm:

Proposition 4. *The modified hybrid algorithm as specified above is geometrically ergodic.*

Proof. The proof is similar to the proof of Proposition 3.3 in [7], so we give only a sketch. Note that the support of the posterior is now $D_{\mathrm{mod}} = \{(x_T, z) : \mathbf{1}_{(x,z)} = 1\} \times I_\lambda \times I_{\sigma^2} \times I_\alpha \times I_\beta \times I_\gamma$. We verify that the associated Markov chain $M_i,\ i = 0, 1, \ldots,$

with states in D_{mod} is in fact V-geometrically ergodic, since the following geometric drift condition is satisfied for the function

$$V(x_T, z, \lambda, \sigma^2, \theta) = M^{n(x_T)}$$

where $M = 1 \vee c$: there exists real constants $a < 1$ and b, and a small set $C \subset D_{\mathrm{mod}}$, so that for all $(x_T, z, \lambda, \sigma^2, \theta) \in D_{\mathrm{mod}}$,

$$E(V(M_1)|M_0 = (x_T, z, \lambda, \sigma^2, \theta)) \leq aV(x_T, z, \lambda, \sigma^2, \theta) + b\mathbf{1}_C(x_T, z, \lambda, \sigma^2, \theta))$$

(see Theorem 15.0.1 in [14]).

Because of the upper bound $c/(n(x_T)+1)$ on the Hastings ratio for birth proposals, following the proof in [7] and using that each iteration in the algorithm starts with an x_T update, we obtain

$$E(V(M_1)|M_0 = (x_T, z, \lambda, \sigma^2, \theta)) \leq ((1 + 1/M)/2)V(M_0) \quad \text{for } n(x_T) \geq K$$

where V is specified as above and $K < \infty$ is chosen sufficiently large. As for the small set, take any integer $m \geq K$ and set $C = \{(x_T, z, \lambda, \sigma^2, \theta) \in D_{\mathrm{mod}} : n(x_T) < K\}$. For $(x_T, z, \lambda, \sigma^2, \theta) \in C$, by continuity of densities, compactness of $I_\lambda, I_{\sigma^2}, I_\alpha, I_\beta, I_\gamma$, and boundedness of the set $\{(x_T, z) : \mathbf{1}_{(x,z)} = 1, n(x_T) < K\}$, there is a uniform lower bound $\kappa > 0$ on the transition probability for updating λ, σ^2, θ and moving a nucleus. Define the probability measure

$$Q(F) = \mathbf{1}[(x_T, z, \lambda, \sigma^2, \theta) \in F \Rightarrow x_T = \emptyset, z_T = \emptyset]$$

for events $F \subseteq D_{\mathrm{mod}}$, where z_T is the component of z corresponding to the disturbances of x_T. Then by similar arguments as in the proof of Proposition 3.2 in [7], using that $n(x_T)/c$ is a lower bound on the Hastings ratio for death proposals, we obtain for any $(x_T, z, \lambda, \sigma^2, \theta) \in C$ and any event $F \subseteq D_{\mathrm{mod}}$,

$$\begin{aligned}
& P^m((x_T, z, \lambda, \sigma^2, \theta), F) \\
\geq\ & [\kappa n(x_T)/(2c)][\kappa(n(x_T) - 1)/(2c)] \cdots [\kappa/(2c)][\kappa/2]^{m-n(x_T)} Q(F) \\
=\ & [\kappa/2]^m n(x_T)!/c^{n(x_T)} \\
\geq\ & [\kappa/(2(c \vee 1))]^m Q(F).
\end{aligned}$$

Hence, by definition of small sets [14], we conclude that C is small.

3.3.2. *Results* The results shown here are based on the same data as in Section 2.3, but with S strictly larger than the observation window W. Specifically, we take S to be a rectangle extending beyond W by 1km in each direction. In Figure 7 the outer rectangle shown is S, the inner one is W; the region between them is T.

The prior for the parameters α, β, γ is the same as in Section 2.3. The prior for the tessellation now needs to specify the joint distribution of λ and σ and of the unknown nuclei and displacements; this is given by (5), with the values of a_1 and a_2, and the upper bound on σ, again as in Section 2.3.

The posterior here is based on 2 MCMC runs, each of 60,000 iterations with a burn-in of 10,000 iterations. Figure 7 shows a typical reconstruction. Note the way in which the nuclei in T can influence the likelihood for the point process in W; for example, the nucleus just below the bottom edge of W leads to edges inside W that run close to some of the observed latrines. Note also that perturbations of vertices outside W may be important; for example, the one vertex that lies above W controls the position of an edge that is mainly inside W and which again runs close to some latrines.

