Object detection and automatic measurements with a data-driven algorithm for marked point process optimization

Claire Coiffard Marre, Ségolen Geffray

HAL Id: hal-00975622
https://hal.science/hal-00975622v2

Preprint submitted on 10 Apr 2014

Object detection and automatic measurements with a data-driven algorithm for marked point process optimization

Long technical extended version

Claire Coiffard-Marre [*], Ségolen Geffray [†]

Université de Strasbourg, IRMA, UMR 7501

April 9, 2014

Abstract

This paper deals with object detection and automatic measurements from optical microscopy and scanning electron microscopy (SEM) 2-dimensional images. More specifically, we consider the problem of detecting and measuring circular cells and rectilinear thick fibers. We adopt the spatial point process viewpoint and follow the maximum a posteriori (MAP) principle to obtain an optimal object configuration. We design a novel data-driven version of the data term in the MAP criterion in the fiber case, of the reference spatial marked point process and of the Markov Chain Monte-Carlo (MCMC) sampler, both in the cell case and in the fiber case. To this purpose, information from the image under consideration is incorporated by means of either the Radon transform or the circular Hough transform. We apply the proposed method to a real cell optical microscopy image and to a real fiber SEM image. Our results show that the proposed algorithm is fast, reliable and stable. We also produce the required characterization of the object morphology.

Keywords: Spatial marked point processes, maximum a posteriori, reversible jump Markov Chain Monte-Carlo, Simulated annealing, Radon transform, circular Hough transform.

1   Introduction

Images are a major source of information when dealing with microscopic objects. For example, think of cells or of polymeric fibers which are roughly of the order of the micrometer. In polymer engineering, systematic measurements are made repeatedly from polymer scanning electron microscopy

---

[*]Email: claire.coiffard@math.unistra.fr

[†]Email: segolen.geffray@math.unistra.fr

(SEM) images, as outlined in Ahirwal et al. [1], Lavielle et al. [14] or Shin et al. [19] for morphology evaluation purposes. In cell biology, the authors of Vigneron et al. [24] developed an automated low-level method of cell detection and counting from optical microscopic images. With these two examples in mind, a crucial goal for biomaterial engineering consists of automatically detecting such objects in a given image and of being able to characterize their location and morphology. Object detection from image data can be done using different techniques. Low-level image analysis methods present the major drawback of being sensitive to local image imperfections such as noise, poor illumination, blur... Moreover, once the detection is obtained with such a low-level image analysis method, it may be tedious to deduce object measurements. As a consequence, we turn to high-level methods in which an image I is considered as an imperfect (noisy, blurred, non-uniformly illuminated,...) observation of a random process X in an observation window $\mathcal{X}$. When the objects of interest can be modelled as simple parametric shape such as circles, squares, triangles, thick straight lines, then a suitable class of random processes for that modelling purpose is the class of spatial marked point processes. We refer the reader to the books by Møller and Waagepetersen [16] or by Van Lieshout [23] for a detailed presentation of spatial point processes, their sampling and practical applications. Object detection and counting can be made by fitting a marked point process to the image as largely demonstrated in the literature. We recommend the book by Descombes et al. [5] and the numerous references therein for applications of spatial point processes to practical image analysis situations.

This paper is concerned with providing an accurate, reliable and robust determination of cell and fiber location and morphology from microscopy images within a reasonable computation time and complexity. Specifically, we aim at detecting cell and fiber objects from microscopy images and at estimating the cell number and the cell radius distribution on the one hand, and the fiber width and orientation distributions on the other hand. We model cells as circles and fibers as thick straight lines within the framework of an appropriate spatial marked point process. We adapt the marked point process method presented by Descombes et al. [5] for its ability to provide straightforward automatic measurements once the object detection is done. Two kinds of inferential approaches may be distinguished when fitting a spatial marked point process to a given image, the Bayesian approach on the one hand, and the so-called detector approach according to by Descombes et al. [5] on the other hand. The Bayesian approach has been used successfully in a number of practical situations and recently in Konomi et al. [10]. However, the Bayesian method requires to specify a likelihood for the grey level intensity of pixels within and outside the objects. This makes the Bayesian method very sensitive to local defects in the image such as noise, blur or illumination gradient and also to the fact that grey level intensities may vary from on object to another which is the case in practice (see Figure 1). For these reasons, we adopt the detector approach as inferential method. Instead of specifying a likelihood for the grey level intensity of pixels within and outside the objects, one only specifies a so-called data term which accounts for the plausibility of a given object configuration with respect to the image at hand. The detection then consists of finding the optimal object configuration which is done by maximizing an objective function. Due to the intractability of the resulting non-convex optimization problem, a simulated annealing algorithm (Kirkpatrick et al. [9]) together with a Reversible Jump Markov Chain Monte-Carlo (RJMCMC) sampler (Green [7]) are used. This method

has demonstrated its wide applicability. However, the algorithm convergence to equilibrium may be really slow and the convergence may be somewhat difficult to diagnose. To speed up the algorithm, Brooks et al. [4] developed a general method for finding efficient proposals for RJMCMC algorithms. Their proposals are constructed from Taylor expansions of acceptance probabilities with the drawback of being computationally demanding. More specific data-driven versions for the RJMCMC algorithm have already been considered by Tu and Zhu [21]. Their paper is segmentation oriented - and also Bayesian oriented. Their method transposed to our cell and fiber detection context would use as proposal distribution a finite mixture of Gaussian distributions the components of which are centered at the local modes of the Radon or Hough transform. Though interesting, their method increases the number of unknown model parameters since numerous additional variance components have to be calibrated making the calibration step tedious.

We present a novel data-driven version of both the objective function in section 2 and of the optimization algorithm in section 3. To this aim, information from the image is incorporated by means of either the Radon or the circular Hough transform. We show in section 4 that our method, though quite simple, leads to improved convergence properties of the algorithm while producing accurate, reliable and robust results. Some technical elements proving the convergence of the proposed algorithm to the target are postponed in the Appendix.
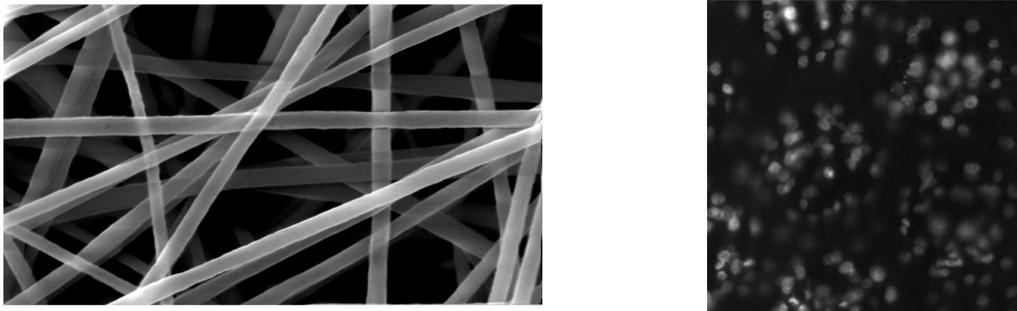

Figure 1: A Scanning Electron Microscopy image of polymer fibers and an optical microscopy image of cells. Image courtesy by Guy Schlatter, ICPEES, University of Strasbourg, France.

.

2 Objective function to be optimized

Object detection is obtained by computing the optimal object configuration according to the MAP criterion:

$$\overline{\mathbf{x}^m} = \mathrm{argmax}_{\mathbf{x}^m \in \Xi} \ P(I|X^M = \mathbf{x}^m)P(X^M = \mathbf{x}^m). \tag{1}$$

The first term $P(I|X^M = \mathbf{x}^m)$ is the so-called data term and represents the probability of obtaining image I under the spatial marked point process model specified for the spatial marked point process denoted as X^M and detailed below. The second term $P(X^M = \mathbf{x}^m)$ is the so-called a priori term and represents the probability of obtaining the object configuration $\mathbf{x}^m$ as a realisation of the random process X^M. Notice that the target distribution of which we are looking for global maxima is the a posteriori distribution hence the denomination of Maximum A Posteriori principle.

We first design the a priori term as a spatial marked point process used to model cells or fibers in subsection 2.1. Then, in subsection 2.2, we design the data term according to the detector approach. We then derive the objective function according to the maximum a posteriori principle. Our main preoccupation is to incorporate information from the image data to make the algorithm efficient, robust and reliable.

2.1 A priori term

We first introduce some useful mathematical formalism. Let λ^{Leb} be the Lebesgue measure in $\mathbb{R}^2$ and let $\mathcal{B}$ be the associated Borel σ-algebra on $\mathbb{R}^2$. We assume that the location space $\mathcal{X}$ is a compact subset of $\mathbb{R}^2$ with positive measure. Let X be a spatial point process on $\mathcal{X}$ so that on can think of X as a random countable subset of $\mathcal{X}$. A point configuration $\mathbf{x}$ is an unordered sequence of locations in $\mathcal{X}$. Let us denote as $n(\mathbf{x})$ the cardinality of $\mathbf{x}$. When $n(\mathbf{x}) = n$ for some n in $\mathbb{N}^*$, then the configuration $\mathbf{x}$ may be written as $\mathbf{x} = \{x_1, ..., x_n\}$. We set $(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, \lambda_{\mathcal{X}}^{\text{Leb}})$ to be the natural restriction to $\mathcal{X}$ of $(\mathbb{R}^2, \mathcal{B}, \lambda^{\text{Leb}})$. In practice, the location space $\mathcal{X}$ is given by the image boundaries that is $\mathcal{X} = [0, x_{max}] \times [0, y_{max}]$.

We then assume that the mark space $\mathcal{M}$ is a compact subset with positive measure of $\mathbb{R}^d$ for some $d \geq 1$. The mark space $\mathcal{M}$ is supposed equipped with the corresponding σ-algebra $\mathcal{F}_{\mathcal{M}}$ and a probability measure $P_{\mathcal{M}}$ having density p_M with respect to $\lambda_{\mathcal{M}}^{\text{Leb}}$ the restriction of Lebesgue measure on $\mathbb{R}^d$ to $\mathcal{M}$.

We introduce a marked point process on $\mathcal{X} \times \mathcal{M}$ as

$$X^M = \left\{ \{X_i^{M_i} := (X_i, M_i)\}_{i=1,...,n} : \{X_i\}_{i=1,...,n} \in X, M_i \in \mathcal{M}, i = 1, ..., n \right\}$$

meaning that a random mark M_i is attached to each random point X_i. A marked point is also termed as an object.

The mark space $\mathcal{M}$ naturally depends on the objects geometry. Cells are modelled as circles while fibers are modelled here as thick straight lines. In the cell case, the parametrization is straightforward. The object location is given by the circle center $x = (u, v)$ in $[0, x_{\text{max}}] \times [0, y_{\text{max}}]$ and the associated circle radius is given by r in $[0, r_{\text{max}}]$. In the fiber case, the object location is given by

(α, p, ε) in $\mathcal{X}$ as defined on figure 2, and the mark of the fiber is w in $\mathcal{M} = [w_{\min}, w_{\max}]$ which corresponds to the fiber width.

