MAXIMUM LIKELIHOOD ESTIMATION FOR RANDOM SEQUENTIAL ADSORPTION

M. N. M. VAN LIESHOUT,* *CWI and EURANDOM*

Abstract

The existence and uniqueness of maximum likelihood estimators for the time and range parameters in random sequential adsorption models are investigated.

Keywords: Maximum likelihood estimation; random sequential adsorption; sequential spatial process; simple sequential inhibition

2000 Mathematics Subject Classification: Primary 60G55; 60D05; 62M30

1. Introduction

Packing problems occur in many physical and biological processes. Hard 'particles', such as cars, monomers, or proteins, arrive in some bounded Euclidean window according to a constant-rate Poisson process and select a position for themselves according to some fixed distribution. If it will result in no overlap with an established particle, the new one attaches itself to the selected location; otherwise, it leaves the system. The process continues either for a given time, until a given number of particles have been adsorbed, or until there is no room for another particle to establish itself without overlapping existing ones.

Complete random packings are those in which the birth process continues until saturation, that is, until no room is left in the window to accommodate another particle. On the line, a rigorous analysis dates back to [15]. Typical results concern the asymptotic coverage fraction of unit-length segments in an interval of increasing length (or, equivalently, the relative vacancy, that is, the length fraction not covered by particles) and central limit theorems for the number of adsorbed particles [6]. The problem is sometimes known as the car parking problem, with the segments representing cars and the interval representing the kerb. In higher dimensions, the analysis is rather harder. Simulation results on the coverage fraction are plentiful. For example, simulated complete random packings of discs in two dimensions can be found in [11] and [19]. Efficient simulation in $d \geq 2$ dimensions was studied in [5]. A law of large numbers for the coverage fraction was proved, from a theoretical point of view, in [13].

Other authors have considered runs of the process over a finite time horizon, or those that terminate when a certain packing density (i.e. number of adsorbed particles) has been achieved. Alternatively, the number of adsorbed particles could be stochastic, and follow some probability mass function. Such models were dubbed simple sequential inhibition models in the spatial statistics literature [4], and are known as random sequential adsorption models [7], used to describe the deposition of colloidal particles onto a substrate, amongst physicists. For finite time horizons, interest is focused on the evolution of summary statistics, for example the mean and variance of the number of particles, the pair correlation function, or the empty space distribution, and especially on the behaviour of the coverage fraction as a function of time, the

so-called kinetic law (cf. the Feder conjecture [8]), mostly using series expansions and Monte Carlo techniques.

The reader is referred to [10, Chapter 1.10], to the review papers [7], [17], and [18], or the textbooks [1], [3], and [20] for overviews and other pointers to the literature. For a flavour of current work in statistical physics, we refer the reader to three special issues of *Colloids and Surfaces A: Physicochemical and Engineering Aspects* (volume 165 (2000)).

To date, most attention has been focused on probabilistic work rather than inference. In this paper we take the statistical viewpoint. We place the random sequential adsorption model in the context of Markov sequential spatial processes [21], [22], derive its Radon–Nikodým derivative, and consider maximum likelihood estimation of its parameters. Note that the maximum likelihood approach is especially convenient in exponential family models, but that the random sequential adsorption model exhibits a much more complicated interaction structure.

The plan of the paper is as follows. In Section 2 we fix notation. In Section 3 some facts from the theory of convolution of exponential distributions are recalled. In Section 4 we present our main results, and in Section 5 we conclude the paper with a brief summary.

2. Definitions and notation

The *random sequential adsorption model* [7], also known as the *simple sequential inhibition model* [3], [4], in some bounded domain of $\mathbb{R}^d$ may be defined as follows. A unit-rate stream of particles arrives in the domain. If the arrival location is within a distance $r > 0$ of some existing particle, the particle leaves; otherwise, it is adsorbed. The process continues for some fixed time $\theta \geq 0$. Note that, in the limit as $\theta \to \infty$, the adsorption process becomes saturated (jammed), that is, particles are adsorbed until there are no open spaces left with radius at least r.

In a recent paper, [22], the author showed that a simple sequential inhibition model can be placed in the context of *finite sequential spatial processes*. Realisations of such a process consist of a finite sequence

$$\boldsymbol{x} = (x_1, \ldots, x_n), \qquad n \in \mathbb{N}_0,$$

of points in some bounded, convex open subset D of $\mathbb{R}^d$ with a nonempty interior. The family of all such configurations is denoted by N^{f}. The notation $n(\boldsymbol{x})$ is used for the length of $\boldsymbol{x}$, and $\boldsymbol{x}_{\leq i}$ is used to denote the subsequence $(x_1, \ldots, x_i)$, $i = 0, \ldots, n$. Note that for $i = 0$, $\boldsymbol{x}_{\leq i}$ has length 0.