To give an indication of the uncertainty in the reconstruction, Figure 8 shows samples from the posteriors for just two of the cells. For the cell nearer the centre (the same one as in Figure 5), allowing for edge effects makes little difference, as expected. For the cell near the top of W, however, acknowledging the possibility of other nuclei, in T, gives some indication of the uncertainty in its outer boundary. With the region T used here, there is still the possibility that this particular cell reaches the boundary of S; if necessary, a larger region T could be used to eliminate this effect, at added computational cost. Finally, to summarise the inference that can be made about the whole process of nuclei, Figure 9 shows the expected intensity of nuclei over T, in nuclei per square kilometre (along with the know positions of the nuclei in W). The surface is shown unsmoothed, and some of the fine-scale variation is due to sampling variation in the MCMC analysis; nevertheless, there are some discernible features. Specifically, there is a low intensity near to W except at a few places where there are latrines observed close to the edge of W, in particular near the centre of the top edge of W, towards the bottom of the left edge, and towards the right of the bottom edge. The algorithm moves rather quickly around the space of possible configurations x_T, essentially because the data are rather uninformative about x_T. The birth and death

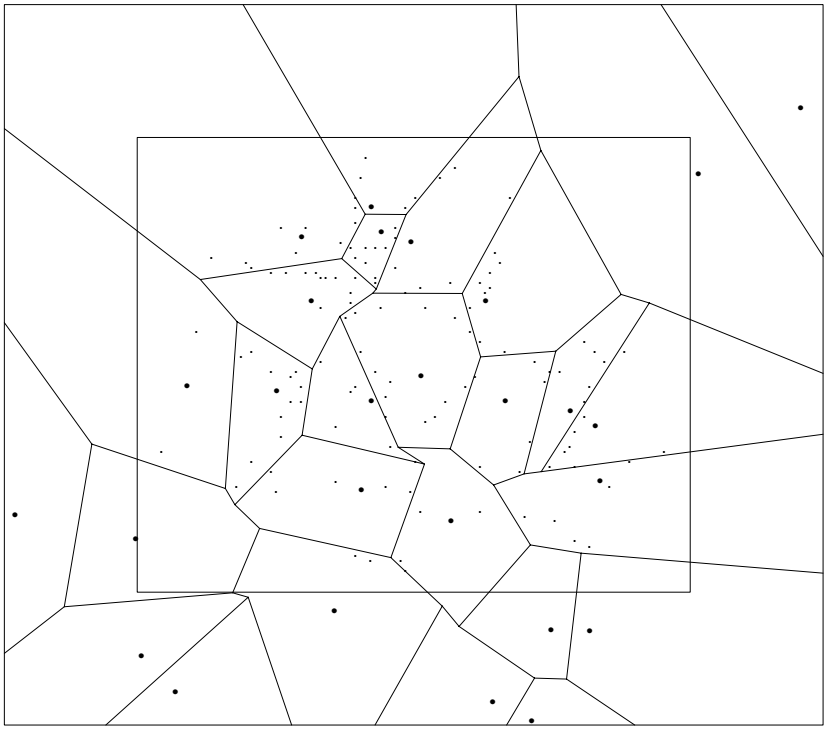

FIGURE 7: An example of a reconstruction allowing for edge effects

steps in the algorithm have acceptance rates of 35 to 45%. (The other types of steps, updating z and the various parameters, have acceptance rates which can be 'tuned' to efficient levels.)

3.4. Reconstruction of the tessellation in a noisy blurred image

In this example we consider the reconstruction of the microscopic grain structure in a two-dimensional cross-section through a sample of metal, using a noisy, grey-scale image. The image used here is shown in Figure 10; it is a 64×64 pixel section of a larger image, on a grey scale with 256 levels. We write y_{ij} for the observed value in the image in pixel (i, j) for $i = 1, \ldots, n_r$ and $j = 1, \ldots, n_c$, where $n_r = n_c = 64$.

In addition to the image shown in Figure 10, we make some use of an empirical, algorithmic reconstruction of the image. This reconstruction is a binary image, at the same resolution as the observed image, partitioning the region into grains of irregular shape, with edges that are a single pixel wide, and is shown in Figure 11.