For n in $\mathbb{N}$, let Ξ_n be the set of all unordered configurations

$$\mathbf{x}^m = (\mathbf{x}, \mathbf{m}) = \{(x_1, m_1), ..., (x_n, m_n)\}$$

consisting of n marked points (x_i, m_i) in $\mathcal{X} \times \mathcal{M}$ with Ξ_0 being the empty configuration. For simplicity, we denote as $n(\mathbf{x})$ the cardinality of a configuration $\mathbf{x}^m$ since the number of marked points corresponds to the number of unmarked points. The configuration space can be written as $\Xi = \bigcup_{n=0}^{\infty} \Xi_n$ and is equipped with the σ-algebra $\mathcal{F}_\Xi$ defined by the mappings

$$\varphi_{A,n} : \{x_1^{m_1}, ..., x_n^{m_n}\} \rightarrow \sum_{i=1}^{n} I(x_i^{m_i} \in A)$$

that count the number of marked points in Borel sets $A \subseteq \mathcal{X} \times \mathcal{M}$.
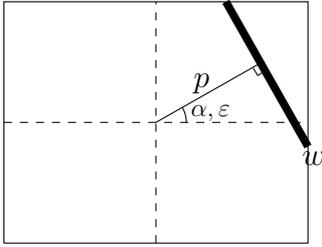
- $\alpha \in [0, \pi]$
- $p \in [0, d_{\max}(\alpha)]$
- $\varepsilon \in \{-1, 1\}$
- $w \in [w_{\min}, w_{\max}]$

Figure 2: Fiber parametrization

The Poisson process is a model for "no interaction" situations or of complete spatial randomness since for any disjoint subsets A and B of $\mathcal{X}$, the events $\{X \in A\}$ and $\{X \in B\}$ are independent. In many applications, the reference Poisson process is the homogeneous Poisson process with independent marks having uniform distribution. It turns out that it is possible to get information about the likely localization and geometry of the objects in an image by using some image processing techniques for example. It seems relevant to use this information within the reference process definition, as in AlAwadhi et al. [2]. We follow the approach of Baddeley and Van Lieshout [3] and Van Lieshout [22] who first define a reference measure denoted as here as π_Λ that incorporates some features of the object distribution but not their mutual interactions and then set the full model as Gibbs point process defined by a Radon-Nikodym derivative with respect to π_Λ. The point process taken as reference process hereafter is the inhomogeneous marked Poisson process defined for $A \in \mathcal{F}_\Xi$ by

$$\pi_\Lambda(\mathbf{x}^m) =$$
$$\sum_{n \in \mathbb{N}} \frac{\exp(-\Lambda(\mathcal{X}))}{n!} \int_{(\mathcal{X} \times \mathcal{M})^n} I(\{(x_1, m_1), ..., (x_n, m_n)\} \in A) \prod_{i=1}^{n} \lambda(x_i) p_M(m_i) d\lambda_{\mathcal{X}}^{\mathrm{Leb}}(x_i) d\lambda_{\mathcal{M}}^{\mathrm{Leb}}(m_i) \quad (2)$$

5

where the integral term for $n = 0$ reads as $I(\emptyset \in A)$, where Λ is the intensity measure, where $\lambda(\cdot) = \frac{d\Lambda(\cdot)}{d\lambda_{\mathcal{X}}^{\text{Leb}}}$ is the intensity function and where p_M is the density of the distribution P_M on $\mathcal{M}$ with respect to $\lambda_{\mathcal{M}}^{\text{Leb}}$. This says that the measure π_Λ distributes the object locations independently in $\mathcal{X}$ according to a Poisson distribution with intensity measure $\Lambda(.)$ and that the corresponding shapes are chosen in $\mathcal{M}$ independently according to the distribution P_M. We design the intensity measure Λ and the mark distribution P_M so that they incorporate some knowledge from the image data according to some deterministic image analysis methods. We set the intensity measure Λ to be equal to the image circular Hough transform in the cell case and to the image Radon transform in the fiber case.

In the cell case, we set the mark distribution P_M to be equal to the radius distribution calibrated from the image in the following way. We compute the image circular Hough transforms for radii within a prespecified range $[r_{min}, r_{max}]$. We then extract the local maxima which correspond to potential circles. We calibrate the radius distribution as the distribution of the radii associated with the obtained potential circles.

We planned to calibrate the fiber width distribution in a similar way. However, in practice, the number of fibers detected from our image by image analysis techniques is too small to derive a corresponding width distribution in a reliable way. Consequently, we kept the fiber width distribution as uniform within a prespecified range $[w_{\min}, w_{\max}]$.

To emphasize the difference with the usual choice, we note that the unit rate homogeneous marked Poisson process with uniformly distributed independent marks is usually chosen as reference process which boils down to taking $\lambda(.) \equiv 1$ and $p_M(.) = I(. \in \mathcal{M})$ in equation (2). Another justification for our reference process choice will also appear in the next section.

Interactions between objects have to be taken into account in order to prevent from excessive object superposition. Contrarily to Poisson processes, Gibbs point process allow the modelling of dependence of spatial point patterns. They can be constructed by specifying a Radon-Nikodym derivative with respect to π_Λ. A repulsive pairwise interaction model which penalises overlap between objects is appropriate for our applications. These processes have densities proportional to

$$p(\mathbf{x}^m) = \prod_{x_i^{m_i} \in \mathbf{x}^m} \beta(x_i^{m_i}) \prod_{x_i^{m_i}, x_j^{m_j} \in \mathbf{x}^m : x_i^{m_i} \neq x_j^{m_j}} \varphi(x_i^{m_i}, x_j^{m_j})$$

with $\beta : \mathcal{X} \to [0, \infty)$ and $\varphi : \mathcal{X} \times \mathcal{X} \to [0, \infty)$. If the interaction function φ satisfies $\varphi \leq 1$, the process is repulsive and if in addition $\int_{\mathcal{X} \times \mathcal{M}} \beta(x_1^{m_1}) \lambda(x_1) \lambda_{\mathcal{X}}^{\text{Leb}}(dx_1) p_M(m_1) \lambda_{\mathcal{M}}^{\text{Leb}}(dm_1) < \infty$, then the process is locally stable which implies that the process is well defined and satisfies the hereditary property (that is $p(\mathbf{x}^m) > 0 \implies p(\mathbf{x}'^{m'}) > 0, \forall \mathbf{x}'^{m'} \subset \mathbf{x}^m$). The local stability condition will also be of use to obtain the convergence properties of the Monte-Carlo dynamics in section 3. For both cell and fiber images, we choose an area-interaction process. This choice is quite classical in the literature, see eg Descombes et al. [5] and the references therein. This process corresponds to the choices $\beta(.) \equiv \beta$ for a $\beta > 0$ and $\varphi(x_i^{m_i}, x_j^{m_j}) = \exp\left(-\gamma_p t\left(x_i^{m_i}, x_j^{m_j}\right)\right)$ with $\gamma_p > 0$ and with $t\left(x_i^{m_i}, x_j^{m_j}\right)$ representing the

overlap degree between the two objects $x_i^{m_i}$ and $x_j^{m_j}$ in the following way

$$t\left(x_i^{m_i}, x_j^{m_j}\right) = \frac{\text{Area}\left(x_i^{m_i} \cap x_j^{m_j}\right)}{\min\left(\text{Area}\left(x_i^{m_i}\right), \text{Area}\left(x_j^{m_j}\right)\right)} \ .$$

The area interaction process then has the following unnormalized density with respect to π_Λ:

$$h_p(\mathbf{x}^m) = \beta^{n(\mathbf{x}^m)} e^{-\gamma_p s(\mathbf{x}^m)}, \tag{3}$$

where $s(\mathbf{x}^m) = \sum_{i<j} t\left(x_i^{m_i}, x_j^{m_j}\right)$ stands for the overall overlap degree in the current object configuration.

2.2 Data term determination

We now deal with the so-called data term. According to the detector approach, its expression is given by

$$h_d(\mathbf{x}^m) = \exp\left(-\gamma_d \sum_{x_i^{m_i} \in \mathbf{x}^m} v_d(x_i^{m_i})\right) \tag{4}$$

where $\gamma_d > 0$ is a weight parameter and where $v_d : \mathcal{X} \times \mathcal{M} \to \mathbb{R}$ is a so-called potential function which quantifies the confidence one may have in the fact that $x_i^{m_i}$ is an object actually present in the image. With such a specification, the plausibility of a configuration is sort of quantified by the addition of the plausibility of each object in the current configuration.

The potential function in $x_i^{m_i}$ is classically given by some statistical distance between the gray levels of the pixels within the object $x_i^{m_i}$ and of those in its neighbourhood. The statement leading to this definition is twofold. Firstly gray levels are homogeneous inside an object. Secondly gray levels inside an object contrast with grey levels in the object neighbourhood. To assess the distance between the grey levels inside and in the neighbourhood of the object, generally either the Bhattacharaya distance is used (Perrin et al. [17]), or a test statistic such as Student's test statistic (Lacoste [12]) or Kolmogorov-Smirnov's test statistic (Imberty and Descombes [8]). The potential function then aims at penalizing objects for which the preceding distance is too small.

We consider here separately the fiber case and the cell case.

In the cell case, we adopt the potential function used in Perrin et al. [17]. Let us denote as $x_i^{m_i}$ the circle in configuration $\mathbf{x}^m$ under consideration and let $b_i(\rho)$ be its neighbourhood defined as the outer pixels no distant than more than ρ pixels from the object boundary. As explained above, the statistical distance between the grey levels inside object $x_i^{m_i}$ and in its neighbourhood is set as the Bhattacharyya distance:

$$d_b(x_i^{m_i}, b_i(\rho)) = \frac{(\widehat{m_0} - \widehat{m_b})^2}{4\sqrt{\widehat{\sigma_0^2} + \widehat{\sigma_b^2}}} - \frac{1}{2}\log\left(\frac{2\sqrt{\widehat{\sigma_0^2}\widehat{\sigma_b^2}}}{\widehat{\sigma_0^2} + \widehat{\sigma_b^2}}\right) \tag{5}$$

where $\widehat{m}_0$ and $\widehat{m}_b$ are the means of the grey levels respectively inside the object $x_i^{m_i}$ and in its neighbourhood $b_i(\rho)$, where $\widehat{\sigma}_0^2$ and $\widehat{\sigma}_b^2$ are the variances of the grey levels respectively inside the object $x_i^{m_i}$ and in its neighbourhood $b_i(\rho)$. The potential function is then defined as

$$v_d(x_i^{m_i}) = \begin{cases} 1 - \left(\frac{d_B(x_i^{m_i},b_i(\rho))}{d_0}\right)^{1/3} & \text{if } d_B(x_i^{m_i}, b_i(\rho)), \\ \exp\left(-\frac{d_B(x_i^{m_i},b_i(\rho))-d_0}{d_0}\right) - 1 & \text{otherwise.} \end{cases} \qquad (6)$$

In the fiber case, we propose a new expression of the potential function which relies on the Radon transform. Indeed, up to our knowledge, there does not exist any implementation of spatial marked point process for thick straight lines detection. One possibility would consist of using a potential function expression similar to that used in the cell case or to that used in the segment detection case by Lacoste et al. [11]. A major drawback is the exploding computation time, since the thick straight lines have a very large number of pixels. Moreover, the presence of intersecting fibers in the neighbourhood of the fiber under consideration entails that the grey level intensities of the pixels in the neighbourhood of the considered fiber are inhomogeneous which makes irrelevant the use of (5) or of any similar distance between the grey levels inside and outside the fiber. Let us denote as $x_i^{m_i}$ the object in configuration $\mathbf{x}^m$ under consideration. Let $RT(x_i)$ be the value of the image Radon transform at the point having location x_i. We recall that the Radon transform is defined as the integral of the image grey levels intensity over straight lines. Consequently, the larger $RT(x_i)$ is, the more likely there is a line at location x_i. To incorporate some additional information about the mark m_i and get a more precise potential, we design a second term that accounts for the fact that the fibers are in light grey on a dark grey background. This term is defined as

$$k(x_i^{m_i}) = \frac{1}{n_p(x_i)} \max\left(1, \#\{\text{pixels of thick line } x_i^{m_i} \text{ having grey level intensity } < \tau\}\right), \qquad (7)$$

where $n_p(x_i^{m_i})$ is the number of pixels in $x_i^{m_i}$ and where the threshold τ needs to be calibrated. Then, the potential $v_d(x_i^{m_i})$ is obtained through a linear transform of

$$h_i = \frac{RT(x_i)}{k(x_i^{m_i})}$$

involving three additional parameters s, s' and V_{max} given by

$$v_d(x_i^{m_i}) = \begin{cases} V_{\max} & \text{if } h_i < s, \\ 1 - 2\frac{h_i-s}{s'-s} & \text{if } s \leq h_i \leq s', \\ -1 & \text{otherwise.} \end{cases} \qquad (8)$$

As a conclusion to this section, we are in position to derive the objective function to be optimized. The data term $P(I = \mathbf{y}|X^M = \mathbf{x}^m)$ is taken proportional to the unnormalized density $h_d(\mathbf{x}^m)$ being

defined by equation (4). The prior term $P(X^M = \mathbf{x}^m)$ is taken proportional to the unnormalized density $h_p(\mathbf{x}^m)$ being defined by equations (3) and (8). With these statements in mind, the optimization problem in (1) may be rewritten as

$$\overline{\mathbf{x}^m} = \mathrm{argmax}_{\mathbf{x}^m} h(\mathbf{x}^m), \tag{9}$$

where $h(\mathbf{x}^m) = h_d(\mathbf{x}^m)h_p(\mathbf{x}^m)$ is the objective function according to the MAP principle. The next section explains how to find a configuration maximizing the unnormalized density h.