Suppose that D is equipped with the Euclidean metric ρ, which defines the topology, and the Borel σ-algebra, denoted by $\mathcal{B}$. Write $\rho(z, A) := \inf\{\rho(z, a) : a \in A\}$ for the distance between $z \in D$ and $A \in \mathcal{B}$. The space N^{f} is equipped with the σ-algebra $\mathcal{N}^{\mathrm{f}}$ generated by the Borel product σ-algebras. The notation $\mu(A)$ is used for the Lebesgue measure of A.

Definition 1. The random sequential adsorption model is a jump process with values in N^{f} on the time interval $[0, \theta]$, starting from an empty vector, with transition rate

$$b(z, \boldsymbol{x}) = \mathbf{1}_{\{\rho(z,\boldsymbol{x})>r\}}$$

with respect to μ for a birth from $\boldsymbol{x}$ to $(\boldsymbol{x}, z)$. Here $r > 0$ is the hard core distance and $\theta \geq 0$ is the time horizon.

A variation that allows for spatial heterogeneity is to set

$$b(z, \boldsymbol{x}) = \pi(z) \, \mathbf{1}_{\{\rho(z,\boldsymbol{x})>r\}},$$

where $\pi(\cdot)$ is a strictly positive probability density on $(D, \mathcal{B})$. Thus, regions with high π-masses are favoured over those with smaller masses. Realisations tend to have particles with low indices concentrated in high-density regions, whereas latecomers (i.e. particles late in sequence) find themselves with locations with smaller π-masses. In the classic set-up of Definition 1, such a competitive advantage is not present.

Interaction structures other than hard core repulsion can be modelled by suitable choices of $b(z, \boldsymbol{x})$; see, e.g. [7]. Such models are known as *cooperative sequential adsorption models*, and can allow for both attraction and repulsion of new particles by previously arrived ones.

Example 1. Consider the limit case $r = 0$, where particles arrive according to a unit-rate Poisson process, that is, $b \equiv 1$ μ-almost everywhere. The number of particles adsorbed in D up to time θ is Poisson distributed with parameter $\theta \mu(D)$. In other words, the probability of exactly n transitions occurring in $[0, \theta]$ is given by

$$q_n = \frac{e^{-\theta \mu(D)}}{n!} (\theta \mu(D))^n$$

for $n = 0, 1, 2, \ldots$, and, given that n particles arrived up to time θ, their joint probability distribution has density $p_n := 1/\mu(D)^n$ with respect to μ^n. Since p_n does not depend on θ, the maximum likelihood estimator is readily derived from

$$\frac{\partial}{\partial \theta} \log q_n = -\mu(D) + \frac{n}{\theta},$$

which equals 0 if and only if $\theta = n/\mu(D) =: \hat{\theta}$. A little closer examination reveals that $\hat{\theta}$ is indeed the unique optimiser.

In order to be able to study maximum likelihood estimation in the presence of an unknown hard core distance, some results on the convolution of exponential distributions are required, as outlined in the next section.

3. Convolution of exponential distributions

The purpose of this section is to recall a result presented as an exercise in [9]. Here we give a proof for the sake of completeness and to correct a misprint in formula $(*)$ of [9, p. 40].

Lemma 1. *For $n \geq 1$, let $X_1, \ldots, X_n$ be independent, exponentially distributed random variables with distinct rates $\lambda_1, \ldots, \lambda_n$. Then the sum $X_1 + \cdots + X_n$ has probability density*

$$f_{X_1 + \cdots + X_n}(t) = (-1)^{n+1} \sum_{i=1}^{n} \lambda_i e^{-\lambda_i t} \prod_{j \neq i} \frac{\lambda_j}{\lambda_i - \lambda_j}, \qquad t \geq 0.$$

