A Contribution to the Nonparametric Statistics of Particle Systems

Consistency and asymptotic normality of a kernel-type estimator for the second-order product density of the stationary point process of exposed tangent points associated with a Boolean model are proved. This estimator is used to estimate characteristics of the typical grain. In the case of spherical grains we obtain an empirical nonparametric distribution function of the radii. This can be used to treat Wicksell’s corpuscle problem even if overlapping grains are unobservable.

1. Introduction

Suppose that a number of spheres with random radii are embedded in an opaque medium. We can observe a planar section of the medium showing overlapping circular sections of some spheres. The aim is to get the distribution function of the sphere radii from that of the circle radii. The following relation between the two distribution functions holds ([5])

\[ F(r) = 1 - \frac{2mV}{\pi} \int_0^r \frac{dF_A(t)}{\sqrt{t^2 - r^2}}, \quad r \geq 0 \]  

(1)

Here, \( F(r) \) resp. \( F_A(r) \) is the distribution function of the sphere radii resp. circle radii and \( m_V \) is the mean sphere radius. Because of the overlapping disks the circle radii cannot be observed directly. We use a tangent-point method to get a nonparametric estimator for \( F_A(r) \) and investigate its properties. First we consider a function of the distribution of the typical grain in a Boolean model in the d-dimensional Euclidean space \( \mathbb{R}^d \). For exact definitions and proofs of the formulae used in the following see [3] and [5].

The Boolean model \( Z \) is defined to be the union of i.i.d. random compact sets (grains) \( Z_0; Z_1, Z_2, \ldots \), shifted by the atoms (germs) of a stationary Poisson point process \( \Phi = \sum_{i \geq 1} \delta_X \), on \( \mathbb{R}^d \) with intensity \( \lambda \).

\[ Z = \bigcup_{i \geq 1} (Z_i + X_i) \]

(2)

\( Z_0 \) is called the typical grain. As in [3] for any \( u \in S^{d-1} \) we assign to \( Z \) the point process of exposed tangent points

\[ \Phi_u = \sum_{i \geq 1} \delta_{l_i(u, Z_i) + X_i} \prod_{j \neq i} (1 - \mathbf{1}_{Z_j + X_j} (l_i(u, Z_i) + X_i)), \]

where \( l_i(u, Z_i) \) is the tangent-point of the grain \( Z_i \) in direction \( u \in S^{d-1} \). \( \Phi_u \) turns out to be a stationary point process with intensity \( \lambda_u \). The second-order product density of \( \Phi_u \) take the following form:

\[ \rho_u^{(2)}(x) = \lambda_u \exp{\lambda |Z_0 \cap (Z_0 - x)|} \mathbb{I}(x \notin Z_0 + l(u, Z_0)) \mathbb{I}(x \notin -Z_0 + l(u, Z_0)) \]

Example. If \( Z_0 = B(0, R_0) = \{x \in \mathbb{R}^d : ||x|| \leq R \} \) then we may write

\[ F(t) = \frac{(1 - p)^2 \rho_u^{(2)}(2tu)}{(1 - 2p + C(2tu))\lambda_u} \]

(3)

in terms of the volume fraction \( p = \mathbb{P}(0 \in Z) \) and the covariance \( C(tu) = \mathbb{P}(0 \in Z, tu \in Z) \) of the Boolean model \( Z \). Here we have used the fact

\[ F(t/2) = \mathbb{P}(tu \notin -Z_0 + l(u, Z_0)) \mathbb{I}(tu \notin -Z_0 + l(u, Z_0)), \quad t > 0. \]

Using suitable unbiased estimators for \( p, \lambda_u \) and \( C(tu) \) (see e.g. [3]) and a kernel-type estimator for \( \rho_u^{(2)}(tu) \) (see [2]) we suggest the following empirical distribution function \( \hat{F}_{u,n}(t) \) for \( F(t) \):

\[ \hat{F}_{u,n}(t) = \frac{(1 - \hat{p}_n)^2 \rho_u^{(2)}(2tu)}{(1 - 2\hat{p}_n + C_n(2tu))(\hat{\lambda}_u,n)^2} \]