3   Optimization

The task corresponding to the non-convex optimization problem in (9) consists of locating a global mode of $h(.)$. For that purpose, we use the simulated annealing algorithm. The simulated annealing (SA) algorithm produces a random sample from $h(.)$ which is biased overwhelmingly in favour of the most probable value $\overline{\mathbf{x}^m}$. Specifically, the SA algorithm consists of repeatedly sampling configurations under the process having unnormalised density $h^{1/T}(.)$ with T going to 0 as the number of iterations grows. Each distribution proportional to $h^{1/T}(.)$ has its mode at $\overline{\mathbf{x}^m}$, and as T decreases the mode becomes more and more exaggerated. Consequently, if we generate a random sample from each successive distribution, there is an increasing probability of producing the target $\overline{\mathbf{x}^m}$. We adopt the classical geometrical cooling scheme which boils down to considering a positive sequence of temperatures $(T_k)_{k \in \mathbb{N}}$ with $T_k = cT_0$ where T_0 stands for the initial temperature and where $0 < c < 1$. Both T_0 and c need to be calibrated. The simulated annealing algorithm is presented in algorithm 1. It involves a sampling step. Gibbs point processes are generally intractable because of the unknown normalizing constant of the likelihood. Therefore the sampling step of the SA algorithm motivates the use of a reversible jump Markov chain Monte-Carlo (RJMCMC) sampler also termed as Metropolis-Hastings-Green sampler. The construction of this sampler was given by Green [7] as an extension of the Metropolis-Hastings algorithm to general state-space. We refer the reader to the monographs by Robert and Casella [18] for a detailed presentation of MCMC methods.

Algorithm 1 SIMULATED ANNEALING

1: **Choose** an initial temperature $T = T_0$
2: **while** maximal number of iterations not reached **do**
3:     **Run** the sampler with unnormalised density $h(.)^{1/T}$
4:     **Change** $T \leftarrow cT$ with $c < 1$

Consequently, the optimization problem (9) moves to building an algorithm that simulates a Markov chain having as equilibrium distribution the measure π of our spatial marked point process

defined as

$$\pi(A) = \int_A h(\mathbf{x}^m)\pi_\Lambda(d\mathbf{x}^m), \qquad A \in \mathcal{F}_\Xi$$

and the realizations of which are configurations of marked points of $\mathcal{X} \times \mathcal{M}$. We build the Markov chain to be π-invariant, aperiodic and geometrically ergodic, as proved in the appendix.

The general Metropolis-Hastings framework simulates Markov chains induced by the transition kernel

$$P(\mathbf{x}^m, A) = \int_A \alpha(\mathbf{x}^m, \mathbf{x}'^{m'})Q(\mathbf{x}^m, d\mathbf{x}'^{m'}), \quad \mathbf{x} \in \Xi,\, A \in \mathcal{F}_\Xi,\, \mathbf{x} \notin A$$

and

$$P(\mathbf{x}^m, \{\mathbf{x}^m\}) = \int_{\mathbf{x}'^{m'} \neq \mathbf{x}^m} (1 - \alpha(\mathbf{x}^m, \mathbf{x}'^{m'}))Q(\mathbf{x}^m, d\mathbf{x}'^{m'})$$

where Q is the proposal kernel and $\alpha(.,.)$ is the acceptance probability. Two types of transitions are considered here. Adding a marked point to the current configuration is proposed with probability p_b and deleting a point from the current configuration is proposed with probability p_d. When relevant, the marked object y^{m_y} to be added to the current marked point configuration $\mathbf{x}^m$ is chosen with probability density $q_b(\mathbf{x}, y)p_M(m_y)$ with respect to $\lambda_\mathcal{X}^{\text{Leb}} \otimes \lambda_\mathcal{M}^{\text{Leb}}$. This means that the location of the new object and its mark are simulated independently. Note that the mark is simulated according to the same distribution as that of the marks of the reference process. When relevant, the object in location y to be deleted from current configuration location $\mathbf{x}^m$ is chosen with probability $q_d(\mathbf{x}, y)$. We write for $(\mathbf{x}, B)$ in $\Xi \times \mathcal{F}_\Xi$

$$Q(\mathbf{x}^m, B) = p_b Q_b(\mathbf{x}^m, B) + p_d Q_d(\mathbf{x}^m, B).$$

The so-called "birth" kernel is

$$Q_b(\mathbf{x}^m, B) = \int_{\mathcal{X} \times \mathcal{M}} I(\mathbf{x}^m \cup \{y^{m_y}\} \in B)q_b(\mathbf{x}, y)p_M(m_y)d\lambda^{Leb}(y)d\lambda^{Leb}(m_y) \qquad (10)$$

Obviously, the new object y^{m_y} needs to be chosen apart from $\mathbf{x}^m$ which happens here almost-surely. The so-called "death" kernel is

$$Q_d(\mathbf{x}^m, B) = \sum_{y^{m_y} \in \mathbf{x}^m} I(\mathbf{x}^m \setminus \{y^{m_y}\} \in B)q_d(\mathbf{x}, y) \qquad (11)$$

Each proposal move is either accepted of rejected according to an acceptance step detailed later. If the proposal is accepted, the chain goes to the new state given by the transition kernel. If the proposal is rejected, the chain remains in its current state. Such a transition kernel may be written

for A in $\mathcal{F}_\Xi$ as

$$P(\mathbf{x}^m, A) = p_b \int_{\mathcal{X} \times \mathcal{M}} q_b(\mathbf{x}, y) p_M(m_y) \alpha(\mathbf{x}^m, \mathbf{x}^m \cup \{y^{m_y}\}) I(\mathbf{x}^m \cup \{y^{m_y}\} \in A) d\lambda_{\mathcal{X}}^{\text{Leb}}(y) d\lambda_{\mathcal{M}}^{\text{Leb}}(m_y)$$

$$+ p_d \sum_{y^{m_y} \in \mathbf{x}^m} q_d(\mathbf{x}, y) \alpha(\mathbf{x}^m, \mathbf{x}^m \setminus \{y^{m_y}\}) I(\mathbf{x}^m \setminus \{y^{m_y}\} \in A)$$

$$+ I(\mathbf{x}^m \in A) \Big[1 - p_b \int_{\mathcal{X} \times \mathcal{M}} q_b(\mathbf{x}, y) p_M(m_y) \alpha(\mathbf{x}^m, \mathbf{x}^m \cup \{y^{m_y}\}) d\lambda_{\mathcal{X}}^{\text{Leb}}(y) d\lambda_{\mathcal{M}}^{\text{Leb}}(m_y)$$

$$- p_d \sum_{y^{m_y} \in \mathbf{x}^m} q_d(\mathbf{x}, y) \alpha(\mathbf{x}^m, \mathbf{x}^m \setminus \{y^{m_y}\}) \Big]$$

If the current configuration is the empty configuration, obviously the transition kernel cannot propose to delete a marked point. In this case, the above transition kernel becomes

$$P(\emptyset, A) = p_b \int_{\mathcal{X} \times \mathcal{M}} q_b(\emptyset, y) p_M(m_y) \alpha(\emptyset, \{y^{m_y}\}) I(\{y^{m_y}\} \in A) d\lambda_{\mathcal{X}}^{\text{Leb}}(y) d\lambda_{\mathcal{M}}^{\text{Leb}}(m_y)$$

$$+ I(\emptyset \in A) \Big[1 - p_b \int_{\mathcal{X} \times \mathcal{M}} q_b(\emptyset, y) p_M(m_y) \alpha(\emptyset, \{y^{m_y}\}) d\lambda_{\mathcal{X}}^{\text{Leb}}(y) d\lambda_{\mathcal{M}}^{\text{Leb}} \Big]$$

In the appendix, we determine the Green ratios associated with our choice of inhomogeneous marked Poisson process as reference. These Green ratios are given by

$$r_b(\mathbf{x}^m, y^{m_y}) = \frac{h(\mathbf{x}^m \cup \{y^{m_y}\}) p_d q_d(\mathbf{x} \cup \{y\}, y) \lambda(y)}{h(\mathbf{x}^m) p_b q_b(\mathbf{x}, y)}$$

and

$$r_d(\mathbf{z}, y) = \frac{h(\mathbf{z} \setminus \{y\}) p_b q_b(\mathbf{z} \setminus \{y\}, y)}{h(\mathbf{x}) p_d q_d(\mathbf{z}, y) \lambda(y)} .$$

Note both the scaling of q_b by $\lambda(.)$ and the absence of p_M in these expressions. This comes from the fact that the proposition kernel uses the mark distribution P_M of the reference process as mark proposition distribution.

An acceptance step then makes use of the Green ratios to determine in a stochastic manner whether or not the proposal should be accepted or rejected in which case the chain remains in its current state. The main steps of the Metropolis-Hastings-Green algorithm are summarized in Algorithm 2.

The definition of p_M has already been given in subsection 2.1. We now detail our choice of q_d and q_b. The death is proposed uniformly among the current configuration, which is a classical choice. This yields

$$q_d(\mathbf{x}, y) = \frac{1}{n(\mathbf{x})} .$$

The novelty of our algorithm lies in the birth proposal choice. In previous work (see eg Descombes et al. [5]), birth is proposed uniformly in $\mathcal{X}$. But it would make more sense to favour birth where an

---
Algorithm 2 METROPOLIS HASTINGS GREEN

1: **Let** $\mathbf{x}^m$ be the current configuration
2: **Propose** birth with probability p_b and death with probability p_d
3: **if** birth **then**
4:     **Propose** $y \sim q_b(\mathbf{x}, .)$
5:     **Propose** $m_y \sim p_M(.)$
6:     **Compute** Green ratio $r_b(\mathbf{x}^m, y^{m_y})$
7:     **Accept** y^{m_y} with probability $\alpha(\mathbf{x}^m, \mathbf{x}^m \cup \{y^{m_y}\}) = \min(1, r_b(\mathbf{x}^m, y^{m_y}))$.
8: **else**
9:     **Propose** $y \sim q_d(\mathbf{x}, .)$
10:     **Compute** Green ratio $r_d(\mathbf{x}^m, y^{m_y})$
11:     **Accept** y^{m_y} with probability $\alpha(\mathbf{x}^m, \mathbf{x}^m \setminus \{y^{m_y}\}) = \min(1, r_d(\mathbf{x}^m, y^{m_y}))$.
---

object is more likely to be. Since it is possible to get approximate information about the location of the objects in the image by image processing techniques, it seems logical to use this information for that purpose. In the cell case, we propose to use the distribution given by the normalized circular Hough transform obtained after taking the local maxima over a prespecified range of radii. Doing this may yield a distribution putting zero weight over quite large areas of the image domain, which may be seen quite often in practice since cell images may be very poorly contrasted. In this case, we advocate for rather using a finite mixture distribution having as two components the preceding normalized circular Hough transform (with an important weight eg 0.8) and the uniform distribution in $\mathcal{X}$ (with a minor weight eg 0.2) so that the normalized circular Hough transform dominates but the whole space is explored anyway. An example of point sampling according to this proposal is displayed in figure 3. In the fiber case, we use the normalized Radon transform. An example of point sampling according to this proposal is displayed in figure 4.