Proof. We proceed by induction, noting that the formula holds for $f_{X_1 + \cdots + X_n}(t)$ if $n = 1$ or $n = 2$. Suppose, then, that Lemma 1 is true up to some $n \geq 2$, and consider the random variable $X_1 + \cdots + X_{n+1}$. Without loss of generality, assume that $\lambda_1 > \lambda_2 > \cdots > \lambda_{n+1} > 0$. By the convolution formula, $f_{X_1 + \cdots + X_{n+1}}(t)$ is given by

$$\int_0^t (-1)^{n+1} \sum_{i=1}^{n} \prod_{\{j \, : \, n \geq j \neq i\}} \frac{\lambda_j}{\lambda_i - \lambda_j} \lambda_i e^{-\lambda_i s} \lambda_{n+1} e^{-\lambda_{n+1}(t-s)} \, \mathrm{d}s$$

$$= (-1)^{n+1} \sum_{i=1}^{n} \prod_{\{j \, : \, n \geq j \neq i\}} \frac{\lambda_j}{\lambda_i - \lambda_j} \int_0^t \lambda_i e^{-\lambda_i s} \lambda_{n+1} e^{-\lambda_{n+1}(t-s)} \, \mathrm{d}s.$$

Now apply the lemma for $n = 2$ to the integral on the right-hand side of this equation, to obtain

$$f_{X_1+\cdots+X_{n+1}}(t) = (-1)^{n+1} \sum_{i=1}^{n} \prod_{\{j:\, n \geq j \neq i\}} \frac{\lambda_j}{\lambda_i - \lambda_j} \left\{ \frac{-\lambda_i \lambda_{n+1}}{\lambda_i - \lambda_{n+1}} e^{-\lambda_i t} + \frac{-\lambda_i \lambda_{n+1}}{\lambda_{n+1} - \lambda_i} e^{-\lambda_{n+1} t} \right\}$$

$$= (-1)^{n+2} \sum_{i=1}^{n} \lambda_i e^{-\lambda_i t} \prod_{\{j:\, n+1 \geq j \neq i\}} \frac{\lambda_j}{\lambda_i - \lambda_j}$$

$$+ (-1)^{n+2} e^{-\lambda_{n+1} t} \left\{ \prod_{j=1}^{n+1} \lambda_j \right\} \sum_{i=1}^{n} \frac{-1}{\prod_{\{j:\, n+1 \geq j \neq i\}} (\lambda_i - \lambda_j)}.$$

The scalar factor associated with $e^{-\lambda_{n+1} t}$ can be identified by multiplying $f_{X_1+\cdots+X_{n+1}}(t)$ by $e^{\lambda_{n+1} t}$ and taking the limit as $t \to \infty$. Hence, a symmetry argument (interchange X_1 and X_{n+1} in the computations above) completes the proof.

In the context of this paper, λ_i will be the total birth rate associated with sequences $\mathbf{x}_{\leq i-1}$ of increasing length, and it is natural to consider sums of exponentially distributed random variables with strictly decreasing rates. However, it is possible to derive convolution densities for all finite sums of exponentially distributed random variables by noting that

$$Y = X_1 + \cdots + X_n = Z_1 + \cdots + Z_m$$

may be rewritten in terms of Erlang-distributed random variables Z_j, $j = 1, \ldots, m$, by pooling those X_i with equal λ_i. The convolution density, f_Y, of the sum of m independent Erlang(k_j)-distributed random variables with distinct rates λ_j is known explicitly [12]:

$$f_Y(t) = \prod_{j=1}^{m} (-\lambda_j)^{k_j} \sum_{j=1}^{m} \sum_{r=1}^{k_j} (-1)^r \frac{b_{jr}}{(r-1)!} t^{r-1} e^{-\lambda_j t}, \tag{1}$$

where

$$b_{jr} = \lim_{t \to \lambda_j} \frac{1}{(k_j - r)!} \frac{\partial^{k_j - r}}{\partial t^{k_j - r}} \left[(t - \lambda_j)^{k_j} \prod_{i=1}^{m} \left(\frac{1}{t - \lambda_i} \right)^{k_i} \right].$$

A probability density f on $(0, \infty)$ is said to be *log-concave* if it can be written as $f = e^{\Psi}$ for some concave function $\Psi \colon (0, \infty) \to [-\infty, \infty)$. Clearly, the exponential distribution has a log-concave density. Since the convolution of log-concave probability densities is log-concave [14], so is $f_{X_1+\cdots+X_n}$ defined above. Consequently, $f_{X_1+\cdots+X_n}$ is unimodal in the sense that the sets $\{t \colon f_{X_1+\cdots+X_n}(t) \geq a\}$ are convex for all $a \geq 0$, a fact that will be used in the proof of Proposition 2 in the next section.