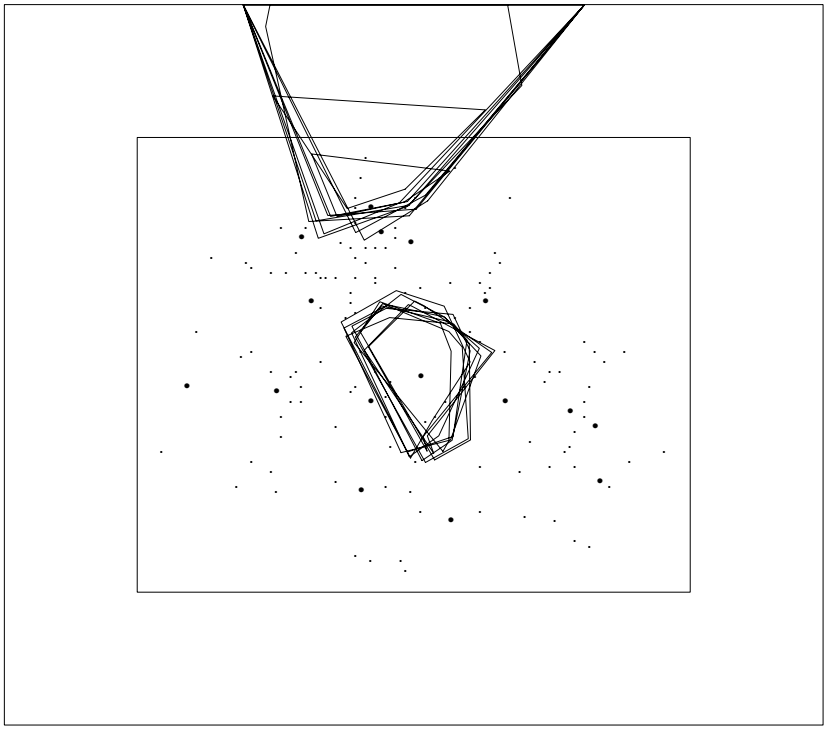

FIGURE 8: Posterior samples for two cells

The reconstruction is computationally easy to obtain, but incorporates no notion of uncertainty. Furthermore, it does not incorporate our expectation that, because of the mechanism by which the fine-scale structure in the metal is formed, the grain structure is likely to be close to a Voronoi tessellation. Thus we aim to reconstruct the grains using a deformed Voronoi tessellation as a representation of the (unobserved) true structure.

To specify a likelihood for this problem, we first consider the observed marginal distribution for y_{ij}, considering separately the two cases where the empirical reconstruction does or does not indicate an edge in a given pixel. Let $p_1(\cdot)$ and $p_0(\cdot)$ respectively denote these two marginal distributions (defined on $0, \ldots, 255$), and let $\tilde{p}_1(\cdot)$ and $\tilde{p}_0(\cdot)$ be versions obtained by smoothing with a locally linear 'lowess' algorithm [4], and satisfying

$$\tilde{p}_k(y_{ij}) > 0, \quad k = 0, 1, \quad y_{ij} = 0, \ldots, 255.$$

Then exploration of the spatial structure in the data suggests the following model. We

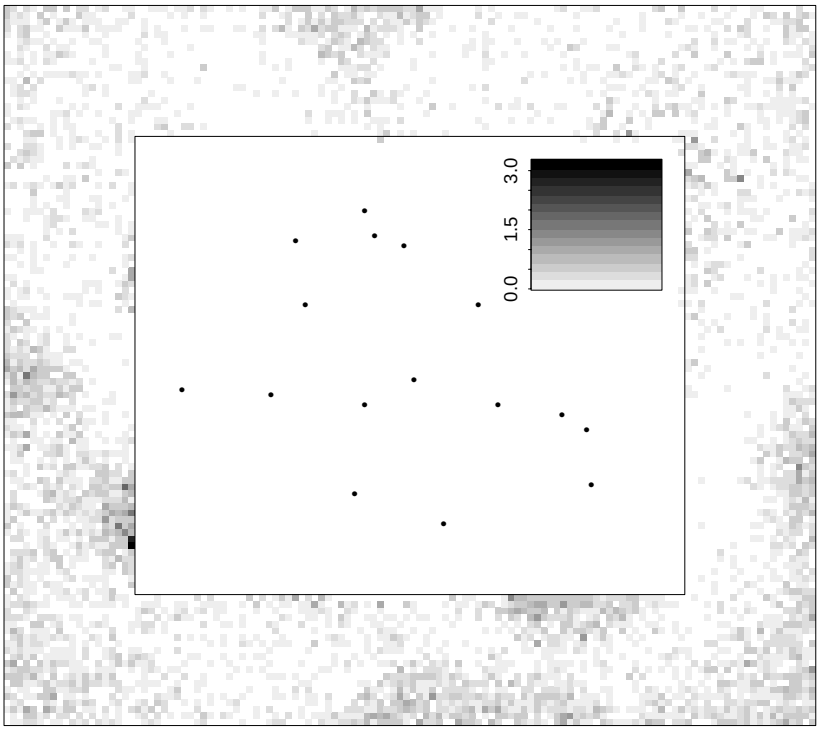

FIGURE 9: Posterior intensity surface for nuclei

take $W = [0, n_r] \times [0, n_c] \subseteq S$. Given a deformed tessellation $V(x, z)$ on S, define a binary $n_r \times n_c$ image $I(x, z)$, a discretized version of $V(x, z)$, by

$$I_{ij}(x, z) = \mathbf{1}\{E(x, z) \cap [i-1, i) \times [j-1, j) \neq \emptyset\}.$$