Here appears another benefit of our choice of inhomogeneous marked Poisson process as reference. Descombes et al. [5] mentioned the difficulty that proposing attractive configurations was not enough to improve the convergence of the algorithm. This was due to the fact that attractive configuration have small Green ratio and the proposition was therefore very likely to be rejected. We recall that in Descombes et al. [5], the reference Poisson process is taken as homogeneous with constant intensity equal to $\lambda > 0$. In this case, when death proposal is uniform in $\mathcal{X}$ and when birth proposal is generated according to a distribution having density denoted as $p(.)$ with respect to $\lambda_{\mathcal{X}}^{\mathrm{Leb}}$, the birth Green ratio writes down as

$$r_b(\mathbf{x}^m, y) = \frac{h(\mathbf{x}^m \cup \{y^{m_y}\})}{h(\mathbf{x}^m)} \frac{\lambda}{(n(\mathbf{x}) + 1)p(y)}.$$

The larger $p(y)$ is, the more likely there is an object in the image at location y. But the larger $p(y)$ is, the smaller r_b is. Though counter-intuitive at a first look, this entails that any reasonable proposition is likely to be rejected. To emphasize the difference with our algorithm, we recall that our reference
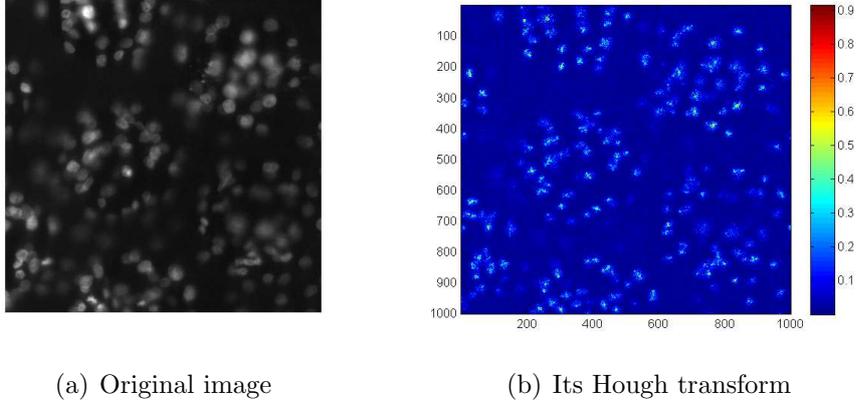
(a) Original image

(b) Its Hough transform
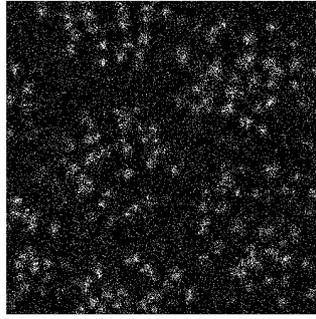
(c) A sampling example of 100000 points according to the proposed mixture

Figure 3: A non-uniform sampling example in the cell case.

Poisson process is inhomogeneous and consider the case where its intensity function is proportional to the birth distribution. We then get

$$r_b(\mathbf{x}, y) \propto \frac{h(\mathbf{x}^m \cup \{y^{m_y}\})}{h(\mathbf{x}^m)} \frac{1}{(n(\mathbf{x}) + 1)}$$

so that the intensity function compensates for the birth proposal density term and prevents the birth Green ratio from becoming too small when the proposal is very likely according to $p(.)$.

4   Results and discussion

In this section, we present the results obtained with the proposed methodology. In the cell case, we compare these results with the results of the reference algorithm that uses uniform birth and

(a) Original image

(b) Its Radon transform
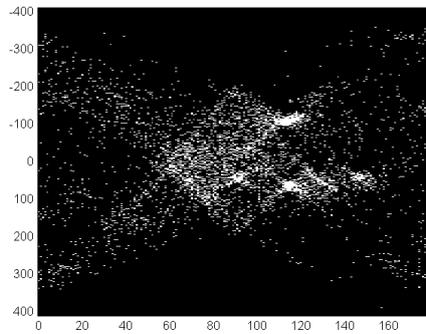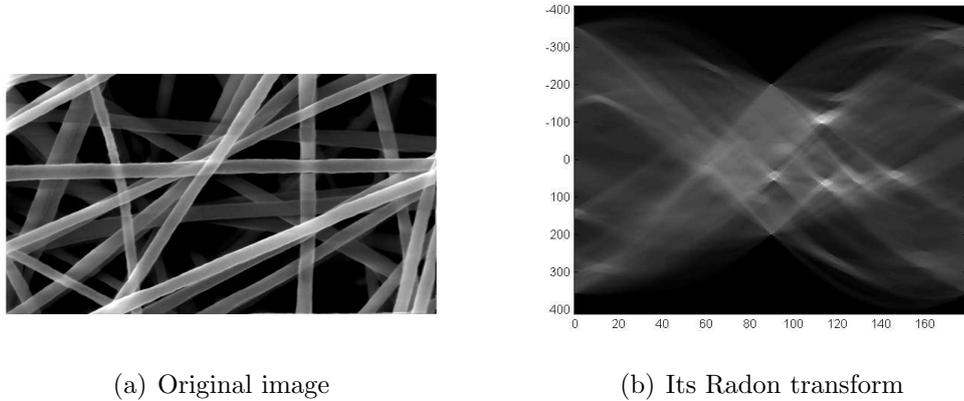
(c) A sampling example of 100000 points according to the normalized Radon transform

Figure 4: A non-uniform sampling example in the fiber case.

death kernels together with a unit rate homogeneous Poisson process as reference process. In the fiber case, we were unable to provide a similar comparison due to the lack of existing algorithm for thick straight line objects. We analyse the results in terms of visual quality of object detection and in terms of computation time and convergence to stability of our algorithm. We also provide the required morphology characterisation. The results presented hereafter are obtained with a dual-core laptop having 2.10 GHz CPU twice.

The parameters are calibrated by basic trial and error method since existing estimation methods did not give satisfying results. The issue of estimating the model parameters is left for future research. For simplicity, we take throughout $p_b = p_d = \frac{1}{2}$.

In the circle case, the parameters of the area interaction process are $\beta = 1$, $\gamma_d = 30$, $\gamma_p = 50$. The initial configuration is the empty configuration. The parameters specific to the simulated annealing part are $T_0 = 0.1$ as initial temperature, $c = 0.9999$ as multiplicative factor in the geometric cooling schedule and $N = 50000$ as prespecified iteration number. The circular Hough transform needed for our sampling step was computed within the radius range $[r_{\min}, r_{\max}]$ with $r_{\min} = 10$ and $r_{\max} = 25$

14

and then normalised to get a proper distribution. The location sampling distribution was then taken as the finite mixture having as two components the normalised circular Hough transform and the uniform distribution with respective weight 0.8 and 0.2. We ran the reference algorithm with the uniform distribution on $[r_{\min}, r_{\max}]$ as radius distribution with $r_{\min} = 10$ and $r_{\max} = 25$. The data term parameters were taken as $\rho = 15$ and $d_0 = 40$.

When implementing our method in the fiber case, the parameters of the area interaction process are taken as $\beta = 1$, $\gamma_d = 50$ and $\gamma_p = 1000$. The initial configuration is the empty configuration. The parameters specific to the simulated annealing part are $T_0 = 10$ as initial temperature, $c = 0.9999$ as multiplicative factor in the cooling schedule and $N = 100000$ as prespecified iteration number. The fiber width was simulated uniformly on $[w_{\min}, w_{\max}]$ with $w_{\min} = 15$ and $w_{\max} = 30$. The data term parameters were taken as $s = 1$, $s' = 5000$, $V_{\max} = 1$ and $\tau = 10$.

The detected objects are represented on the original image on figure 5 in the cell case and on figure 8 in the fiber case. A visual examination leads to the conclusion that our method provides a valuable object detection.

For the cell image, our algorithm runs in 43 minutes for 50000 iterations while the reference algorithm runs in 41 minutes with the same iteration number making the computation times similar. For the fiber image, our algorithm runs in 18 minutes for 100000 iterations.

The evolution of the number of detected cells is represented as a function of the number of iterations in the cell case on figure 6 and in the fiber case on figure 9. Noticeably, in the cell case, our algorithm converges quickly to equilibrium in a stable way which eases the convergence diagnosis.

The required morphology characterisation is provided in figure 7 in the cell case and in figure in the fiber case 10. Noticeably, in the cell case, our results illustrate the fact that the estimated radius distribution shares a similar shape with the mark proposal distribution. This highlights the importance of calibrating the proposal in a data-driven way when possible.

As a conclusion, our results illustrate the reliability and the efficiency of our data-driven technique to guide the Markov chain search and get quickly the required morphology characterisation.

5   Appendix: some mathematical details

We present here some technical results to prove the convergence of the proposed algorithm. We follow the lines of Stoica [20].

More generally, we show the convergence of the proposed algorithm when the following conditions are simultaneously fulfilled:

- the target distribution π is a locally stable process having density $h(.)$ with respect to an inhomogeneous Poisson process with intensity function $\lambda(.)$,

- the birth and death probabilities satisfy $0 < p_b, p_d < 1$ with $0 < p_b + p_d < 1$,

-

$$u_n := \sup_{y^{m_y} \in \mathcal{X} \times \mathcal{M}, \mathbf{x}^m \in \Xi_n} \frac{q_d(\mathbf{x}, y)\lambda(y)}{q_b(\mathbf{x} \setminus \{y\}, y)} \xrightarrow{n \to \infty} 0, \tag{12}$$

and

$$\frac{q_b(\mathbf{x} \setminus \{y\}, y)}{q_d(\mathbf{x}, y)} \frac{1}{\lambda(y)} > 0, \quad \forall \mathbf{x}^m \in \Xi, \forall y^{m_y} \in \mathcal{X} \times \mathcal{M}. \tag{13}$$

First, the algorithm should reproduce a transition kernel P such that the target distribution π is an invariant distribution of the chain. It turns out that the target distribution π is an invariant distribution of the chain if the proposal kernel P and the acceptance probabilities $\alpha(.,.)$ are chosen so that the Markov chain is reversible with respect to π. The reversibility property is fulfilled when for any $A, B \in \mathcal{F}_\Xi$

$$\pi P(A, B) = \int_A \int_B I_A(\mathbf{x}^m) I_B(\mathbf{x}'^{m'}) \pi(d\mathbf{x}^m) P(\mathbf{x}^m, d\mathbf{x}'^{m'})$$

is symmetric when exchanging the sets A and B. We require that the proposal kernel and the acceptance probabilities satisfy the stronger property of detailed balance condition which writes down here as

$$p_b h(\mathbf{x}^m) q_b(\mathbf{x}, y) \alpha(\mathbf{x}^m, \mathbf{x}^m \cup \{y^{m_y}\}) = p_d h(\mathbf{x}^m \cup \{y^{m_y}\}) q_d(\mathbf{x} \cup \{y\}, y) \alpha(\mathbf{x}^m \cup \{y^{m_y}\}, \mathbf{x}^m) \lambda(y) \tag{14}$$

where $\lambda(.)$ is the intensity function of the reference inhomogeneous Poisson process.

We first show that the detailed balance condition (14) implies the reversibility of the transition kernel.