4. Maximum likelihood estimation

The random sequential adsorption model has two parameters: r and θ. The purpose of this section is to study maximum likelihood estimation. In order to do so, we derive a Radon–Nikodým derivative f (see Theorem 1) with respect to the distribution of a sequence of (mean-$\mu(D)$) Poisson length with independently, uniformly distributed components, thus placing the random sequential adsorption model into the framework of Markov sequential spatial processes [21], [22]. We proceed to establish the existence and uniqueness of

$$(\hat{r}, \hat{\theta}) := \operatorname{argsup}\{ f(\mathbf{x}; r, \theta) \colon r > 0, \theta \geq 0 \}$$

(see Propositions 1 and 2).

Theorem 1. *For the model of Definition 1, the distribution of particles adsorbed in the time interval $[0, \theta]$ is absolutely continuous with respect to the measure v on $(N^{\mathrm{f}}, \mathcal{N}^{\mathrm{f}})$ defined by*

$$v(F) := \sum_{n=0}^{\infty} \frac{\mathrm{e}^{-\mu(D)}}{n!} \int_{D} \cdots \int_{D} \mathbf{1}_{\{(x_1,\ldots,x_n)\in F\}} \, \mathrm{d}x_1 \cdots \mathrm{d}x_n$$

for $F \in \mathcal{N}^{\mathrm{f}}$. For $n = 0$, the Janossy density at the empty vector is $j_0(\varnothing) = \mathrm{e}^{-\theta\mu(D)}$. With the notation $B(\boldsymbol{x}) := \int_D b(z, \boldsymbol{x}) \, \mathrm{d}z$ for the total birth rate, n-point Janossy densities j_n, $n \geq 1$, at $\boldsymbol{x} = (x_1, \ldots, x_n) \in D^n$ are as follows.

- $j_n(\boldsymbol{x}) = 0$ *if $\rho(x_i, x_j) \leq r$ for some $i, j \in \{1, \ldots, n\}$, $i \neq j$.*

- *If $\rho(x_i, x_j) > r$ for all $i \neq j$ and $B(\boldsymbol{x}) > 0$, then*

$$j_n(\boldsymbol{x}) = \frac{g_{n+1}(x_1, \ldots, x_n; \theta)}{B(x_1, \ldots, x_n)} \prod_{i=1}^{n} \frac{b(x_i, \boldsymbol{x}_{\leq i-1})}{B(\boldsymbol{x}_{\leq i-1})},$$

  where $g_{n+1}(x_1, \ldots, x_n; \theta)$ is the probability density (see Lemma 1 and (1)) at θ of the sum of $n + 1$ independent random variables that are exponentially distributed with rates $B(\boldsymbol{x}_{\leq i-1})$, $i = 1, \ldots, n+1$.

- *If $\rho(x_i, x_j) > r$ for all $i \neq j$ and $B(\boldsymbol{x}) = 0$, then*

$$j_n(\boldsymbol{x}) = G_n(x_1, \ldots, x_{n-1}; \theta) \prod_{i=1}^{n} \frac{b(x_i, \boldsymbol{x}_{\leq i-1})}{B(\boldsymbol{x}_{\leq i-1})},$$

  where $G_n(x_1, \ldots, x_{n-1}; \theta)$ is the cumulative probability distribution function at θ of the sum of n independent random variables that are exponentially distributed with rates $B(\boldsymbol{x}_{\leq i-1})$, $i = 1, \ldots, n$.

The various cases listed in Theorem 1 correspond respectively to impossible configurations which contain a pair of particles that violate the hard core condition, to configurations with room for additional particles, and to saturated configurations. The n-point Janossy density may be interpreted heuristically as the probability of a sequence of n particles being successively located in infinitesimal regions centred at $x_1, \ldots, x_n$. A Radon–Nikodým derivative with respect to v is given by $f(\boldsymbol{x}) = \mathrm{e}^{\mu(D)} n(\boldsymbol{x})! \, j_{n(\boldsymbol{x})}(\boldsymbol{x})$; see [2, Chapter 5.3], [21], and [22].