Then a simple 1-pixel "thickening" of the edges is obtained by defining a new image $\tilde{I}(x, z)$, where $\tilde{I}_{ij}(x, z)$ is the maximum of $I_{i'j'}(x, z)$ for $(i - i')^2 + (j - j')^2 \leq 1$, i.e. when either $(i', j') = (i, j)$ or (i', j') is a horizontal or vertical nearest-neighbour to (i, j). Finally, we define a model for the observed image by taking the associated random variables Y_{ij}, $i = 1, \ldots, n_r$, $j = 1, \ldots, n_c$, to be conditionally independent given (x, z) with

$$P(Y_{ij} = y_{ij}|x, z) = P(Y_{ij} = y_{ij}|\tilde{I}_{ij}(x, z)) = \tilde{p}_{\tilde{I}_{ij}(x,z)}(y_{ij}).$$

Hence the required likelihood is

$$L(y|x, z) = \prod_{ij} \tilde{p}_{\tilde{I}_{ij}(x,z)}(y_{ij}).$$

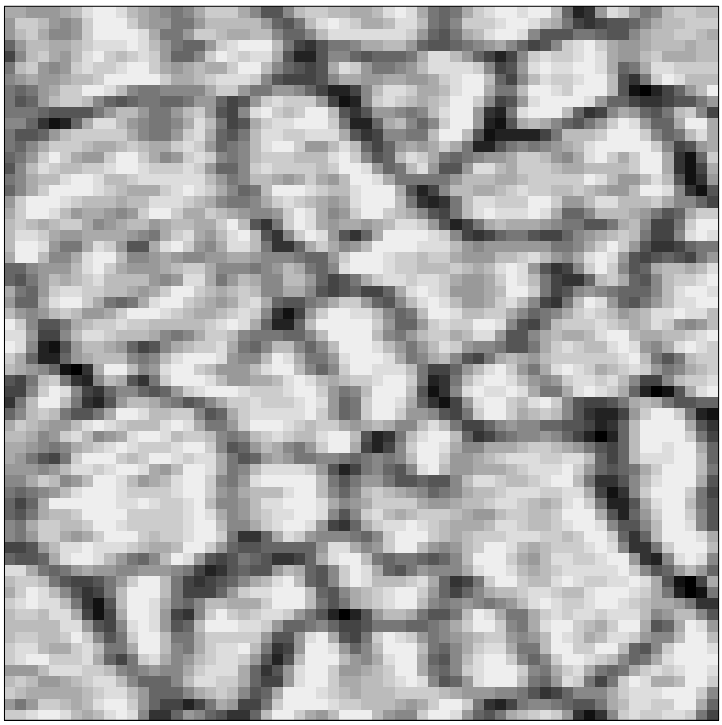

FIGURE 10: A grey-scale image of a cross-section of a sample of metal

Of course, this likelihood is rather simple, ignoring the local correlation structure in the observed image. But there is no obstacle, other than computational issues, to incorporating more sophisticated models into our framework. A natural alternative approach would be to approximate the distributions over the grey scale by Gaussian distributions, and write

$$Y = B(\mu_0(I - I(x,z)) + \mu_1 I(x,z) + \epsilon),$$

where ϵ consists of independent $\mathrm{N}(0,\psi^2)$ noise for some $\psi^2 > 0$, B is a $(n_r n_c) \times (n_r n_c)$ matrix representing "blurring" in the image, μ_0 and μ_1 are real parameters, and I is the $(n_r n_c) \times (n_r n_c)$ identity matrix. Then μ_0, μ_1 and ψ^2 could be readily estimated from the data.