Since the proposal kernel Q only enables moves from Ξ_n to Ξ_{n+1} and from Ξ_{n+1} to Ξ_n and since the acceptance step only allows in addition the possibility of no move, we only need to prove that $\pi P(A_n, B_{n+1}) = \pi P(B_{n+1}, A_n)$ for $A_n \subseteq \Xi_n$ and $B_{n+1} \subseteq \Xi_{n+1}$ for any n. On the one hand, the definition of πP gives

$$\pi P(A_n, B_{n+1}) = \int_{(\mathcal{X} \times \mathcal{M})^n} I(\mathbf{x}^m \in A_n) P(\mathbf{x}^m, B_{n+1}) \pi(d\mathbf{x}^m)$$

$$= \int_{\mathbf{x}^m \in A_n} \int_{\mathbf{x}'^{m'} \in B_{n+1}} \alpha(\mathbf{x}^m, \mathbf{x}'^{m'}) p_b Q_b(\mathbf{x}, d\mathbf{x}'^{m'}) \pi(d\mathbf{x}^m)$$

where $\pi(d\mathbf{x}^m) = h(\mathbf{x}^m)\pi_\Lambda(d\mathbf{x}^m)$ and where Q_b is given by (10). We get

$$\pi P(A_n, B_{n+1}) = \int_{(\mathcal{X} \times \mathcal{M})^{n+1}} I(\mathbf{x}^m \in A_n, \mathbf{x}^m \cup \{y^{m_y}\} \in B_{n+1})$$

$$\times \alpha(\mathbf{x}^m, \mathbf{x}^m \cup \{y^{m_y}\}) p_b q_b(\mathbf{x}, y) p_M(m_y) h(\mathbf{x}^m) \lambda_{\mathcal{X}}^{\text{Leb}}(dy) \lambda_{\mathcal{M}}^{\text{Leb}}(dm_y) \pi_\Lambda(d\mathbf{x}^m)$$

Now, we substitute π_Λ by its expression in (2). This gives

$$\pi P(A_n, B_{n+1}) = \frac{\exp(-\Lambda(\mathcal{X}))}{n!} \int_{(\mathcal{X} \times M)^{n+1}} I(\mathbf{x}^m \in A_n, \mathbf{x}^m \cup \{y^{m_y}\} \in B_{n+1}) \alpha(\mathbf{x}^m, \mathbf{x}^m \cup \{y^{m_y}\})$$

$$\times p_b q_b(\mathbf{x}, y) p_M(m_y) h(\mathbf{x}^m) \lambda_{\mathcal{X}}^{\mathrm{Leb}}(dy) \lambda_{\mathcal{M}}^{\mathrm{Leb}}(dm_y) \prod_{i=1}^{n} \lambda(x_i) p_M(m_i) \lambda_{\mathcal{X}}^{\mathrm{Leb}}(dx_i) \lambda_{\mathcal{M}}^{\mathrm{Leb}}(dm_i)$$

On the other hand, we have

$$\pi P(B_{n+1}, A_n) = \int_{(\mathcal{X} \times \mathcal{M})^{n+1}} I(\mathbf{z}^m \in B_{n+1}) P(\mathbf{z}^m \in A_n) \pi(d\mathbf{z}^m)$$

$$= \int_{\mathbf{z}^m \in B_{n+1}} \int_{\mathbf{z}'^{m'} \in A_n} \alpha(\mathbf{z}^m, \mathbf{z}'^{m'}) p_d Q_d(\mathbf{z}^m, d\mathbf{z}'^{m'}) \pi(d\mathbf{z}^m)$$

where $\pi(d\mathbf{z}^m) = h(\mathbf{z}^m) \pi_\Lambda(d\mathbf{z}^m)$ and where Q_d is given by (11). We get

$$\pi P(B_{n+1}, A_n) = \int_{(\mathcal{X} \times \mathcal{M})^{n+1}} \sum_{y^{m_y} \in \mathcal{X} \times \mathcal{M}} I(y^{m_y} \in \mathbf{z}^m) I(\mathbf{z}^m \in B_{n+1}, \mathbf{z}^m \setminus \{y^{m_y}\} \in A_n)$$

$$\times \alpha(\mathbf{z}^m, \mathbf{z}^m \setminus \{y^{m_y}\}) p_d q_d(\mathbf{z}, y) h(\mathbf{z}^m) \pi_\Lambda(d\mathbf{z}^m)$$

We reverse the order of the sum and the integral, all the terms being positive:

$$\pi P(B_{n+1}, A_n) = \sum_{y^{m_y} \in \mathcal{X} \times \mathcal{M}} \int_{(\mathcal{X} \times \mathcal{M})^{n+1}} I(y^{m_y} \in \mathbf{z}^m) I(\mathbf{z}^m \in B_{n+1}, \mathbf{z}^m \setminus \{y^{m_y}\} \in A_n)$$

$$\times \alpha(\mathbf{z}^m, \mathbf{z}^m \setminus \{y^{m_y}\}) p_d q_d(\mathbf{z}, y) h(\mathbf{z}^m) \pi_\Lambda(d\mathbf{z}^m)$$

Now, we substitute π_Λ by its expression on (2). This gives

$$\pi P(B_{n+1}, A_n) = \frac{\exp(-\Lambda(\mathcal{X}))}{(n+1)!} \sum_{y^{m_y} \in \mathcal{X} \times \mathcal{M}} \int_{(\mathcal{X} \times \mathcal{M})^{n+1}} I(y^{m_y} \in \mathbf{z}^m) I(\mathbf{z}^m \in B_{n+1}, \mathbf{z}^m \setminus \{y^{m_y}\} \in A_n)$$

$$\times \alpha(\mathbf{z}^m, \mathbf{z}^m \setminus \{y^{m_y}\}) p_d q_d(\mathbf{z}, y) h(\mathbf{z}^m) \prod_{i=1}^{n+1} \lambda(z_i) d\lambda_{\mathcal{X}}^{\mathrm{Leb}}(z_i) p_M(dm_i) d\lambda_{\mathcal{M}}^{\mathrm{Leb}}(m_i)$$

Letting $\mathbf{x}^m = \mathbf{z}^m \setminus \{y^{m_y}\}$ in the integrals, we get

$$\pi P(B_{n+1}, A_n) = \frac{\exp(-\Lambda(\mathcal{X}))}{(n+1)!} \sum_{y^{m_y} \in \mathcal{X} \times \mathcal{M}} \int_{(\mathcal{X} \times \mathcal{M})^{n+1}} I(\mathbf{x}^m \cup \{y^{m_y}\} \in B_{n+1}, \mathbf{x}^m \in A_n) \alpha(\mathbf{x}^m \cup \{y^{m_y}\}, \mathbf{x}^m)$$

$$\times p_d q_d(\mathbf{x} \cup \{y\}, y) h(\mathbf{x}^m \cup \{y^{m_y}\}) \prod_{i=1}^{n} \lambda(x_i) d\lambda_{\mathcal{X}}^{\mathrm{Leb}}(x_i) p_M(dm_i) d\lambda_{\mathcal{M}}^{\mathrm{Leb}}(m_i) \lambda(y) d\lambda_{\mathcal{X}}^{\mathrm{Leb}}(y) p_M(dm_y) d\lambda_{\mathcal{M}}^{\mathrm{Leb}}(m_y)$$

17

Only the terms for which the indicator $I(\mathbf{x}^m \cup \{y^{m_y}\} \in B_{n+1}, \mathbf{x}^m \in A_n)$ is non zero are non-zeros. There are $(n+1)$ such terms in which case they are all equal to the same integral (since it does not depend on the integration order). This entails

$$\pi P(B_{n+1}, A_n) = \frac{\exp(-\Lambda(\mathcal{X}))}{n!} \int_{(\mathcal{X} \times \mathcal{M})^{n+1}} I(\mathbf{x}^m \cup \{y^{m_y}\} \in B_{n+1}, \mathbf{x}^m \in A_n) \alpha(\mathbf{x}^m \cup \{y^{m_y}\}, \mathbf{x}^m)$$

$$\times p_d q_d(\mathbf{x} \cup \{y\}, y) h(\mathbf{x}^m \cup \{y^{m_y}\}) \prod_{i=1}^{n} \lambda(x_i) d\lambda_{\mathcal{X}}^{\text{Leb}}(x_i) p_M(dm_i) d\lambda_{\mathcal{M}}^{\text{Leb}}(m_i) \lambda(y) d\lambda_{\mathcal{X}}^{\text{Leb}}(y) p_M(dm_y) d\lambda_{\mathcal{M}}^{\text{Leb}}(m_y)$$

The detailed balance condition in (14) gives the reversility which concludes this proof.

We determine the Green ratios for both the birth and death proposals according to the method of Green [7]. To this aim, we define a symmetric measure $\psi : \mathcal{F}_\Xi \times \mathcal{F}_\Xi \to \mathbb{R}^+$ such that the Radon-Nikodym derivative of πQ with respect to ψ exists. This derivative is denoted by D. The proof runs along similar lines than the above one but the calculus are presented for the sake of clarity and completeness. Let us recall that the measure πQ is defined for (A, B) in $\mathcal{F}_\Xi \times \mathcal{F}_\Xi$ by

$$\pi Q(A, B) = \int_{\mathbf{x}^m \in A} Q(\mathbf{x}^m, B) \pi(d\mathbf{x}^m).$$

Since the proposal kernel Q only enables moves from Ξ_n to Ξ_{n+1} and from Ξ_{n+1} to Ξ_n, we only have to study the derivability of πQ on sets (A_n, B_{n+1}) or (A_{n+1}, B_n) with $A_n, B_n \subseteq \Xi_n$ and $A_{n+1}, B_{n+1} \subseteq \Xi_{n+1}$ for any arbitrary n. On the first hand,

$$\pi Q(A_n, B_{n+1}) = \int_{(\mathcal{X} \times \mathcal{M})^n} I(\mathbf{x}^m \in A_n) Q(\mathbf{x}^m, B_{n+1}) \pi(d\mathbf{x}^m)$$

$$= \int_{(\mathcal{X} \times \mathcal{M})^n} I(\mathbf{x}^m \in A_n) p_b Q_b(\mathbf{x}^m, B_{n+1}) \pi(d\mathbf{x}^m)$$

with $\pi(d\mathbf{x}^m) = h(\mathbf{x}^m)\pi_\Lambda(d\mathbf{x}^m)$ and where Q_b is given by (10). We get

$$\pi Q(A_n, B_{n+1}) = \int_{(\mathcal{X} \times \mathcal{M})^{n+1}} I(\mathbf{x}^m \in A_n, \mathbf{x}^m \cup \{y^{m_y}\} \in B_{n+1})$$

$$p_b q_b(\mathbf{x}, y) h(\mathbf{x}^m) \lambda_{\mathcal{X}}^{\text{Leb}}(dy) p_M(m_y) \lambda_{\mathcal{M}}^{\text{Leb}}(dm_y) \pi_\Lambda(d\mathbf{x}^m)$$

We define a measure ψ_n^+ on $\Xi_n \times \Xi_{n+1}$ by

$$\psi_n^+(A_n, B_{n+1}) = \int_{(\mathcal{X} \times \mathcal{M})^{n+1}} I(\mathbf{x}^m \in A_n, \mathbf{x}^m \cup \{y^{m_y}\} \in B_{n+1})$$

$$\lambda(y) \lambda_{\mathcal{X}}^{\text{Leb}}(dy) p_M(m_y) \lambda_{\mathcal{M}}^{\text{Leb}}(dm_y) \pi_\Lambda(d\mathbf{x}^m)$$

18

With such a choice, the Radon-Nikodym derivative of πQ with respect to ψ_n^+ on $\Xi_n \times \Xi_{n+1}$ is

$$D_b(\mathbf{x}^m, \mathbf{x}^m \cup \{y^{m_y}\}) = \frac{p_b q_b(\mathbf{x}, y) h(\mathbf{x}^m)}{\lambda(y)}$$

On the other hand,

$$\pi Q(A_{n+1}, B_n) = \int_{(\mathcal{X} \times \mathcal{M})^{n+1}} I(\mathbf{z}^m \in A_{n+1}) Q(\mathbf{z}^m, B_n) \pi(d\mathbf{z}^m)$$

$$= \int_{(\mathcal{X} \times \mathcal{M})^{n+1}} I(\mathbf{z}^m \in A_{n+1}) p_d Q_d(\mathbf{z}^m, B_n) \pi(d\mathbf{z}^m)$$

with $\pi(d\mathbf{z}^m) = h(\mathbf{z}^m) \pi_\Lambda(d\mathbf{z}^m)$ and where Q_d is given by (11). We get

$$\pi Q(A_{n+1}, B_n) = \int_{(\mathcal{X} \times \mathcal{M})^{n+1}} \sum_{y^{m_y} \in \mathbf{z}^m} I(\mathbf{z}^m \in A_{n+1}, \mathbf{z}^m \setminus \{y^{m_y}\} \in B_n) p_d q_d(\mathbf{z}, y) h(\mathbf{z}^m) \pi_\Lambda(d\mathbf{z}^m)$$