Proof of Theorem 1. Since $j_0(\varnothing)$ is the probability that no transition occurs up to time θ, the formula for $n = 0$ holds. Thus, from now on let $n \geq 1$. Note that the assumption that $\rho(x_i, x_j) > r$ for all $i \neq j \in \{1, \ldots, n\}$ implies that $B(\boldsymbol{x}_{\leq n-1}) > 0$, since the particle x_n may be added to $\boldsymbol{x}_{\leq n-1}$ without violation of the hard core condition and sets of the form $\{z \in D : \rho(z, x_j) > r\}$ are open. Moreover, under the same assumption, the total birth rates $B(\boldsymbol{x}_{\leq i-1})$ are strictly decreasing in $i = 1, \ldots, n$. To see this, note that $B(\boldsymbol{x}_{\leq i-1}) - B(\boldsymbol{x}_{\leq i})$ is the volume of the set $\{z \in D : \rho(z, x_k) > r \text{ for } k \leq i - 1, \rho(z, x_i) \leq r\}$. Since this set contains x_i in its interior, its volume is strictly positive. Consequently, $B(\boldsymbol{x}_{\leq n-1}) > 0$ implies that all $B(\boldsymbol{x}_{\leq i-1})$, $i < n$, are strictly positive; hence, the ratios in the posited Janossy densities are well defined. In order to compute $j_n(\boldsymbol{x})$, note that the first exponentially distributed waiting time (with rate $B(\varnothing) = \mu(D)$) must be less than θ; the first particle has to be located in $\mathrm{d}x_1$; all subsequent cumulative waiting times must fall before θ until the last particle, in $\mathrm{d}x_n$, is

selected; and no further transitions may take place within the time horizon θ. Integration with respect to the waiting times yields

$$\prod_{i=1}^{n} \frac{b(x_i, \boldsymbol{x}_{\leq i-1})}{B(\boldsymbol{x}_{\leq i-1})}$$

$$\times \int_0^\theta \int_0^{\theta-s_1} \cdots \int_0^{\theta-\sum_{i=1}^{n-1} s_i} \prod_{i=1}^{n} f_{B(\boldsymbol{x}_{\leq i-1})}(s_i) \left(1 - F_{B(\boldsymbol{x}_{\leq n})}\left(\theta - \sum_{i=1}^{n} s_i\right)\right) \mathrm{d}s_1 \cdots \mathrm{d}s_n \tag{2}$$

with the notation f_λ for the density of an exponential distribution of rate λ and F_λ for the corresponding cumulative distribution function.

In order to evaluate (2), first consider the unsaturated case, where $B(\boldsymbol{x}) > 0$, and proceed by induction to prove that the multiple integral in (2) is equal to $g_{n+1}(x_1, \ldots, x_n; \theta)$ divided by $B(x_1, \ldots, x_n)$. For $n = 1$,

$$\int_0^\theta f_{B(\varnothing)}(s)(1 - F_{B(x)}(\theta - s)) \,\mathrm{d}s = \int_0^\theta f_{B(\varnothing)}(s) f_{B(x)}(\theta - s) \frac{(1 - F_{B(x)}(\theta - s))}{f_{B(x)}(\theta - s)} \,\mathrm{d}s$$

$$= \frac{g_2(x; \theta)}{B(x)},$$

as an exponential distribution has constant hazard equal to its rate. Next, suppose that (2) holds for $n \geq 1$ and consider the $n + 1$ case. Then

$$\int_0^\theta f_{B(\varnothing)}(s_1) \,\mathrm{d}s_1 \int_0^{\theta-s_1} \cdots \int_0^{\theta-s_1-\sum_{i=2}^{n} s_i} \left\{ \prod_{i=2}^{n+1} f_{B(\boldsymbol{x}_{\leq i-1})}(s_i) \right\}$$

$$\times \left(1 - F_{B(\boldsymbol{x}_{\leq n+1})}\left(\theta - s_1 - \sum_{i=2}^{n+1} s_i\right)\right) \mathrm{d}s_2 \cdots \mathrm{d}s_{n+1}$$

$$= \int_0^\theta f_{B(\varnothing)}(s_1) \frac{g_{n+1}(x_1, \ldots, x_{n+1}; \theta - s_1)}{B(\boldsymbol{x}_{\leq n+1})} \,\mathrm{d}s_1$$

$$= \frac{g_{n+2}(x_1, \ldots, x_{n+1}; \theta)}{B(\boldsymbol{x}_{n+1})}$$

by the induction assumption.

To conclude the proof, note that in the saturated case, where $B(\boldsymbol{x}) = 0$, the survival probability $1 - F_{B(\boldsymbol{x})}$ is identically equal to 1 on $[0, \infty)$. Hence, the integral in (2) is equal to $G_n(x_1, \ldots, x_{n-1}; \theta)$.

Next we turn to statistical inference. From now on, where necessary, the notation will take into explicit account the parameters r and θ.

