*Quaderni di Statistica*
*Vol. 4, 2002*

# Simultaneous confidence bands for pair correlation functions in Markov point processes

Andrea Pallini
*Dipartimento di Scienze Statistiche, Università di Bologna, Italy*
*E-mail: pallini@stat.unibo.it*

*Summary*: This article considers simultaneous bootstrap confidence bands for correlation functions in Markov point processes, with a bounded interaction radius $r$. The resampling scheme is defined by a preliminary subdivision of the observed region of the point process into a set of conditionally independent subregions $\{B_v\}$, selected at distance $d \geq 2r$. Bootstrap confidence bands are obtained by resampling the suprema of kernel density estimates calculated from all these subregions, as conditionally independent replicates.

*Key words*: bootstrap; conditional independence; correlation function; kernel density estimation; pair correlation function; point processes; simultaneous confidence band; spatial dependence.

## *1. Introduction and basic notation*

Stationary Markov point processes are relevant stochastic models for the statistical analysis of configurations of points in a planar region $A \subset \mathbb{R}^2$ (Isham, 1984; Baddeley and Möller, 1989). In forestry statistics, for example, the region $A$ may be a forest stand with different species of trees and the points are the observed positions of the trees in $A$ (cf. Penttinen *et al.*, 1992, Penttinen and Stoyan, 2000, and Pallini, 2002). The theory of Markov densities is useful for characterizing the neighbors of each

observed point and the point process itself. Correlation functions are typically estimated for analyzing the attraction or repulsion (interactions) of points of the same type, or the attraction or repulsion between points of different types.

Let $\Omega$ be the sample space of all finite subsets in $A$, representing all configurations $x_A$ of points in $A$. Finite point processes with absolutely continuous distributions (with respect to the Poisson process on $A$ with mean $\lambda(\cdot)$) can be specified by their density function $f : \Omega \to [0, \infty)$. Suppose that the cardinality of a point configuration $x_A \in \Omega$ is $n$; that is, $x_A = \{x_1, \ldots, x_n\}$, where point $x_i$ has components $x_i^{(1)}$ and $x_i^{(2)}$, generated by the process in $A$. A pairwise interaction point process is a process in $A$ with density

$$f(x_A) = \alpha\beta^n \prod_{i<j} \gamma(d(x_i, x_j)), \tag{1}$$

where $d$ is the Euclidean distance, $\alpha, \beta > 0$ and $\gamma(d(x_i, x_j))$ is a nonnegative, integrable, real function, modelling interaction between the pair of points $x_i$ and $x_j$. More specifically, $\alpha$ is the normalizing constant. Various functions $\gamma(d(x_i, x_j))$ are reviewed and studied in Ripley (1988). A relevant example of function $\gamma(d(x_i, x_j))$ is the Strauss interaction

$$\gamma(d(x_i, x_j)) = \begin{cases} \gamma & \text{if} \quad d(x_i, x_j) < r \\ 1 & \text{if} \quad d(x_i, x_j) \geq r \end{cases}, \tag{2}$$

where $0 \leq \gamma \leq 1$ and $r > 0$ (Kelly and Ripley, 1976). Interaction (2) in (1) defines the Strauss point process with density

$$f(x_A) = \alpha\beta^n \gamma^{\theta_r(x_A)}, \tag{3}$$

where $\theta_r(x_A)$ is the number of pairs of points in $A$, which lie within a distance $r$. Assume that the interaction has bounded interaction radius $r$, in the sense that $\gamma(d(x_i, x_j)) = 1$ occurs for all $d(x_i, x_j) \geq r$. Two points $x_i$ and $x_j$ are neighbors with respect to a symmetric, reflexive relation $\sim$ on $A$ (neighborhood relation), $x_i \sim x_j$, if $d(x_i, x_j) < r$. Generally speaking, interactions occur only between neighbors. Density (1) may

then be expressed as

$$f(x_A) = \alpha\beta^n \prod_{x_i \sim x_j\,,\, i<j} \gamma(d(x_i, x_j))\,. \tag{4}$$