We reverse the order of the sum and the integral, all the terms being positive:

$$\pi Q(A_{n+1}, B_n) = \sum_{y^{m_y} \in \mathcal{X} \times \mathcal{M}} \int_{(\mathcal{X} \times \mathcal{M})^{n+1}} I(\mathbf{z}^m \in A_{n+1},\, y^{m_y} \in \mathbf{z}^m,\, \mathbf{z}^m \setminus \{y^{m_y}\} \in B_n) p_d q_d(\mathbf{z}, y) h(\mathbf{z}^m) \pi_\Lambda(d\mathbf{z}^m)$$

Letting $\mathbf{x}^m = \mathbf{z}^m \setminus \{y^{m_y}\}$ in the integrals, we get

$$\pi Q(A_{n+1}, B_n) = \sum_{y^{m_y} \in \mathcal{X} \times \mathcal{M}} \int_{(\mathcal{X} \times \mathcal{M})^{n+1}} I(\mathbf{x}^m \cup \{y^{m_y}\} \in A_{n+1},\, \mathbf{x}^m \in B_n)$$

$$p_d q_d(\mathbf{x} \cup \{y\}, y) h(\mathbf{x}^m \cup \{y^{m_y}\}) \lambda(y) \lambda_\mathcal{X}^{\mathrm{Leb}}(dy) p_M(m_y) \lambda_\mathcal{M}^{\mathrm{Leb}}(dm_y) \pi_\Lambda(d\mathbf{x}^m)$$

since in view of (2), we have that $\pi_\Lambda(d\mathbf{z}^m) = \pi_\Lambda(d\mathbf{x}^m) \lambda(y) \lambda_\mathcal{X}^{\mathrm{Leb}}(dy) p_M(m_y) \lambda_\mathcal{M}^{\mathrm{Leb}}(dm_y)$. Only the terms for which the indicator $I(\mathbf{x}^m \cup \{y^{m_y}\} \in A_{n+1},\, \mathbf{x}^m \in B_n)$ is non zero are non-zeros. There are $(n+1)$ such terms in which case they are all equal to the same integral (since it does not depend on the integration order). This entails

$$\pi Q(A_{n+1}, B_n) = (n+1) \int_{(\mathcal{X} \times \mathcal{M})^{n+1}} I(\mathbf{x}^m \cup \{y^{m_y}\} \in A_{n+1},\, \mathbf{x}^m \in B_n)$$

$$p_d q_d(\mathbf{x} \cup \{y\}, y) h(\mathbf{x}^m \cup \{y^{m_y}\}) \lambda(y) \lambda_\mathcal{X}^{\mathrm{Leb}}(dy) p_M(m_y) \lambda_\mathcal{M}^{\mathrm{Leb}}(dm_y) \pi_\Lambda(d\mathbf{x}^m)$$

We define a measure ψ_n^- on $\Xi_{n+1} \times \Xi_n$ by

$$\psi_n^-(A_{n+1}, B_n) = (n+1) \int_{(\mathcal{X} \times \mathcal{M})^{n+1}} I(\mathbf{x}^m \cup \{y^{m_y}\} \in A_{n+1},\, \mathbf{x}^m \in B_n)$$

$$\lambda(y) \lambda_\mathcal{X}^{\mathrm{Leb}}(dy) p_M(m_y) \lambda_\mathcal{M}^{\mathrm{Leb}}(dm_y) \pi_\Lambda(d\mathbf{x}^m)$$

The Radon-Nikodym derivative of πQ with respect to ψ_n^- on $\Xi_{n+1} \times \Xi_n$ is then defined by

$$D_d(\mathbf{x}^m \cup \{y^{m_y}\}, \mathbf{x}^m) = p_d q_d(\mathbf{z}, y) h(\mathbf{x}^m \cup \{y^{m_y}\})$$

We then define a measure ψ on $\Xi \times \Xi$ as the measure concentrated on the set

$$\bigcup_{\substack{n=0 \\ disjoint}}^{\infty} \left((\Xi_n \times \Xi_{n+1}) \cup_{disjoint} (\Xi_{n+1} \times \Xi_n) \right)$$

which coincides with ψ_n^+ on $\Xi_n \times \Xi_{n+1}$ and with ψ_n^- on $\Xi_{n+1} \times \Xi_n$. It is straightforward to see that this actually defines a measure on $\Xi \times \Xi$. To prove the symmetry of ψ, we only have to prove that $\psi(A_{n+1}, B_n) = \psi(B_n, A_{n+1})$ for $A_n \in \Xi_n$ and $B_{n+1} \in \Xi_{n+1}$. This is immediate since

$$\psi(B_n, A_{n+1}) = \psi_n^+(B_n, A_{n+1})$$
$$= \frac{\exp(-\Lambda(\mathcal{X}))}{n!} \int_{(\mathcal{X} \times \mathcal{M})^{n+1}} I(\mathbf{x}^m \in B_n,, \mathbf{x}^m \cup \{y^{m_y}\} \in A_{n+1}) \lambda(y) \lambda_{\mathcal{X}}^{\text{Leb}}(dy)$$
$$\times p_M(m_y) \lambda_{\mathcal{M}}^{\text{Leb}}(dm_y) \prod_{i=1}^{n} \lambda(x_i) \lambda_{\mathcal{X}}^{\text{Leb}}(dx_i) p_M(m_i) \lambda_{\mathcal{M}}^{\text{Leb}}(dm_i)$$

and

$$\psi(A_{n+1}, B_n) = \psi_n^-(A_{n+1}, B_n)$$
$$= \frac{(n+1)\exp(-\Lambda(\mathcal{X}))}{(n+1)!} \int_{(\mathcal{X} \times \mathcal{M})^{n+1}} I(\mathbf{x}^m \cup \{y^{m_y}\} \in A_{n+1}, \mathbf{x}^m \in B_n)$$
$$\times \lambda(y) \lambda_{\mathcal{X}}^{\text{Leb}}(dy) p_M(m_y) \lambda_{\mathcal{M}}^{\text{Leb}}(dm_y) \prod_{i=1}^{n} \lambda(x_i) \lambda_{\mathcal{X}}^{\text{Leb}}(dx_i) p_M(m_i) \lambda_{\mathcal{M}}^{\text{Leb}}(dm_i)$$

The derivative D exists and coincides with D_b on $\Xi_n \times \Xi_{n+1}$ and with D_d on $\Xi_{n+1} \times \Xi_n$. We then deduce the following Green ratios:

$$r_b(\mathbf{x}^m, y) = \frac{D_d(\mathbf{x}^m \cup \{y^{m_y}\}, \mathbf{x}^m)}{D_b(\mathbf{x}^m, \mathbf{x}^m \cup \{y^{m_y}\})} = \frac{h(\mathbf{x}^m \cup \{y^{m_y}\})}{h(\mathbf{x}^m)} \frac{p_d}{p_b} \frac{q_d(\mathbf{x} \cup \{y\}, y)}{q_b(\mathbf{x}, y)} \lambda(y)$$

and

$$r_d(\mathbf{z}^m, y^{m_y}) = \frac{D_d(\mathbf{z}^m \setminus \{y^{m_y}\}, \mathbf{z}^m)}{D_b(\mathbf{z}^m, \mathbf{z}^m \setminus \{y^{m_y}\})} = \frac{h(\mathbf{z}^m \setminus \{y^{m_y}\})}{h(\mathbf{z}^m)} \frac{p_b}{p_d} \frac{q_b(\mathbf{z} \setminus \{y\}, y)}{q_d(\mathbf{z}, y)} \frac{1}{\lambda(y)}$$

The acceptance probabilities are then defined in the birth case as

$$\alpha(\mathbf{x}^m, \mathbf{x}^m \cup \{y^{m_y}\}) = \min(1, r_b(\mathbf{x}^m, y^{m_y}))$$

and in the death case as

$$\alpha(\mathbf{z}^m, \mathbf{z}^m \setminus \{y^{m_y}\}) = \min(1, r_d(\mathbf{z}^m, y^{m_y})). \tag{15}$$

It is easy to check that

$$r_b(\mathbf{x}^m, \mathbf{x}^m \cup \{y^{m_y}\}) = \frac{1}{r_d(\mathbf{x}^m \cup \{y^{m_y}\}, \mathbf{x}^m)}$$

so that the detailed balance condition in (14) is satisfied.

A potential problem is that the condition of reversibility alone implies the initial distribution needs to be chosen according to the distribution of interest. This problem is solved if the chain is irreducible, aperiodic and Harris recurrent.

Irreducibility is a crucial property in the set-up of MCMC algorithms. Roughly speaking, a ϕ-irreducible Markov chain is able to reach any set A which is large enough in the sense that $\phi(A) > 0$. For any positive integer n, let P^n be the n-step transition probability kernel which is defined inductively for $\mathbf{x}^m \in \Xi$ and $B \in \mathcal{F}_\Xi$ by

$$P^n(\mathbf{x}^m, B) = \int_\Xi P(\mathbf{x}^m, d\mathbf{z}^m) P^{n-1}(\mathbf{z}^m, B)$$

where $P^0(\mathbf{x}^m, B) = \delta_B(\mathbf{x}^m)$ with δ_B being the Dirac measure concentrated at B.

We show that our chain satisfies the irreducibility property provided the conditions (13) and (12) are fulfilled. To this aim, we exhibit a positive measure ϕ such that, for any set A in $\mathcal{F}_\Xi$ such that $\phi(A) > 0$, there exists a positive integer n_0 such that $P^{n_0}(\mathbf{x}^m, A) > 0$ for any $\mathbf{x}^m$ in Ξ. We choose as positive measure the Dirac measure concentrated at $\emptyset$, that is $\phi = \delta_{\{\emptyset\}}$, so that the only set A in $\mathcal{F}_\Xi$ having positive measure is $\{\emptyset\}$. The Markov property of the chain entails that, for any $\mathbf{x}^m \in \Xi$, for any positive integer n_0 such that $n_0 \geq n(\mathbf{x})$, we have

$$P^{n_0}(\mathbf{x}^m, \{\emptyset\}) \geq P^{n(\mathbf{x})}(\mathbf{x}^m, \{\emptyset\}) P^{n_0 - n(\mathbf{x})}(\emptyset, \{\emptyset\})$$

where

$$P(\emptyset, \{\emptyset\}) = 1 - p_b \int_{\mathcal{X} \times \mathcal{M}} q_p(\emptyset, y) p_M(m_y) \alpha(\emptyset, \{y^{m_y}\}) d\lambda_{\mathcal{X}}^{\text{Leb}}(y) d\lambda_{\mathcal{M}}^{\text{Leb}}(m_y) \geq 1 - p_b = p_d$$

and where

$$P^{n(\mathbf{x})}(\mathbf{x}, \{\emptyset\}) \geq p_d^{n(\mathbf{x})} \left(\min_{\mathbf{z}^m \in \Xi, y^{m_y} \in \mathbf{z}^m : n(\mathbf{z}) \leq n_0} \{\alpha(\mathbf{z}^m, \mathbf{z}^m \setminus \{y^{m_y}\})\} \right)^{n(\mathbf{x})}.$$

For $\mathbf{z}^m$ in Ξ with $n(\mathbf{z}) \leq n_0$ and $y^{m_y} \in \mathbf{z}^m$, the death acceptance probability is defined by (15). Since the target distribution is a locally stable process, there exists a $K > 0$ such that for $\mathbf{z}^m$ in Ξ and $y^{m_y} \in \mathbf{z}^m$,

$$\frac{h(\mathbf{z}^m \setminus \{y^{m_y}\})}{h(\mathbf{z}^m)} \geq \frac{1}{K}.$$

On the other hand, letting

$$\Delta(n_0) = \min_{\mathbf{z}^m \in \Xi, y^{m_y} \in \mathbf{z}^m : n(\mathbf{z}) \leq n_0} \left\{ \frac{q_b(\mathbf{z} \setminus \{y\}, y)}{q_d(\mathbf{z}, y)} \frac{1}{\lambda(y)} \right\},$$

21

we deduce that for any $\mathbf{x}^m$ in Ξ such that $n(\mathbf{x}) \leq n_0$, for any y^{m_y} in $\mathbf{x}^m$

$$\alpha(\mathbf{x}^m, \mathbf{x}^m \setminus \{y^{m_y}\}) \geq \min\left(1, \frac{\Delta(n_0)}{K}\frac{p_b}{p_d}\right)$$

In view of (12), one may find a n_0 large enough so that

$$\Delta(n_0) < \frac{1}{K}\frac{p_b}{p_d}.$$

in which case for any $\mathbf{x}^m$ in Ξ such that $n(\mathbf{x}) \leq n_0$, for any y^{m_y} in $\mathbf{x}^m$

$$\alpha(\mathbf{x}^m, \mathbf{x}^m \setminus \{y^{m_y}\}) \geq \frac{\Delta(n_0)}{K}\frac{p_b}{p_d}.$$

We get that for any $\mathbf{x}^m \in \Xi$, one may find a n_0 large enough so that

$$P^{n_0}(\mathbf{x}^m, \{\emptyset\}) \geq p_d^{n(\mathbf{x})}\left(\frac{\Delta(n_0)}{K}\frac{p_b}{p_d}\right)^{n_0} p_d^{n_0 - n(\mathbf{x})} = \left(\frac{p_b \Delta(n_0)}{K}\right)^{n_0}. \tag{16}$$

In view of (13), this gives the ϕ-irreducibility of P.