Example 2. Upon observation of the empty set, $j_0(\varnothing; r, \theta) = \mathrm{e}^{-\theta \mu(D)}$ is strictly decreasing in θ and does not depend on r. Hence, we cannot carry out inference on the interaction distance, whereas the maximum likelihood estimator of the time horizon is $\hat{\theta}(\varnothing) = 0$.

If the data sequence is a singleton, (x), then

$$j_1(x; r, \theta) = \int_0^\theta \exp[-B(\varnothing)s - B(x; r)(\theta - s)] \,\mathrm{d}s.$$

For any given θ, the exponent in the integrand is strictly increasing in r on the interval $(0, \sup_{z \in D} \rho(x, z))$, with limit value $(1 - e^{-\mu(D)\theta})/\mu(D)$ for larger r. Identifying the interval $[\sup_{z \in D} \rho(x, z), \infty)$ with ∞, we obtain $\hat{r}(x) = \infty$. Optimising over θ yields $\hat{\theta}(x) = \infty$.

As for the classic hard core point process, upon observation of a nonempty sequence, r may be estimated using the minimum interparticle distance [16].

Proposition 1. *For the model of Definition 1, suppose that a sequence $x = (x_1, \ldots, x_n)$, $n \geq 2$, is observed. Then $\hat{r}(x) = r_{\min}(x) := \min_{i < j} \rho(x_i, x_j)$ is the maximum likelihood estimator of the hard core distance for any value of $\theta \geq 0$.*

The supremum argument is not attained (i.e. it is not a maximum argument), as indicated by the proof.

Proof of Proposition 1. Note that $j_n(x; r, \theta) = 0$ unless $\rho(x_i, x_{\leq i-1}) > r$ for all $i = 1, \ldots, n$. Therefore, $\hat{r}(x) \leq r_{\min}(x)$. Now, (2) can be expressed as

$$I(s_1, \ldots, s_n; r, \theta) \prod_{i=1}^{n} b(x_i, x_{\leq i-1}; r),$$

where

$$I(s_1, \ldots, s_n; r, \theta)$$
$$:= \int_0^{\theta} \cdots \int_0^{\theta - \sum_{i=1}^{n-1} s_i} \exp\left[-\sum_{i=1}^{n} s_i B(x_{\leq i-1}; r) - \left(\theta - \sum_{i=1}^{n} s_i\right) B(x; r) \right] ds_1 \cdots ds_n.$$

If $0 < r_1 < r_2 < r_{\min}(x)$ then $b(x_i, x_{\leq i-1}; r_1) = b(x_i, x_{\leq i-1}; r_2)$, while it follows from the convexity of D and the fact that x_n may be added to $x_{\leq i-1}$ under a hard core distance-r_2 constraint that $B(x_{\leq i-1}; r_1) > B(x_{\leq i-1}; r_2)$ for all $i = 2, \ldots, n$. Hence, $I(s_1, \ldots, s_n; r_1, \theta) < I(s_1, \ldots, s_n; r_2, \theta)$, which implies that the maximum likelihood estimator is as stated.

It remains to estimate the time horizon, θ.

Example 3. As we saw in Example 2, for $n = 0$ and fixed r, we have $\hat{\theta} = 0$. For $n \geq 1$, consider the case where, for a given r, we observe a sequence x of length n such that $j_n(x; \theta)$ is strictly positive but $B(x) = 0$. To optimise $j_n(x; \theta)$ over θ, according to Theorem 1 we have to optimise $G_n(x_1, \ldots, x_{n-1}; \theta)$, which, being a cumulative distribution function, is monotonically increasing in θ. Furthermore, by the convolution formula, it is easily seen that $G_n(x_1, \ldots, x_{n-1}; t) < 1$ for any finite $t \geq 0$. Hence, $\hat{\theta}(x) = \infty$.

Proposition 2. *For the model of Definition 1 with fixed $r > 0$, suppose that a sequence $x = (x_1, \ldots, x_n)$, $n \geq 1$, is observed for which $B(x)$, $j_n(x; \theta) > 0$. Then the maximum likelihood estimator*

$$\hat{\theta} \equiv \hat{\theta}(x_1, \ldots, x_n) := \operatorname*{argmax}_{\theta \geq 0} g_{n+1}(x_1, \ldots, x_n; \theta)$$

exists, is unique, and solves the score equation

$$\frac{\partial}{\partial \theta} g_{n+1}(x_1, \ldots, x_n; \theta) = B(x_1, \ldots, x_n)[g_n(x_1, \ldots, x_{n-1}; \theta) - g_{n+1}(x_1, \ldots, x_n; \theta)] = 0.$$