The number $\theta_r(x_A)$ in Strauss density (3) may be regarded as the number of pairs of neighbors. This density may also be defined from (1) by taking $\gamma(d(x_i, x_j)) = \gamma$. A nearest-neighbor Markov point process on $A$ is a finite point process with density $f$, such that $f$ is a Markov function (Baddeley and Møller, 1989). A function $f : \Omega \to [0, \infty)$ is a Markov function if and only if there is an interaction function $\varphi$, such that

$$f(x_A) = \prod_{y \subset x_A} \varphi(y)\,,$$

for all $x_A \in \Omega$. Density $f$ given by (1) is a Markov function with respect to the relation $x_i \sim x_j$. In particular, $\varphi(\emptyset) = \alpha$, where $\emptyset$ is the empty configuration, $\varphi(\{x_i\}) = \beta$, $\varphi(\{x_i, x_j\}) = \gamma(d(x_i, x_j))$ for 2 neighbors, and $\varphi(y) = 1$ for 3 or more neighbors.

In Pallini (2000, 2002), a resampling scheme for Markov (Gibbs) point processes of bounded interaction radius $r$ is studied, from an original sample of conditionally independent replicates of the statistic of interest.

In this paper, we attempt to obtain simultaneous bootstrap confidence bands for pair correlation functions by resampling the suprema of kernel density estimates calculated from a set of conditionally independent subregions (at distance $d \geq 2\,r$) in region $A$.

### *1.1. Gibbs random fields*

Pairwise interaction point processes given by densities (4) are Gibbs point processes. A Gibbs point process may typically be viewed as a Gibbs random field, and vice versa. See Preston (1976), Chapter 6, and Klein (1982). A point configuration $x_A$ can be split into a set of point configurations, organized as configurations around the sites of a Gibbs field.

Let us suppose that the observed region $A$ is rectangular. We can partition $A$ into several rectangles $\{A_v\}$, each containing a subset of $x_A$. Let $\left\{X_v,\, v \in \mathbf{Z}^2\right\}$ be a Gibbs random field. Denote by $e$ a real value, a product of the components $e^{(1)}$ and $e^{(2)}$, $e = e^{(1)}e^{(2)}$, such that $\min(e^{(1)}, e^{(2)}) \geq r$, where $r$ is the interaction radius. For every $v \in \mathbf{Z}^2$, define rectangles $\{A_v\}$ of size $e$ as

$$A_v = \left\{t \in \mathbf{R}^2 \,:\, e^{(j)}\left(v^{(j)} - 0.5\right) \leq t^{(j)} < e^{(j)}\left(v^{(j)} + 0.5\right)\,,\; j = 1, 2\right\}\,, \tag{5}$$

where $t^{(j)}$ are the components of a generic point $t$ in $A_v$. Rectangles $\{A_v\}$, $v \in \mathbf{Z}^2$, partition $\mathbf{R}^2$, while $e$ determines their common area.

We can transform the Gibbs point process observed in $A$ into the Gibbs field $\left\{X_v,\, v \in \mathbf{Z}^2\right\}$, by taking the point configuration $x_{A_v}$ in rectangle (5) as the site variable $X_z$, $X_v = x_{A_v}$. The spatial dependence between sites in $\left\{X_v,\, v \in \mathbf{Z}^2\right\}$ is studied in Klein (1982).

### *1.2. A spatial Markov property*

Let us consider two rectangles $B_v$ and $B_z$ in observed region $A \subset \mathbf{R}^2$, for some $v, z \in \mathbf{Z}^2$. Suppose that $B_v$ and $B_z$ are at distance $d = d(B_v, B_z) \geq 2\,r$, where $r$ is the interaction radius. In the manner of Baddeley and Møller (1989), we denote by $N_v$ the neighborhood of $B_v$,

$$N_v = \{x_i \in A : x_j \sim x_i \; for \; some \; x_j \in B_v\},$$

$v \in \mathbf{Z}^2$. We define $W_v = B_v \cup N_v$, $v \in \mathbf{Z}^2$. A set $y \in \Omega$ is a clique if all points of $y$ are neighbors, $x_i \sim x_j$, for every $x_i, x_j \in y$. In Pallini (2002), it is seen that

$$\begin{aligned} f(x_A \cap W_v, x_A \cap W_z) &= \prod_{cliques\; y \in W_v} \varphi(y) \prod_{cliques\; y \in W_z} \varphi(y) \\ &= f(x_A \cap W_v)\, f(x_A \cap W_z)\,. \end{aligned} \tag{6}$$

See also Pallini (2000), for an alternative proof.