The Harris recurrence is the property that guarantees the non-existence of sets of configurations $\Xi' \subset \Xi$ such that $\pi(\Xi') = 1$ and $\lim_{n \to \infty} \sup_{A \in \mathcal{F}_\Xi} |P^n(\mathbf{x}^m, A) - \pi(A)| \neq 0$ for each $\mathbf{x}^m$ in Ξ'. Hence the recurrence property guarantees the asymptotic convergence of a MCMC sampling algorithm regardless of the initial conditions. According to the paper of Meyn and Tweedie [15], a sufficient condition for Harris recurrence is the drift condition. The drift condition says that, if the transition kernel P is ϕ-irreducible, if there is a small set C in $\mathcal{F}_\Xi$ and a function unbounded off small sets $V : \mathcal{X}_i \to \mathbb{R}^+$ such that

$$\int_\Xi P(\mathbf{x}^m, d\mathbf{z}^m)V(\mathbf{z}^m) \leq V(\mathbf{x}^m), \quad \forall \mathbf{x}^m \notin C$$

then P is Harris recurrent. We recall that a set C in $\mathcal{F}_\Xi$ is small if there are a positive integer n_0 and a non-zero measure ν such that

$$P^{n_0}(\mathbf{x}^m, B) \geq \nu(B), \quad \forall \mathbf{x}^m \in C, \forall B \in \mathcal{F}_\Xi$$

First, set a number $A > 1$ and then set a number $0 < \varepsilon < \min\left(1, \frac{p_d}{p_b A}\right)$. In view of (12), one may find a N_ε large enough so that for ny $n \geq N_\varepsilon$

$$\frac{p_d}{p_b}K u_n \leq \varepsilon \tag{17}$$

Equation (16) shows that the set

$$C(\varepsilon) = \{\mathbf{x}^m \in \Xi : n(\mathbf{x}) \leq N_\varepsilon\}$$

is small with the choice $\nu(.) = (p_d \Delta(N_\varepsilon)/K)^{N_\varepsilon} \delta_{\{\emptyset\}}(.)$. We also recall that a function $V : \Xi \to \mathbb{R}^+$ is unbounded off small sets if the level sets

$$C_V(\alpha) = \{\mathbf{x}^m \in \Xi : V(\mathbf{x}^m) \leq \alpha\}$$

are small for any $\alpha > 0$. The choice $V(\mathbf{x}^m) = A^{n(\mathbf{x})}$ entails as previously that the sets

$$C_V(\alpha) = \{\mathbf{x}^m \in \Xi : n(\mathbf{x}) \leq \log\left(\frac{\alpha}{A}\right)\}$$

are small for any $\alpha > 0$. It remains to get the negativity of the drift. Applying the transition kernel P to the function V leads to

$$\int_\Xi V(\mathbf{z}^m) P(\mathbf{x}^m, d\mathbf{z}^m) = p_b A^{n(\mathbf{x})} (A-1) \int_{\mathcal{X} \times \mathcal{M}} q_b(\mathbf{x}, y) \alpha(\mathbf{x}^m, \mathbf{x}^m \cup \{y^{m_y}\}) p_M(m_y) d\lambda_{\mathcal{X}}^{\mathrm{Leb}}(y) d\lambda_{\mathcal{M}}^{\mathrm{Leb}}(m_y)$$

$$+ p_d A^{n(\mathbf{x})} \left(\frac{1}{A} - 1\right) \sum_{y^{m_y} \in \mathbf{x}^m} q_d(\mathbf{x}, y) \alpha(\mathbf{x}^m, \mathbf{x}^m \setminus \{y^{m_y}\}) + A^{n(\mathbf{x})}$$

In view of (17), we deduce for $\mathbf{x}^m \notin C(\varepsilon)$ that

$$\alpha(\mathbf{x}^m, \mathbf{x}^m \cup \{y^{m_y}\}) \leq \varepsilon$$

and that

$$\alpha(\mathbf{x}^m, \mathbf{x}^m \setminus \{y^{m_y}\}) = 1.$$

This gives for $\mathbf{x}^m \notin C(\varepsilon)$

$$\int_\Xi V(\mathbf{z}^m) P(\mathbf{x}^m, d\mathbf{z}^m) \leq \left(p_b(A-1)\varepsilon + p_d\left(\frac{1}{A} - 1\right) + 1\right) V(\mathbf{x}^m) \tag{18}$$

$$\leq V(\mathbf{x}^m).$$

So far, we know that the Harris recurrent chain should reach the equilibrium asymptotically but we do not know when this will happen. One way of solving this problem is to build an ergodic chain (that is Harris recurrent and aperiodic). In this case, there is no need to wait infinitely until the equilibrium. The property of aperiodicity is immediate since the self-transition probabilities

$$P(\mathbf{x}^m, \{\mathbf{x}^m\}) = 1 - p_b \int_{\mathcal{X} \times \mathcal{M}} q_p(\mathbf{x}, y) p_M(m_y) \alpha(\mathbf{x}^m, \mathbf{x}^m \cup \{y^{m_y}\}) d\lambda_{\mathcal{X}}^{\mathrm{Leb}}(y) d\lambda_{\mathcal{M}}^{\mathrm{Leb}}(m_y)$$

$$- p_d \sum_{y^{m_y} \in \mathbf{x}^m} q_d(\mathbf{x}, y) \alpha(\mathbf{x}^m, \mathbf{x}^m \setminus \{y^{m_y}\})$$

are positive for $\mathbf{x}^m$ in A such that $\pi(A) > 0$.

The last refinement consists of proving that the geometric drift property is fulfilled so that the convergence of the chain to its equilibrium distribution happens at a geometric speed. The geometric drift condition says that a ϕ-irreducible and aperiodic Markov chain with transition kernel P is geometrically ergodic if there exists a function $V : \Xi \to (1, \infty)$, two numbers b in $\mathbb{R}$ and $a < 1$ and a small set C in $\mathcal{F}_\Xi$ such that

$$\int_\Xi V(\mathbf{z}^m) P(\mathbf{x}^m, d\mathbf{z}^m) \leq aV(\mathbf{x}^m) + bI_C(\mathbf{x}^m), \quad \forall \mathbf{x}^m \in \Xi.$$

Taking the set $C(\varepsilon)$ and the function V as previously defined for some $1 < A < 1 + 1/p_b$ and taking $b = A^{N_\varepsilon + 1}$ gives the desired property in view of (18).

Indeed, if $\mathbf{x}^m \notin C(\varepsilon)$, then (18) reads as

$$\int_\Xi V(\mathbf{z}^m) P(\mathbf{x}^m, d\mathbf{z}^m) \leq \underbrace{\left(p_b(A-1)\varepsilon + p_d \left(\frac{1}{A} - 1 \right) + 1 \right)}_{:=a' < 1} V(\mathbf{x}^m) + b \times 0.$$

If $\mathbf{x}^m \in C(\varepsilon)$, then

$$\int_\Xi V(\mathbf{z}^m) P(\mathbf{x}^m, d\mathbf{z}^m) = p_b A^{n(\mathbf{x})}(A-1) \underbrace{\int_{\mathcal{X} \times \mathcal{M}} q_b(\mathbf{x}, y) \alpha(\mathbf{x}^m, \mathbf{x}^m \cup \{y^{m_y}\}) p_M(m_y) d\lambda_{\mathcal{X}}^{\mathrm{Leb}}(y) d\lambda_{\mathcal{M}}^{\mathrm{Leb}}(m_y)}_{\leq 1}$$

$$+ p_d A^{n(\mathbf{x})} \left(\frac{1}{A} - 1 \right) \underbrace{\sum_{y^{m_y} \in \mathbf{x}^m} q_d(\mathbf{x}, y) \alpha(\mathbf{x}^m, \mathbf{x}^m \setminus \{y^{m_y}\})}_{<0} + A^{n(\mathbf{x})}$$

$$\leq \underbrace{p_b(A-1)}_{:=a'' < 1} A^{n(\mathbf{x})} + A^{N_\varepsilon + 1}$$

Taking $a = \max(a', a'')$ gives the result.

References

[1] Ahirwal D., Hébraud A., Kádár R., Wilhelm M. and Schlatter G. (2013) From Self-assembly of Electrospun NanoFibers to 3D cm-thick Hierarchical Foams, Soft Matter, 9: 3164-3172.

[2] Al-Awadhi F., Jennison C. and Hurn M. (2004) Statistical image analysis for a confocal microscopy two-dimensional section of cartilage growth, Applied Statistics, 53: 31-49.

[3] Baddeley A. and Van Lieshout M.N.M. (1993) Stochastic geometry model in high-level vision. In *Statistics and Images: 1* (Eds K.V. Mardia and G.K. Kanji), pp. 2335-256.

[4] Brooks S.P., Giudici P. and Roberts G.O. (2003) Efficient construction of reversible jump MCMC proposal distributions (with discussion), Journal of the Royal Statistical Society, series B, 65:3-55.

[5] Descombes X. (Editor) Stochastic geometry for image analysis (2011), Wiley.

[6] Eap S., Ferrand A., Palomares C.M., Hébraud A., Stoltz J.F., Mainard D., Schlatter G. and Benkirane-Jessel N. (2012) Electrospun nanofibrous 3D scaffold for bone tissue engineering, Bio-Medical Materials and Engineering, 22: 137-141.

[7] Green P.J. (1995) Reversible jump Markov Chain Monte Carlo computation and Bayesian model determination, Biometrika, 82: 711-732.

[8] Imberty M. and Descombes X. (2000) Simulation de processus objets : Etude de faisabilité pour une application à la segmentation d'image, Research Report 3881-INRIA.

[9] Kirkpatrick S., Gelatt C.D. and Vecchi M.P. (1983) Optimization by Simulated Annealing, Science, 220: 671-680.

[10] Konomi B., Dahval S., Huang J., Kundu S., Huitink D., Liang H., Ding Y. and Mallick B. (2013) Bayesian object classification of gold nanoparticles, Annals of Applied Statistics, 7:640-668.

[11] Lacoste C., Descombes X. and Zerubia J. (2002) A Comparative Study of Point Processes for Line Network Extraction in Remote Sensing, Research Report 4516-INRIA.

[12] Lacoste C., Descombes X. and Zerubia J. (2003) Road network extraction in remote sensing by a Markov Object process, Proceedings of the ICIP'03, Barcelona, Spain.

[13] Lavielle N., Hébraud A., Mendoza C., Ferrand A., Benkirane-Jessel N. and Schlatter G. (2012) Structuring and molding of electrospun nanofibers: Effect of electrical and topographical local properties of micro-patterned collectors, Macromolecular Materials and Engineering, 297: 958-968.

[14] Lavielle N., de Geus M., Hébraud A., Schlatter G., Rossi R., Thöny-Meyer L. and Popa A.-M. (2013) Controlled formation of polycaprolactone ultrathin electrospun nanofibers by hydrolytic degradation approach, European Polymer Journal, 49: 1331-1336.

[15] Meyn S.P. and Tweedie R.L. (1993) Markov chains and stochastic stability, Springer.