Proof. According to Theorem 1, the maximum likelihood estimator for the time horizon is $\hat{\theta}(x_1, \ldots, x_n) := \mathrm{argsup}_{\theta \geq 0}\, g_{n+1}(x_1, \ldots, x_n; \theta)$, provided that the supremum argument exists. Because the exponential distribution is log-concave as a function of $\theta > 0$, so too is $g_{n+1}(x_1, \ldots, x_n; \cdot)$, from a result of [14]. Consequently, $g_{n+1}(x_1, \ldots, x_n; \cdot)$ is unimodal. Moreover, as it is continuous on $[0, \infty)$ with

$$g_{n+1}(x_1, \ldots, x_n; 0) = 0 = \lim_{t \to \infty} g_{n+1}(x_1, \ldots, x_n; t),$$

the supremum argument is attained.

Next consider the score function, and write $f_\lambda(t) = \lambda e^{-\lambda t}$ for $t \geq 0$. By the convolution formula,

$$g_{n+1}(x_1, \ldots, x_n; \theta) = \int_0^\theta g_n(x_1, \ldots, x_{n-1}; s)\, f_{B(x_{\leq n})}(\theta - s)\, \mathrm{d}s.$$

The functions f_λ and $g_n(x_1, \ldots, x_{n-1}; \cdot)$ are uniformly bounded, so by dominated convergence we have

$$\lim_{h \downarrow 0} \frac{1}{h} \int_0^\theta g_n(x_1, \ldots, x_{n-1}; s)[f_{B(x_{\leq n})}(\theta + h - s) - f_{B(x_{\leq n})}(\theta - s)]\, \mathrm{d}s$$
$$= \int_0^\theta g_n(x_1, \ldots, x_{n-1}; s)\, f'_{B(x_{\leq n})}(\theta - s)\, \mathrm{d}s.$$

Furthermore, the continuity of f_λ and the fundamental theorem of analysis imply that

$$\lim_{h \downarrow 0} \frac{1}{h} \int_\theta^{\theta + h} g_n(x_1, \ldots, x_{n-1}; s)\, f_{B(x_{\leq n})}(\theta + h - s)\, \mathrm{d}s = g_n(x_1, \ldots, x_{n-1}; \theta)\, f_{B(x_{\leq n})}(0).$$

Hence, the derivative of $g_{n+1}(x_1, \ldots, x_n; \theta)$ can be written as

$$- B(x_{\leq n}) \int_0^\theta g_n(x_1, \ldots, x_{n-1}; s)\, f_{B(x_{\leq n})}(\theta - s)\, \mathrm{d}s + B(x_{\leq n}) g_n(x_1, \ldots, x_{n-1}; \theta)$$
$$= B(x_{\leq n})[g_n(x_1, \ldots, x_{n-1}; \theta) - g_{n+1}(x_1, \ldots, x_n; \theta)].$$

A similar argument for the limit $h \uparrow 0$ completes the score calculation.

Finally, we must show that the maximum likelihood estimator is unique, that is, that the function to be optimised has no plateaux. This can be done by induction, as the statement clearly holds for $n = 1$. Suppose, then, that unique maximum likelihood estimators exist for sequences of length at most $n - 1$, but that for the sequence $(x_1, \ldots, x_n)$ the function $g_{n+1}(x_1, \ldots, x_n; t)$ is constant on the interval (t_0, t_1), $t_0 < t_1$. Then its derivative equals 0 on (t_0, t_1), which by the score equation implies that $g_n(x_1, \ldots, x_{n-1}; t)$ is constant for $t \in (t_0, t_1)$. Since that would contradict the induction assumption, g_{n+1} cannot have any plateaux either.

Proposition 2 holds in particular for $r = r_{\min}(x)^-$, the range just before the minimum interparticle distance. Both Theorem 1 and Proposition 2 hold in the Poisson case, where $r = 0$, as well (see Example 1).

Note that $g_1(\varnothing; 0) = B(\varnothing)$ and that $g_{n+1}(x_1, \ldots, x_n; 0) = 0$ for $n \geq 1$. Thus, for $n = 1$ the latter function has a boundary minimum at 0 that does not correspond to a solution to the score equation. For $n \geq 2$, $\theta = 0$ is a solution to the score equation, and corresponds to the global minimum.

5. Conclusion

In this paper we have proved that the maximum likelihood method can be applied to estimate the parameters in random sequential adsorption models. The set-up can be generalised to metrics that are equivalent to the Euclidean metric, or indeed to general metric spaces, provided that the total birth rate is a strictly decreasing and positive function of both the sequence and the range until saturation.