The restriction of a finite Gibbs point process to $x_A \cap W_v$ is conditionally independent of the restriction to $x_A \cap W_z$, given the restriction

to the boundary $(x_A \setminus (x_A \cap W_v)) \setminus (x_A \cap W_z)$ . Property (6) allows us to isolate subregions $\{B_z\}$ in observed region $A$ , which are defined by the union of contiguous rectangles (5) and contain conditionally independent configurations of points, according to a self-reproducing coding scheme.

### *1.3. Coding schemes*

We denote by $Z$ the set of sites in $\mathbf{Z}^2$, such that $A = \bigcup_{v \in Z} A_v$ . Let $q$ be the number of sites in $Z$ . Since $A$ is rectangular, $q$ is a product of the integers $q^{(1)}$ and $q^{(2)}$, $q = q^{(1)}\, q^{(2)}$ .

Conditionally independent, rectangular subregions $\{B_v\}$ in $A$ , where $v \in E \subseteq Z$ , may be determined by defining fixed coding schemes.

Every subregion $B_v$ is the union of $p = p^{(1)}\, p^{(2)}$ contiguous rectangles (5) and contains the point configuration $x_{B_v}$ . Let $s = s^{(1)} s^{(2)}$ denote the designated number of subregions $\{B_v\}$ . By choosing subregions $\{B_v\}$ at Euclidean distance $d \geq r$ , the desired conditional independence is obtained, as implied by the Markov property (6).

### *1.4. Increasing domain asymptotics*

We refer to $q$ as the sample size defining all sequences of finite subsets of interest in $\mathbf{Z}^2$ (cf. Bolthausen, 1982). The size $e$ of rectangles $\{A_v\}$ given by (5) is taken as fixed. Let $\mid Z \mid$ be the cardinality $q$ of the set $Z$ in $\mathbf{Z}^2$ . We assume that $Z$ belongs to an infinite sequence $Z_q$ going up to $\mathbf{Z}^2$, such that $\mid Z_q \cup N \mid^{-1} \mid N \mid \to 0$ , as $q \to \infty$ , where neighborhood $N$ is defined as

$$N = \left\{ v \notin Z_q : max_{i=1,2}\, d(v^{(i)}, z^{(i)}) \; = 1 \; for \; every \; z \in Z_q \right\} .$$

We also assume condition (C1) (cf. Pallini, 2002) as follows.

(C1). $\quad p^{(i)} = O(q^{(i)})$ , $i = 1, 2$ , as $q \to \infty$ .

The coding scheme reproduces itself according to a fixed rule, which maintains $s^{(i)} p^{(i)} \leq q^{(i)}$ , so that $q^{-1} sp \to 1$ , as $q \to \infty$ . In particular,

$$p = c_1\, s^{1/2} , \qquad q = c_2\, s^{3/2} ,$$

where $c_1, c_2 > 0$, and $c_2^{-1} c_1 \to 1$, as $q \to \infty$. We write $q \Rightarrow \infty$, to indicate this increasing domain asymptotics.

Stationarity of the point process implies that means $\lambda(A) = \lambda a$ and $\lambda(B_v) = \lambda b$ (cf. Stoyan *et al.*, 1995, Chapter 4), $v \in E$, where the intensity $\lambda$ is a non-negative constant (possibly infinite), where $a = a^{(1)} a^{(2)}$ and $b = b^{(1)} b^{(2)}$ are the areas of $A$ and $B_v$. Observe that

$$\lambda(A \setminus (\cup_{v \in E} B_v)) = \lambda(q - sp)e \,,$$

where $e$ is the area of rectangles (5). Under condition (C1), it follows that $\lambda(q - sp)e \to 0$ and $a^{-1} b \to s^{-1}$, as $q \Rightarrow \infty$.

## *2. The pair correlation function*

Let us consider two infinitesimally small discs $D_1$ and $D_2$ of area $dD_1$ and $dD_2$ with intercenter distance $t$. The pair correlation function may be defined in terms of probabilities for points in $D_1$ and $D_2$ (cf. Daley and Vere-Jones, 1988, Chapter 5, and Stoyan *et al.*, 1995, Chapter 4). We indicate $\Pr(t)$ as the probability that both discs each contain a point.

The pair correlation function $g(t)$ is given by

$$\Pr(t) = \lambda^2 \, g(t) \, dD_1 \, dD_2 \,, \tag{7}$$

where $t \geq 0$; function $g(t)$ is real-valued.