3.4.1. *Algorithmic details and stability properties* The algorithm in this example is rather straightforward. Updates of λ are carried out by Gibbs sampling, as in Section

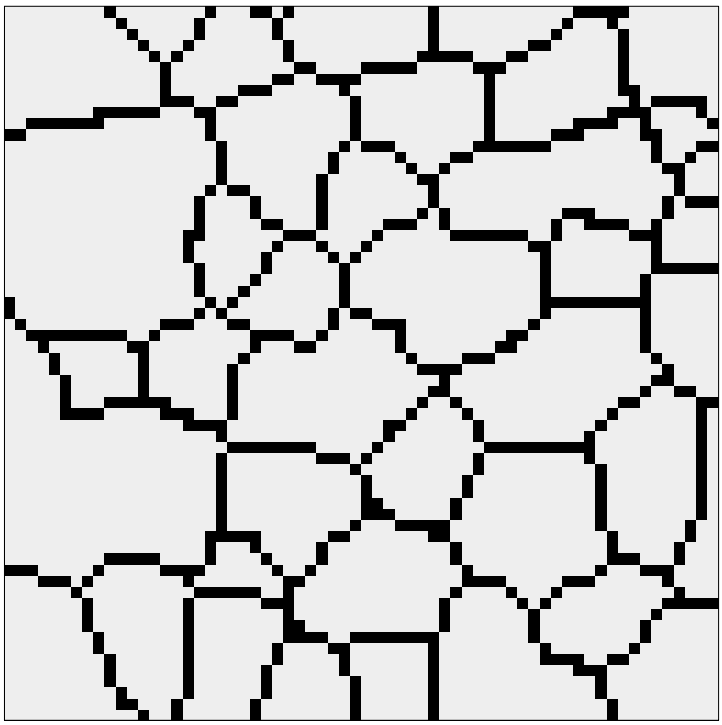

FIGURE 11: An empirical reconstruction of the image in Figure 10

3.2. Updates of z and x are carried out as in Sections 2.3 and 3.2 respectively; in each case, the required likelihood ratio $L(y|x',z')/L(y|x,z)$ can be calculated as a product of terms corresponding only to those (i,j) for which $\tilde{I}_{ij}(x',z') \neq \tilde{I}_{ij}(x,z)$.

As in Section 3.3.1, it is straightforward to see that the algorithm is aperiodic and irreducible. Furthermore, there is an upper bound on the Hastings ratio for a birth proposal, given by

$$\frac{|T|\lambda}{n(x_T)+1}\left(\frac{\max_{k,y_{ij}}\{\tilde{p}_k(y_{ij})\}}{\min_{k,y_{ij}}\{\tilde{p}_k(y_{ij})\}}\right)^{n_r n_c}.$$

Thus provided λ and σ^2 are restricted to compact intervals $I_\lambda, I_{\sigma^2} \subset (0,\infty)$, we immediately have that the modified algorithm is Harris recurrent and geometrically ergodic, by the same proof as in Proposition 4.

3.4.2. *Results* The results here are based on analysis of the image shown in Figure 10 (mapped on to the unit square).

The prior for λ and σ is specified jointly as in (5), with $e_1 = 400$, $e_2 = 10$ and

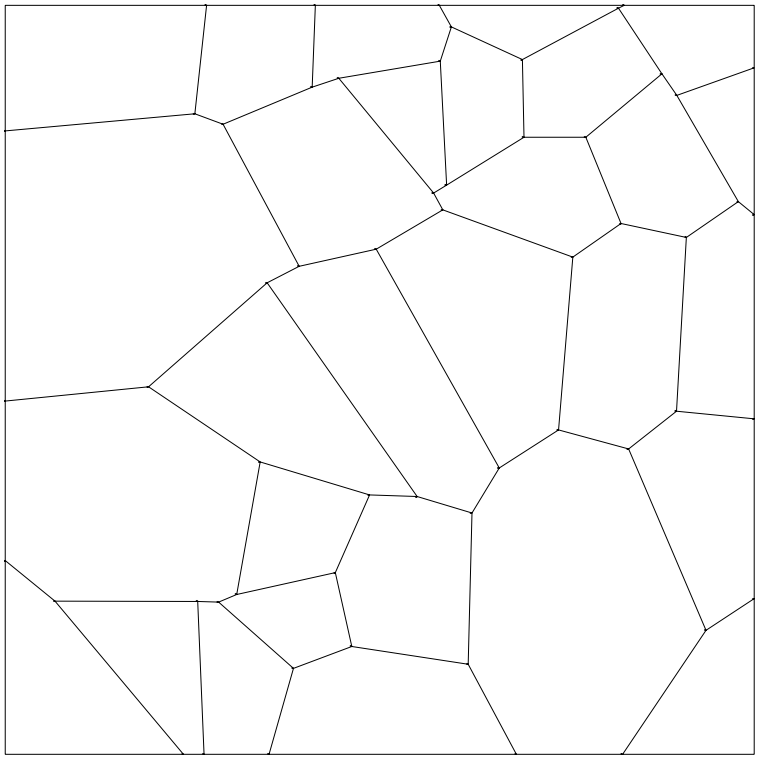

FIGURE 12: The posterior modal reconstruction

$a_1 = 8$, $a_2 = 2$. This gives λ a prior which is approximately the Gamma(400,10) distribution—the marginal prior is not exactly Gamma, because of the indicator term in (5)—representing reasonably strong prior information about the mean size of cells; similarly, the prior for σ^2 given λ is approximately InverseGamma(8,2/λ) representing strong prior belief that the perturbations will be reasonably small compared with the cells, i.e. that the tessellation will be close to Voronoi.