[16] Møller J. and Waagepetersen R.P. (2004) Statistical Inference and Simulation for Spatial Point Processes. Chapman and Hall/CRC Press.

[17] Perrin G., Descombes X. and Zerubia J. (2006) A Non-Bayesian Model for Tree Crown Extraction using Marked Point Processes, Research report 5846-INRIA.

[18] Robert C. and Casella G. (2004) Monte Carlo Statistical Methods, Springer.

[19] Shin E.H., Cho K.S., Seo M.H. and Kim H. (2008) Determination of electrospun fiber diameter distributions using image analysis processing, Macromolecular Research, 314-319

[20] Stoica R. (2010) Marked point processes for statistical and morphological analysis of astronomical data, The European Physical Journal Special Topics, 186: 123-165.

[21] Tu Z. and Zhu S.-C. (2002) Image segmentation by data-driven Markov Chain Monte Carlo, IEEE Transactions on Pattern Analysis and Machine Intelligence, 24: 657-673.

[22] Van Lieshout M.N.M. (1995) Markov point processes and their applications in high-level imaging, Bulletin of the International Statistical Institute LVI, book 2: 559-576.

[23] Van Lieshout M.N.M. (2000) Markov point processes and their applications, Imperial College Press, London.

[24] Vigneron V., Syed T.Q., Barlovatz-Meimon G., Malo M., Montagne C. and Lelandais S. (2010) Adaptive filtering and hypothesis testing: Application to cancerous cells detection, Pattern Recognition Letters, 31: 2214-2224.
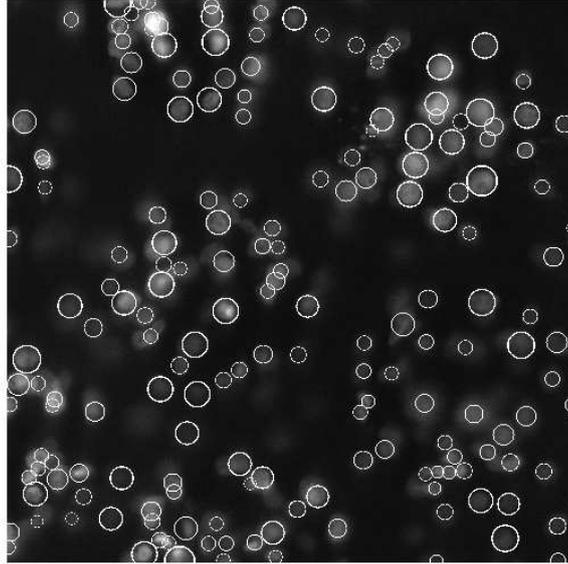
(a) Detected cells using our method
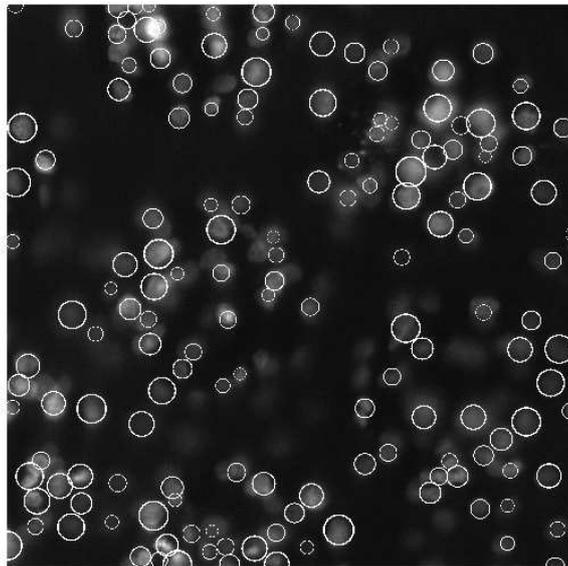
(b) Detected cells using the reference algorithm

Figure 5: Detected cells in the same SEM image
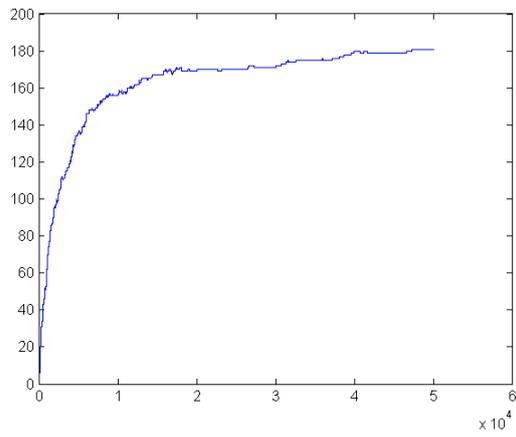
(a) Number of detected cells using our method
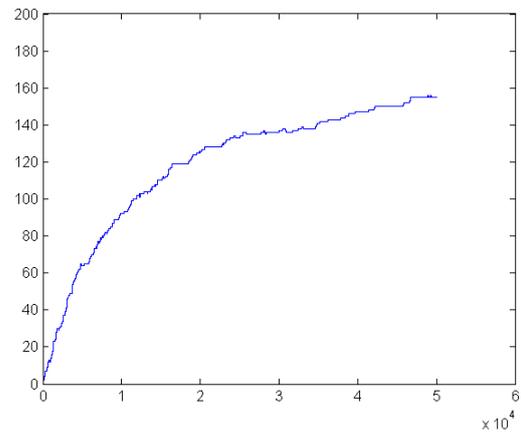
(b) Number of detected cells using the reference algorithm

Figure 6: Number of detected cells as a function of iteration number

28
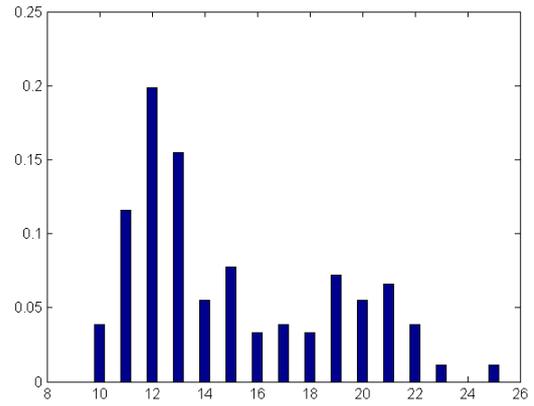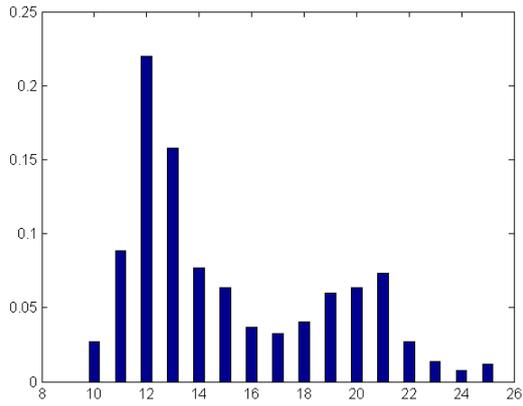
(a) Distribution used to simulate cell radius using our method

(b) Estimated distribution of cell radius using our method
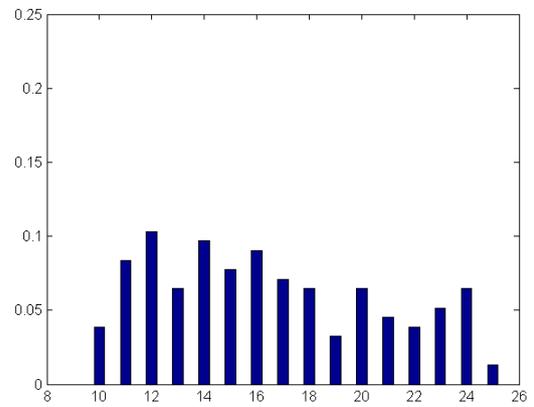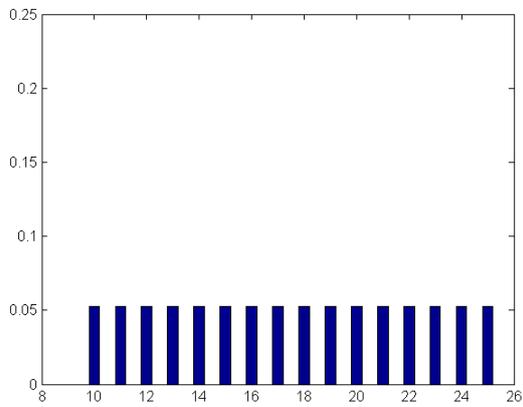
(c) Uniform distribution used to simulate cell radius using the reference algorithm

(d) Estimated distribution of cell radius using the reference algorithm

Figure 7: Histograms of cell radius distribution
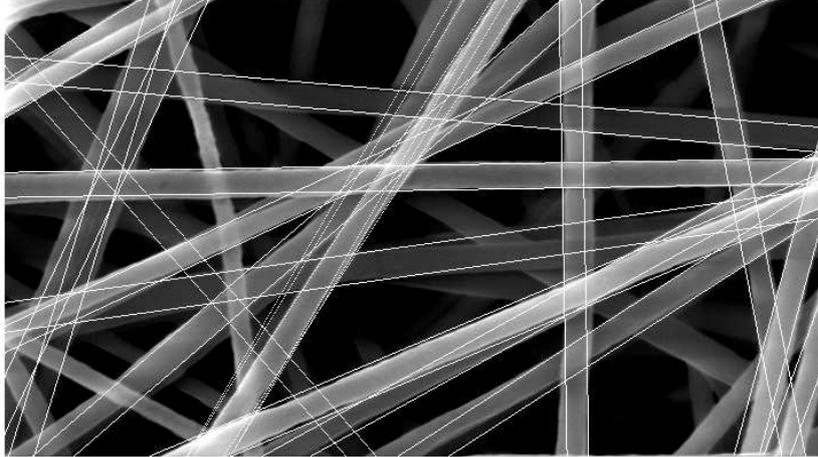
Figure 8: Detected fibers using our method
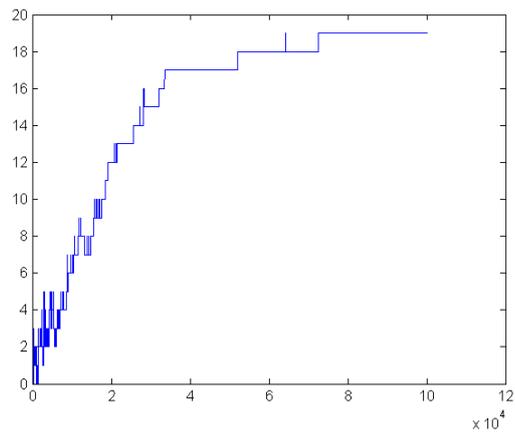
Figure 9: Number of detected fibers as a function of iteration number using our method
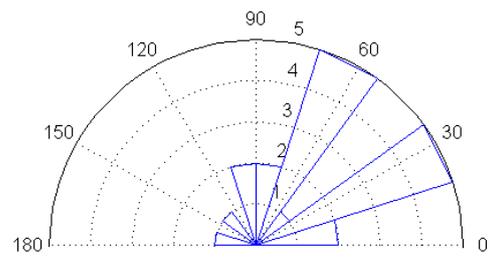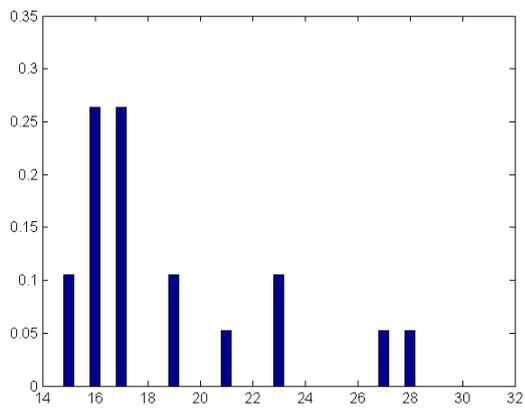
(a) Estimated distribution of fiber width using our method

(b) Estimated distribution of fiber orientation using our method

Figure 10: Histogram of fiber width and polar histogram of fiber orientation using our method

31