The value $g(t) = 1$ means that the points are spatially uncorrelated. Whereas $g(t) > 1$, the interpoint distances around $t$ occur more frequently, and we can thus say that the underlying process exhibits attraction between points or a positive correlation. Whereas $g(t) < 1$, we can say that the process exhibits inhibition between points or a negative correlation.

### *2.1. Kernel density estimates*

Estimates for correlation functions in point processes may conveniently be obtained by kernel density estimation (cf. Penttinen *et al.*, 1992,

Doguwa and Kukoyi, 1993, Stoyan *et al.*, 1993, and Wand and Jones, 1995).

Point configuration $x_A = \{x_1, \ldots, x_n\}$ is observed in the region $A$. We estimate $\lambda^2$, where $\lambda$ is the intensity, with

$$\widehat{\lambda^2} = a^{-2}\, n(n-1)\,,$$

$a$ being the area of the observed region $A$. The pair correlation function $g(t)$, $t \geq 0$, given by (7), is typically estimated by

$$\widehat{g}(t) = \sum_{i \neq j\,;\, i,j=1,\ldots,n} \frac{k_h(d(x_i, x_j) - t)}{\widehat{\lambda^2}\, 2\pi t}\,, \tag{8}$$

where $k_h(u) = h^{-1}\, k(h^{-1}u)$ is a kernel function with bandwidth $h$.

A suitable choice (cf. Penttinen *et al.*, 1992; Stoyan *et al.*, 1993) for the kernel function $k_h(u)$, $|u| \leq h$, in kernel density estimates (8) is the Epanechnikov kernel,

$$k_h(u) = \frac{3}{4h}\left(1 - \frac{u^2}{h^2}\right). \tag{9}$$

Selection of bandwidth $h$ is discussed below in section 3.2.

### *2.2. Edge-corrections*

Treatment of edge-effects is crucial in estimating the pair correlation functions $g(t)$, $t \geq 0$. In kernel density estimates (8), a minus (plus)-sampling idea seems the most flexible. We suggest calculating (8) for $i$ labelling points in the restricted region $A \ominus r$ and $j$ labelling points in the whole region $A$.

Alternative edge-corrections are available, for instance, in Ripley (1988), Chapter 3, and Stoyan *et al.* (1995), Chapter 4.

### *3. Simultaneous bootstrap confidence bands*

For every $v \in E$, suppose that the cardinality of the point configuration $x_{B_v}$ in subregion $B_v$ of area $b$ is $n_v$. We denote by $\widehat{g}(t; x_{B_v})$ the kernel density estimate (8) of $g(t)$ from $x_{B_v}$. Expectation $\mathbb{E}(\cdot)$ is calculated with respect to density (1). We assume conditions (C2) and (C3) (cf. Pallini, 2002 as follows.

(C2). $n_v(n_v - 1) = O_p(p^2)$, $h \to 0$, $p^2 h \to \infty$, as $q \Rightarrow \infty$.

(C3). $\mathbb{E}\,(\widehat{g}(t; x_{B_v}))^{\rho+2} < \infty$ for some integer $\rho \geq 0$, where $t \geq 0$.

For every $t \geq 0$, under conditions (C1) to (C3), it follows that

$$\sum_{v \in E} \widehat{g}(t; x_{B_v}) = \widehat{g}(t) + R(t)\,,$$

where $R(t)$ is asymptotically negligible, as $q \Rightarrow \infty$.

Stationarity of the point process implies that an appropriate intensity estimate $\widehat{\lambda}_1^2$ (equivalent to $\widehat{\lambda}^2$) may be defined by the equation

$$a_1^{-2} s^2 p^2 = a^{-2}\, n(n-1)\,, \tag{10}$$

where solution $a_1$ is the area of a bounded region.

The kernel density estimate $\widehat{g}(t)$ may thus be approximated as

$$\overline{y}(t) = s^{-2} \sum_{v \in E} y_v(t)\,, \tag{11}$$

where $\overline{y}(t) = \widehat{g}(t) + O_p(q^{-1}h^{-1/2})$, as $q \Rightarrow \infty$, and the random variables

$$y_v(t) = p^{-2} \sum_{i \neq j;\; x_i\,, x_j \in B_v} \frac{a_1^2\, k_h(d(x_i, x_j) - t)}{2\pi t}\,, \tag{12}$$

are conditionally independent according to the Markov property (6).

Estimate (11) is edge-corrected (see subsection 2.2 above) for every $B_v$, taken as a restriction of $B_v \oplus r$.