The posterior distributions illustrated here are based on 2 runs of length 150,000 iterations, each with a burn-in of 10,000 iterations. Figure 12 shows the single most likely tessellation obtained (i.e. the one with the highest posterior density). It can be seen that it successfully identifies the clearest 'grains' or microcrystals from Figure 10. Figure 13 shows another tessellation chosen for contrast as the one with the highest number of cells obtained. Clearly, some features are common to these tessellations, at least approximately; these include much of the bottom part of the figure. However, there are substantial differences, most noticeably in the top left

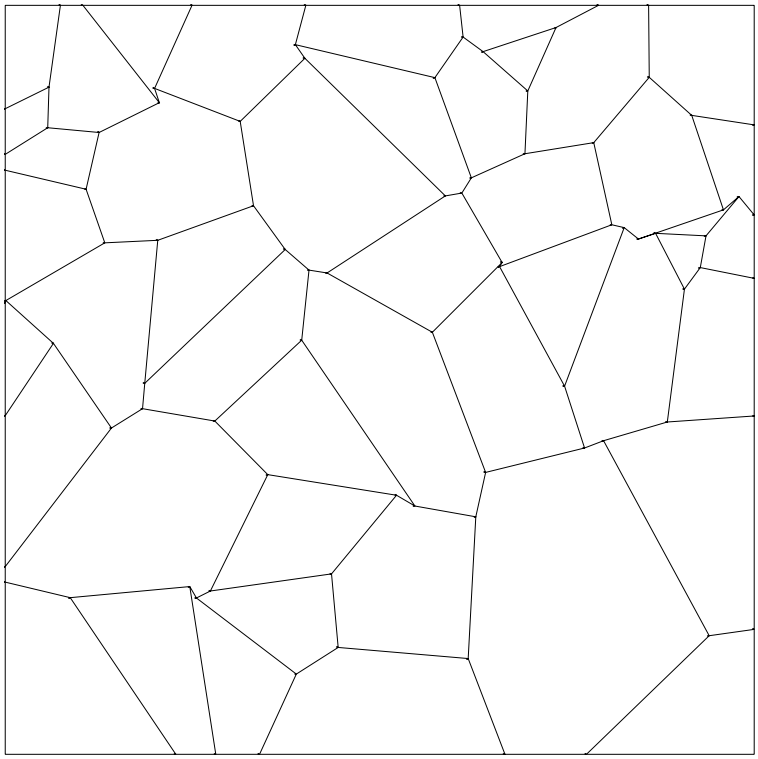

FIGURE 13: A reconstruction with a large number of cells

corner. Not surprisingly, the common features tend to coincide with obvious structure in the raw image, Figure 10, while the differences are in areas where the data are less informative. Similarly, in the clearer regions, the tessellations show some similarity with the algorithmic reconstruction in Figure 11, but there are substantial differences, due partly to the more constrained shapes of cells in the model-based reconstruction.

A summary of one aspect of the posterior variability of the fitted tessellations is given by the posterior distribution for the number of cells. Figure 14 shows this posterior and also the corresponding prior. Note that although the prior is rather informative about λ (as mentioned above) and hence about the number of cells, the posterior for the number of cells is much more concentrated: not surprisingly, we can learn a lot about the actual number of cells or microcrystals in the region represented by this image. Nevertheless, there remains considerable uncertainty in this number, and an advantage of our Bayesian approach is the ability to quantify that uncertainty. For comparison, the number of cells in the algorithmic reconstruction is 40, a value out in