(4)
2. Estimators for the product density and the distribution function

According to [2] we use a kernel-type estimator to determine the product density

\[
\hat{p}_u^{(2)}(x) = \frac{1}{|W_n| b_n^d} \sum_{z_1, z_2 \in S_n} 1_{W_n}(z) k\left(\frac{x - z_1 - z_2}{b_n}\right)
\]  

(5)

Here, \( W_n = [0, n]^d \) denotes the sampling window and \( k : \mathbb{R}^d \to \mathbb{R} \) the kernel function. The sequence of bandwidths \( b_n > 0, n \in \mathbb{N} \) is chosen such that \( \lim_{n \to \infty} b_n = 0 \) and \( \lim_{n \to \infty} n b_n = \infty. \)

Lemma 1. Let \( Z \) be a Boolean model with convex grains satisfying \( E|Z_0|^2 < \infty. \) If the function 
\[
\tilde{P}(u|Z_0) := \tilde{P}(x \not\in Z_0 + l(u, Z_0)) \tilde{P}(x \not\in Z_0 + l(u, Z_0))
\]

is Lipschitz-continuous for \( x \neq 0 \) and the kernel \( k \) is a non-negative bounded function with bounded support, then for \( x \neq 0 \)

\[
\exists L > 0: \quad \left| E \hat{p}_u^{(2)}(x) - \lambda_u \hat{p}_u^{(2)}(x) \right| \leq \lambda_u L b_n \int_{\mathbb{R}^d} ||y|| k(y) dy
\]

\[
\left| W_n |b_n^d | \text{Var} \hat{p}_u^{(2)}(x) - \lambda_u \hat{p}_u^{(2)}(x) \right| \int_{\mathbb{R}^d} k^2(y) dy \leq c_1 b_n + c_2 b_n^d
\]

hold, where

\[
c_1 = \lambda_u L \int_{\mathbb{R}^d} ||y|| k^2(y) dy \quad \text{and} \quad c_2 = 4 \left( \lambda^3 + \lambda^4 (2E|Z_0| + \lambda E|Z_0|^2) \right).
\]

Moreover, for \( x \neq y \) and \( x \neq -y \) we have

\[
\left| W_n | \text{Cov} \left( \hat{p}_u^{(2)}(x), \hat{p}_u^{(2)}(y) \right) \right| \leq c_2.
\]

Applying a central limit theorem for \( m \)-dependent random variables from [1] yields

Theorem 2. If \( Z_0 \) is bounded and \( nb_n^{d+1} \to \infty \) for some \( \gamma > 0 \), and the conditions of Lemma 1 are satisfied, then

\[
\sqrt{b_n^d |W_n|} \left( \sqrt{\hat{p}_u^{(2)}(x) - \lambda_u \hat{p}_u^{(2)}(x)} \right) \to N(0, \sigma_p^2)
\]

with \( \sigma_p^2 = \int k^2(y) dy / 4 \).

Using the consistency of the estimators for \( p, \lambda_u \) and \( C(tu) \) we obtain

Corollary 3. For \( Z_0 = b(0, R_0) \) with bounded \( R_0 \) having a Lipschitz-continuous distribution function \( F(t) \) and the kernel \( k \) as in Lemma 1, we have for \( t > 0 \)

\[
\sqrt{b_n^d |W_n|} \left( \sqrt{\tilde{F}_u^{(2)}(t) - \tilde{F}(t)} \right) \to N(0, \sigma_F^2(t)) \quad \text{with} \quad \sigma_F^2(t) = \frac{(1 - p)^2 \int k^2(y) dy}{4\lambda_u^3(1 - 2p + C(2tu))}.
\]

3. References


Addresses: PROF. DR. LOTHAR HEINRICHS, Institut für Stochastik, TU Freiberg, 09596 Freiberg

DIPL.-MATH. MICHAEL WERNER, Graduiertenkolleg Räumliche Statistik, TU Freiberg, 09596 Freiberg