Simultaneous bootstrap confidence bands for $\mathbb{E}\{\widehat{g}(t)\}$, $t \in \tau$, can be obtained from the bootstrap distribution of the suprema of

$$T^*(t) = \frac{s^{1/2} p\,(\overline{y}^*(t) - \overline{y}(t))}{(\overline{y}(t))^{1/2}}\,, \tag{13}$$

$t \geq 0$, where $\overline{y}^*(t)$ is the bootstrap version of $\overline{y}(t)$ from a generic resample $\{y_v^*(t)\}$, drawn with replacement from the original set $\{y_v(t)\}$, $v \in E$, given by (12).

Note that (13) is the bootstrap version of a well-known statistic studied in Bickel and Rosenblatt (1973), based on the mean of $s$ independent and identically distributed replicates.

Let $\xi \in (0,1)$ be the nominal level of confidence. Let $\varepsilon = (1+\xi)/2$. We want to construct a simultaneous bootstrap confidence band for $\mathbb{E}\{\widehat{g}(t)\}$, over the interval $\tau = (t_1, t_2)$. Let $\widehat{w}_L$ and $\widehat{w}_U$ be two numbers such that

$$P^*\Big\{\sup_\tau T^*(t) \leq \widehat{w}_L\Big\} = P^*\Big\{\inf_\tau T^*(t) > \widehat{w}_U\Big\} = \varepsilon\,.$$

As in Hall (1993), $\widehat{w}_L$ and $\widehat{w}_U$ must satisfy

$$P^*\Big\{\widehat{w}_L \leq \inf_\tau T^*(t) \leq \sup_\tau T^*(t) \leq \widehat{w}_U\Big\} = \xi\,.$$

The level-$\xi$ bootstrap confidence band for $\mathbb{E}\{\widehat{g}(t)\}$ over the interval $\tau = (t_1, t_2)$ is defined as

$$\mathcal{B}_\xi(\widehat{w}_L, \widehat{w}_U) = \left\{(t,z) : t_1 \leq t \leq t_2\,;\; \widehat{w}_L \leq \frac{s^{1/2} p\,(\overline{y}^*(t) - \overline{y}(t))}{(\overline{y}(t))^{1/2}} \leq \widehat{w}_U\right\}. \tag{14}$$

In particular,

$$\mathcal{B}_\xi(\widehat{w}_L, \widehat{w}_U) = \Big[(t,z) : t_1 \leq t \leq t_2\,;\; \omega_L\{\widehat{g}(t)|\widehat{w}_L\} \leq z \leq \omega_U\{\widehat{g}(t)|\widehat{w}_U\}\Big],$$

where

$$\begin{aligned}
\omega_L\{v|w\} &= v + w^2/2 - (vw^2 + w^4/4)^{1/2}\,,\\
\omega_U\{v|w\} &= v + w^2/2 + (vw^2 + w^4/4)^{1/2}\,.
\end{aligned}$$

Under conditions (C1) to (C3), bootstrap confidence bands (14) are asymptotically of exact level $\xi$, as $s \to \infty$, $q \Rightarrow \infty$.

### *3.1. Bias-corrected simultaneous bootstrap confidence bands*

Simultaneous bootstrap confidence bands based on kernel density estimates are affected by their bias. Better bootstrap confidence bands may be estimated by explicit bias-removal (Hall, 1993).

The Epanechnikov kernel function (9) is a 2nd order kernel; (9) satisfies

$$\int_{-\infty}^{\infty} k(u)\,du = 1\,,\ \int_{-\infty}^{\infty} u\,k(u)\,du = 0\,,\ \int_{-\infty}^{\infty} u^2\,k(u)\,du \neq 0\,.$$

Let

$$\kappa_2 = \frac{1}{2\,!}\int_{-\infty}^{\infty} u^2\,k(u)\,du\,.$$

Under condition (C2), the bias of the kernel density estimate $\widehat{g}(t)$ given by (8) is

$$\eta(t) = \kappa_2\,h^2\,g^{(2)}(t) + o(h^2)\,,$$

$t \geq 0$, where derivative $g^{(2)}$ is bounded and continuous. Bias $\eta(t)$ may be estimated by a kernel function $l_{h_1}(u) = h_1^{-3}\,l(h_1^{-1}u)$, with bandwidth $h_1$, having at least 2 derivatives.