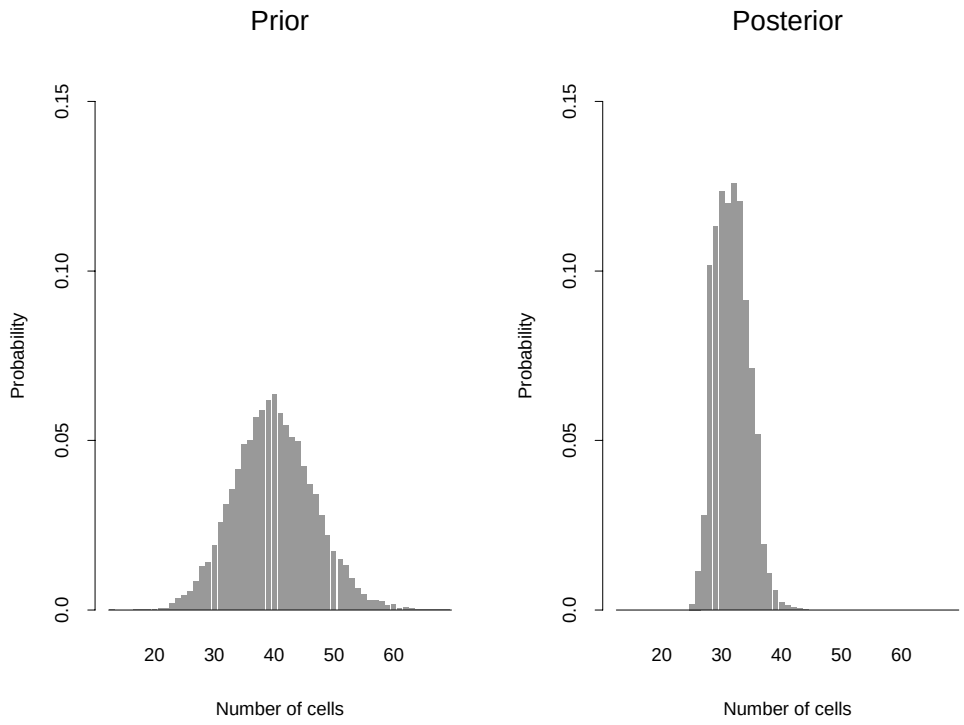

FIGURE 14: Prior and posterior distributions for the number of cells

the upper tail of the posterior distribution.

The algorithm mixes rather slowly in this example, since births and deaths here inherently give large changes in the tessellation and (since the whole tessellation is observed, albeit with noise) in the image. The acceptance rates for birth and death moves are each approximately 4%. (The acceptance rates for other types of moves can be 'tuned', and for these runs have acceptance rates of 20 to 30%.) Somewhat longer runs would be required for a definitive analysis of these data, but the results here are enough to indicate the feasibility of our approach. In addition, further analysis should perhaps incorporate both more detailed modelling of local correlation, along the lines suggested above, and allowance for edge effects, as in Section 3.3.

4. Generalizations

The class of deformed Voronoi tessellations is very flexible and widely applicable. There are, however, some obvious straightforward extensions to the model as discussed

below.

Remark 4.1. The specific distribution of individual vertex perturbations, approximately $\mathrm{N}(0, \sigma^2 I)$, does not really affect the details above (except in choosing a "conjugate" prior for σ^2, for convenience); any other bivariate distribution could be chosen. Having zero mean and circularity seem natural, so the obvious variations would be heavier, or perhaps lighter, even finite, tails.

Remark 4.2. Taking the individual perturbations z_j to be nearly independent, as we have done here, seems very natural when our prior simply represents the belief that the deformed tessellation is close to a Voronoi tessellation. If we have more specific prior information, this may well induce prior dependence in the z_js. This might be because we have prior knowledge about the shapes of cells, which will lead to dependence between perturbations of vertices of the same cell, or because we have information about the mechanism causing the perturbation from an underlying Voronoi model. For example, if we believe that an image of a Voronoi tessellation has been deformed by some smooth transformation of the plane, this might lead to positive dependence between perturbation of vertices that are close together. It would also be natural to incorporate such dependence in the proposal distributions for z_js.

Remark 4.3. Let $\{v_1, \ldots, v_m\}$ denote the set of boundary vertices of $V(x)$. As mentioned in Section 1, our approach may easily be extended to perturb these vertices into $v'_j = z_j + v_j, j = 1, \ldots, m$; let $\partial z = \{z_1, \ldots, z_m\}$ denote the set of such perturbations. It is then natural to add the condition that the ordering of boundary vertices is preserved. More precisely, consider any continuous parametrisation $s(t),\ 0 \le t < 1$, of the closed curve ∂S. Then we may require that

(C4) if $v_j = s(t_1),\ j = 1, \ldots, m$, where $0 \le t_1 < \ldots < t_m < 1$, and $v'_j = s(t'_j),\ j = 1, \ldots, m$, then each $v'_j \in \partial S$, and $0 \le t'_i < \ldots < t'_m < t'_1 < \ldots < t'_{i-1} < 1$ for some integer i with $1 \le i \le m$.

This condition does not depend on the choice of parametrisation of ∂S. The prior (5) may then be extended to a joint prior for ∂z and $(x, z, \lambda, \sigma^2)$ so that (C4) is satisfied almost surely. Finally, the MCMC algorithm in Section 3.2 can be straightforwardly extended to include updating of ∂z.