Were the inhomogeneity measure π in the discussion that follows Definition 1 to contain nuisance parameters, these could be estimated separately from observable covariates, or by profile likelihood methods.

We tested the approach in practice for a range of parameter values, inhomogeneity measures, and metrics. Some software and results can be found online at http://db.cwi.nl/projecten/project.php4?prjnr=187.

It is important to note that the method is not restricted to the random sequential adsorption model, but can be applied in principle to estimate the time horizon in other cooperative sequential inhibition processes [7]. In particular, the sequential spatial process framework applies, so densities can be derived as in Theorem 1 provided that proper care is taken of any zero rates that might arise.

Acknowledgements

References

[1] CRESSIE, N. A. C. (1991). *Statistics for Spatial Data*. John Wiley, New York.
[2] DALEY, D. J. AND VERE-JONES, D. (1988). *An Introduction to The Theory of Point Processes*. Springer, New York.
[3] DIGGLE, P. J. (1983). *Statistical Analysis of Spatial Point Patterns*. Academic Press, London.
[4] DIGGLE, P. J., BESAG, J. AND GLEAVES, J. T. (1976). Statistical analysis of spatial point patterns by means of distance methods. *Biometrics* **32,** 659–667.
[5] DÖGE, G. (2001). Perfect simulation for random sequential adsorption of d-dimensional spheres with random radii. *J. Statist. Comput. Simul.* **69,** 141–156.
[6] DVORETZKY, A. AND ROBBINS, H. (1964). On the 'parking' problem. *Publ. Math. Res. Inst. Hungar. Acad. Sci.* **9,** 209–225.
[7] EVANS, J. W. (1993). Random and cooperative sequential adsorption. *Rev. Modern Phys.* **65,** 1281–1329.
[8] FEDER, J. (1980). Random sequential adsorption. *J. Theoret. Biol.* **87,** 237–254.
[9] FELLER, W. (1970). *An Introduction to Probability Theory and Its Applications*, Vol. II, 2nd edn. John Wiley, New York.
[10] HALL, P. G. (1988). *Introduction to the Theory of Coverage Processes*. John Wiley, New York.
[11] LOTWICK, H. W. (1982). Simulations of some spatial hard core models, and the complete packing problem. *J. Statist. Comput. Simul.* **15,** 295–314.
[12] MATHAI, A. M. (1982). Storage capacity of a dam with gamma type inputs. *Ann. Inst. Statist. Math.* **34,** 591–597.
[13] PENROSE, M. D. (2001). Random parking, sequential adsorption, and the jamming limit. *Commun. Math. Phys.* **218,** 153–176.
[14] PRÉKOPA, A. (1973). On logarithmic concave measures and functions. *Acta Sci. Math. (Szeged)* **34,** 335–343.
[15] RÉNYI, A. (1958). On a one-dimensional problem concerning random space filling. *Publ. Math. Inst. Hungar. Acad. Sci.* **3,** 109–125.
[16] RIPLEY, B. D. (1988). *Statistical Inference for Spatial Processes*. Cambridge University Press.
[17] SENGER, B., VOEGEL, J.-C. AND SCHAAF, P. (2000). Irreversible adsorption of colloidal particles on solid substrates. *Colloids Surfaces A* **165,** 255–285.

[18] STOYAN, D. (1998). Random sets: models and statistics. *Internat. Statist. Rev.* **66,** 1–27.
[19] TANEMURA, M. (1979). On random complete packing by disks. *Ann. Inst. Statist. Math.* **31,** 351–365.
[20] UPTON, G. J. G. AND FINGLETON, B. (1985). *Spatial Data Analysis by Example*, Vol. I, *Point Pattern and Quantitative Data*. John Wiley, Chichester.
[21] VAN LIESHOUT, M. N. M. (2006). Campbell and moment measures for finite sequential spatial processes. In *Proc. Prague Stoch. 2006*, eds M. Hušková and M. Janžura, Matfyzpress, Prague, pp. 215–224.
[22] VAN LIESHOUT, M. N. M. (2006). Markovianity in space and time. In *Dynamics and Stochastics: Festschrift in Honour of M. S. Keane* (Lecture Notes Monogr. Ser. **48**), eds D. Denteneer, F. den Hollander and E. Verbitskiy, Institute for Mathematical Statistics, Beachwood, OH, pp. 154–168.