In particular, bias $\eta(t)$ may be estimated by

$$\widehat{\eta}(t) = \kappa_2\,h^2\,p^{-2}s^{-2}\sum_{v\in E}\sum_{i\neq j;\,i,j=1,\ldots,n}\frac{a_1^2\,l_{h_1}^{(2)}(d(x_i,x_j)-t)}{2\pi t}\,, \qquad (15)$$

$t \geq 0$, with edge-correction for each $B_v$ as a restriction of $B_v \oplus r$.

A convenient choice for $l_{h_1}^{(2)}(u)$, $|u| \leq h_1$, in (16) is the $(2,4)$th optimal order kernel of Gasser *et al.* (1985), defined as

$$l_{h_1}^{(2)}(u) = \frac{105}{16h_1}\left(\frac{6u^2}{h_1^2} - \frac{5u^4}{h_1^4} - 1\right).$$

Selection of bandwidth $h_1$ is discussed below in section 3.2.

For every $v \in Z$, we assume condition (C4) as follows.

(C4). $n(n-1) = O_p(q^2)$, $h_1 \to 0$, $q^2 h_1^5 \to \infty$, $q h^{5/2} h_1^2 \to 0$, as $q \Rightarrow \infty$.

The bias-corrected level-$\xi$ bootstrap confidence band for the pair correlation function $\widehat{g}(t)$, over the interval $\tau = (t_1, t_2)$, is defined as

$$\mathcal{B}'_\xi(\widehat{w}_L, \widehat{w}_U) = \left\{ (t, z - \widehat{\eta}(t)) : (t, z) \in \mathcal{B}_\xi(\widehat{w}_L, \widehat{w}_U) \right\}. \qquad (16)$$

Under conditions (C1) to (C4), bootstrap confidence bands (16) are asymptotically of exact level $\xi$, as $s \to \infty$, $q \Rightarrow \infty$.

### *3.2. Bandwidths*

Simulations described in Section 4 below suggest that values for bandwidths $h$ and $h_1$, in the kernel density estimates (8) and (15), which may maximize the empirical coverage of bootstrap confidence bands (14) and (16), are

$$h = \delta_1 \, c_3 \, a_1^{1/2}, \qquad h_1 = \delta_2 \, c_3 \, a_1^{5/18}, \qquad (17)$$

where $\delta_1, \delta_2 \in (0, 1]$ and $c_3 > 0$, and area $a_1$ (of a bounded region) solves equation (10).

Analogous selection criteria for bandwidths $h$ and $h_1$ may be found in Penttinen *et al.* (1992), Stoyan *et al.* (1993) and Pallini (2002). Asymptotically optimal formulae for $h$ and $h_1$, which maximize the coverage of bootstrap confidence bands (14) and (16), as $q \Rightarrow \infty$, can be obtained, following (10) and Hall (1993).

## *4. A simulation study*

Simultaneous bootstrap confidence bands (14) and (16) typically produce interesting results in terms of nominal levels of confidence.

In this section, we present a simulation experiment performed by generating various configurations of points from Strauss processes, given by

density (3). The region $A$, containing configurations of different sizes, was always of size $30 \times 40$, with different coding schemes of size $s = 4$, $6$, $9$ and an interaction radius $r$ ranging from $0.925$ to $1.025$. We considered the kernel estimates of correlation functions as population values to be estimated with simultaneous bootstrap confidence bands (14) and (16). The approach may be justified by presuming a sufficiently large region $A$, which yields estimates very close to the true values of the point process. Empirical coverages were based on $500$ independent configurations of points, with bootstrap approximations based on $400$ resamples. In Table 1, we report results for empirical coverages of simultaneous bootstrap confidence bands (14) and (16), with $\delta_1, \delta_2 = 1$ and $c_3 = 1$ in bandwidths (17), and a nominal level of confidence $\xi = 0.9$. They confirm the effectiveness and consistency of the proposed resampling method.

Other simulation experiments show equivalent and similar results, where the worst confidence bands and situations may exhibit over-coverage of nominal confidence levels.

### *4.1. Computational details*

Strauss point processes were generated with values $r = 0.925$, $0.95$, $1.025$ for the interaction radius $r$, through density (3) (with $\theta_r(x_A) = 2$). In particular, we generated any point $x$ in a configuration $x_A$, with density proportional to ratio $f(x_A \cup \{x\})/f(x_A)$, by using Ripley's (1977) alternating birth-death technique and the acceptance-rejection method. In order to speed up simulations, we considered values $\alpha = 1$ and $\beta = 1$ for the parameters in Strauss density (3).