Remark 4.4. The starting tessellation itself need not be Voronoi. The key requirements are that there should be some way of defining the topology of the tessellation, i.e. what vertices exist, and how they are connected, and that the tessellation boundaries themselves should be defined purely by the positions of the vertices and by the topology — essentially that cells should be polygonal. The nuclei are just a set of objects that define the topology — they need not be points in S as in Section 3. Nor need they be in one-to-one correspondence with the cells; the sets $\mathcal{V}_+(\xi|x \cup \{\xi\})$ and $\mathcal{V}_-(\xi|x \cup \{\xi\})$ can simply be defined as the vertices added or removed, respectively, when adding nucleus ξ to existing nuclei x. For example, the undisturbed tessellation could be a line tessellation, with the "nuclei" being points in the space of lines in S. Then there are typically many more cells than nuclei; $\mathcal{V}_+(\xi|x \cup \{\xi\})$ is simply the set of new vertices where the line ξ intersects the existing lines in x, and $\mathcal{V}_-(\xi|x \cup \{\xi\})$ is always empty. If the prior for x is based on a stationary Poisson line process restricted to the space of lines intersecting S (see e.g. [27]), then very little of the detail in Section 3 need change. In fact this model would be much easier to program than our posterior based on deformed Voronoi tessellations.

Acknowledgements

This research was initiated in November 2001 when PGB was visiting JM at the Department of Mathematical Sciences, Aalborg University. The hospitality and financial support of this department is gratefully acknowledged. JM was supported by the European Union's research network "Statistical and Computational Methods for the Analysis of Spatial Data, ERB-FMRX-CT96-0096", by the Centre for Mathematical Physics and Stochastics (MaPhySto), funded by a grant from the Danish National Research Foundation, and by the Danish Natural Science Research Council.

We are grateful to Joachim Ohser for access to the data in Section 3.4.

References

[1] Blackwell, P. G. (2000). Inference for generalized tessellation processes. Research Report 504/00, Department of Probability and Statistics, University of Sheffield.

[2] Blackwell, P. G. (2001). Bayesian inference for a random tessellation process. *Biometrics* **57,** 502–507.

[3] Chan, K. S. and Geyer, C. J. (1994). Discussion of the paper 'Markov chains for exploring posterior distributions' by Luke Tierney. *Annals of Statistics* **22,** 1747–1747.

[4] Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* **74,** 829–836.

[5] Dagpunar, J. S. (1989). An easily implemented Generalized Inverse Gaussian generator. *Communications in Statistics B: Simulation and Computing* **18,** 703–710.

[6] Gelman, A. and Meng, X.-L. (1998). Simulating normalizing constants: from importance sampling to bridge sampling to path sampling. *Statistical Science* **13,** 163–185.

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

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

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

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

[11] Grenander, U. (1993). *General Pattern Theory*. Oxford University Press, Oxford.

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

[13] Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions, volume 1, second edition.* Wiley, New York.

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

[15] Møller, J. (1989). Random tessellations in $\mathbf{R}^d$. *Advances in Applied Probability* **21,** 37–73.

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

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

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

[19] Møller, J. (1999). Topics in Voronoi and Johnson-Mehl tessellations. In *Stochastic Geometry: Likelihood and Computation.* ed. O. E. Barndorff-Nielsen, W. S. Kendall, and M. N. M. van Lieshout. No. 80 in Monographs on Statistics and Applied Probability. Chapman and Hall/CRC, Boca Raton. pp. 173–198.

[20] Møller, J. and Skare, Ø. (2001). Bayesian image analysis with coloured Voronoi tessellations and a view to applications in reservoir modelling. *Statistical Modelling* **1,** 213–232.

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

[22] Okabe, A., Boots, B., Sugihara, K. and Chiu, S. N. (2000). *Spatial Tessellations. Concepts and Applications of Voronoi Diagrams.* Second Edition. Wiley, Chichester.

[23] Robert, C. P. and Casella, G. (1999). *Monte Carlo Statistical Methods.* Springer-Verlag, New York.

[24] Roberts, G. O., Gelman, A. and Gilks, W. R. (1997). Weak convergence and optimal scaling of random walk Metropolis algorithms. *Annals of Applied Probability* **7,** 110–120.

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

[26] Roberts, G. O. and Rosenthal, J. S. (1998). Two convergence properties of hybrid samplers. *Annals of Applied Probability* **8,** 397–407.

[27] Stoyan, D., Kendall, W. S. and Mecke, J. (1995). *Stochastic Geometry and Its Applications* second ed. Wiley, Chichester.