Labels of bootstrap resamples $\{y_v^*(t)\}$, $v \in E$, were generated from the original set of integers $\{1, 2, \ldots, s\}$, by the S-plus (cf. Becker *et al.*, 1988) command `samp.boot.mc(s,b)`, where $s = 4, 6, 9$ and $b = 400$. In order to calculate empirical coverages, we used a vector $u$ of $25$ differences between $\widehat{g}(t)$ and the corresponding lower bounds given by (14) and (16), another vector $u$ of $25$ differences between upper bounds given by (14) and (16) and $\widehat{g}(t)$, for $25$ equispaced values $t \in [0.25, 1.5]$. In particular, we simultaneously compared all differences in $u$ with the command `(mean(u > 0)==1)`, which is $1$, if all $25$ components in $u$ are

positive, $0$, otherwise.

Software is available from the Author on request.

*Table 1. Empirical coverages $\widehat{\xi}^*$ and $\widehat{\xi}^*_{BC}$ of confidence bands for the pair correlation function $g(t)$, over the interval $\tau = [0.25, 1.5]$, from a region $A$ of size $30 \times 40$ (level of confidence $\xi = 0.9$). Coding schemes based on $s$ rectangular subregions in $A$, for point configurations with interaction radius $r$ of size $n$.*

| r | s | n | $\widehat{\xi}^*$ | $\widehat{\xi}^*_{BC}$ | r | s | n | $\widehat{\xi}^*$ | $\widehat{\xi}^*_{BC}$ |
|---|---|---|---|---|---|---|---|---|---|
| 0.925 | 4 | 20 | 0.948 | 0.898 | 0.925 | 6 | 10 | 0.816 | 0.844 |
| 0.950 | 4 | 20 | 0.928 | 0.886 | 0.950 | 6 | 10 | 0.866 | 0.874 |
| 1.025 | 4 | 20 | 0.930 | 0.892 | 1.025 | 6 | 10 | 0.842 | 0.860 |
| 0.925 | 4 | 22 | 0.886 | 0.886 | 0.925 | 6 | 14 | 0.894 | 0.896 |
| 0.950 | 4 | 22 | 0.906 | 0.906 | 0.950 | 6 | 14 | 0.888 | 0.888 |
| 1.025 | 4 | 22 | 0.916 | 0.916 | 1.025 | 6 | 14 | 0.886 | 0.892 |
| 0.925 | 4 | 26 | 0.908 | 0.908 | 0.950 | 9 | 14 | 0.896 | 0.898 |
| 0.950 | 4 | 26 | 0.894 | 0.892 | 1.025 | 9 | 10 | 0.870 | 0.876 |
| 1.025 | 4 | 26 | 0.896 | 0.896 | 1.025 | 9 | 14 | 0.894 | 0.898 |

### *4.2. Interaction radius*

In all the simulations performed, the interaction radius $r$, which characterizes conditional independence in a coding scheme, is taken as known. Methods for estimating $r$ are discussed in Baddeley and Turner (2000) and Pallini (2002), along with their applications.

We may also estimate $r$, routinely, by calculating the infimum of a subset of values for $r$, which yield approximately stable simultaneous bootstrap confidence bands (14) and (16).

### *4.3. Conclusions*

Bootstrap methods for conditionally independent subregions allow straightforward applications of well-known statistical tools for independent and identically distributed observations.

By splitting the observed region $A$ into conditionally independent subregions, part of the point configuration in $A$ is not used. In any case, simulation experiments show that such bootstrap methods may provide good and efficient compromises.

Another key result is equation (10) (proposed in Pallini, 2002), where the random size of the point configuration in $A$ is conveniently defined by the number $s$ of conditionally independent subregions in $A$ and the number $p$ of rectangles (5) in each subregion.

### *References*


Baddeley, A.J. and Møller, J. (1989). Nearest-neighbour Markov point processes and random sets. *International Statistical Review, 57*, 89-121.

Baddeley, A.J. and Turner, R. (2000). Practical maximum pseudolikelihood for spatial point patterns (with discussion). *Australian and New Zealand Journal of Statist*ics, 42, 901-940.

Becker, R.A:, Chambers, J.M. and Wilks, A.R. (1988). *The New S Language*. Wadsworth and Brooks/Cole, Pacific Grove, California.

Bickel, P.J. and Rosenblatt, M. (1973). On some global measures of the deviations of density function estimates. *Annals of Statistics*, 1, 1071-1095.

Bolthausen, E. (1982). On the central limit theorem for stationary random fields. *Annals of Probability*, 10, 1047-1050.

Capobianco, R. and Renshaw, E. (1998). The autocovariance function for marked point processes: a comparison between two different approaches. *Biometrical Journal*, 40, 431-446.

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

Doguwa, S.I. and Kukoyi, K.O. (1993). Some remarks on the estimation of the pair-correlation function for point processes. *Biometrical Journal*, 35, 379-384.

Gasser, T., Müller, H.-G. and Mammitzsch, V. (1985). Kernels for nonparametric curve estimation. *Journal of the Royal Statistical Society B*, 47, 238-252.

Hall, P. (1979). The rate of convergence of normal extremes. *Journal of Applied Probability*, 16, 433-439.

Hall, P. (1993). On Edgeworth expansion and bootstrap confidence bands in nonparametric curve estimation. *Annals of Statistics*, 55, 291-304.

Jensen, J.L. (1993a). Asymptotic normality of estimates in spatial point processes. *Scandinavian Journal of Statistics*, 20, 97-109.

Kaiser, M.S., Hsu, N.-J., Cressie, N. and Lahiri, S.N. (1997). Inference for spatial processes using subsampling: a simulation study. *Environmetrics*, 8, 485-502.

Kelly, F.P. and Ripley, B.D. (1976). A note on Strauss's model for clustering. *Biometrika*, 63, 357-360.

Klein, D. (1982). Dobrushin uniqueness techniques and the decay of correlations in continuum statistical mechanics. *Communications in Mathematical Physics*, 86, 227-246.

Isham, V. (1984). Multitype Markov point processes: some approximations. *Proceedings of the Royal Society of London*, 391, 39-53.

Lahiri, S.N., Kaiser, M.S., Cressie, N. and Hsu, N.-J. (1999). Prediction of spatial cumulative functions using subsampling (with discussion). *Journal of the American Statistical Association*, 94, 86-110.

Lieshout, M.N.M. van and Baddeley, A.J. (1996). A nonparametric measure of spatial interaction in point patterns. *Statistica Neerlandica*, 50, 344-361.

Lieshout, M.N.M. van and Baddeley, A.J. (1999). Indices of dependence between types in multivariate point patterns. *Scandinavian Journal of Statistics*, 26, 511-532.

Pallini, A. (2000). Resampling configurations of points through coding schemes. *Journal of the Italian Statistical Society*, 10, 1-27.

Pallini, A. (2002). Nonparametric confidence intervals for correla-

tions in nearest-neighbour Markov point processes. *Environmetrics*, 13, 187-207.

Penttinen, A. and Stoyan, D. (2000). Recent applications of point process methods in forestry statistics. *Statistical Science, 15*, 61-78.

Penttinen, A., Stoyan, D. and Henttonen, H.M. (1992). Marked point processes in forest statistics. *Forest Science*, 38, 806-824.

Politis, D.N. and Sherman, M. (2001). Moment estimation for statistics from marked point processes. *Journal of the Royal Statistical Society B*, 63, 261-275.

Preston, C. (1976). *Random Fields*. Lectures Notes in Mathematics 534. Springer-Verlag, Berlin.

Ripley, B.D. (1988). *Statistical Inference for Spatial Processes*. Cambridge: Cambridge University Press.

Schlather, M. (2001). On the second-order characteristics of marked point processes. *Bernoulli*, 7, 99-117.

Sherman, M. and Carlstein, E. (1994). Nonparametric estimation of the moments of a general statistic computed from spatial data. *Journal of the American Statistical Association*, 89, 496-500.

Stoyan, D., Bertram, U. and Wendrock, H. (1993). Estimation variances for estimators of product densities and pair correlation functions of planar point processes. *Annals of the Institute of Statistical Mathematics*, 45, 211-221.

Stoyan, D., Kendall, W.S. and Mecke, J. (1995). *Stochastic Geometry and Its Applications*. 2nd Ed., Chichester: Wiley.

Wand, M.P. and Jones, M.C. (1995). *Kernel Smoothing*. Chapman and Hall, London.