Asymptotic properties of estimators for the volume fractions of stationarily connected random sets

STEPHAN BÖHM
University of Ulm

LOTHAR HEINRICH
University of Augsburg

VOLKER SCHMIDT
University of Ulm

ABSTRACT. In the present paper, we show how a consistent estimator can be derived for the asymptotic covariance matrix of stationary 0-1-valued vector fields in \( \mathbb{R}^d \), whose supports are stationarily connected random closed sets. As an example, which is of particular interest for statistical applications, we consider stationarily connected random closed sets associated with the Boolean model in \( \mathbb{R}^d \) such that the components indicate the frequency of coverage by the single grains of the Boolean model. For this model, a representation formula for the entries of the covariance matrix is obtained.

Key words: Stationary vector field; covariance matrix; estimation; consistency; Boolean model; frequency of coverage

1 Preliminaries

In the following we recall some basic notions and results from stochastic geometry. More details on this can be found, for example, in Stoyan, Kendall, & Mecke (1995), Cressie (1991), Ohser & Mücklich (2000), and Molchanov (1997). Let \( \xi(x) = (\xi_1(x), \ldots, \xi_r(x))^T \), \( x \in \mathbb{R}^d \), be a stationary \( r \)-dimensional vector field in \( \mathbb{R}^d \) such that the components \( \xi_k(x) \), \( 1 \leq k \leq r \), are given by

\[
\xi_k(x) = \mathbb{1}_{\Xi_k}(x), \quad 1 \leq k \leq r,
\]

where \( \Xi_1, \ldots, \Xi_r \) are arbitrary stationary random closed sets (RACS) in \( \mathbb{R}^d \), which are stationarily connected, and \( \mathbb{1}_{\Xi_k}(x) \) denotes the indicator function of \( \Xi_k \), i.e.,

\[
\mathbb{1}_{\Xi_k}(x) = \begin{cases} 
1 & \text{if } x \in \Xi_k, \\
0 & \text{if } x \notin \Xi_k.
\end{cases}
\]

Notice that for each pair \( k, l \) with \( 1 \leq k, l \leq r \), the covariance

\[
\text{Cov}(\xi_k(x), \xi_l(y)) = \mathbb{E}\xi_k(x)\xi_l(y) - \mathbb{E}\xi_k(x)\mathbb{E}\xi_l(y), \quad x, y \in \mathbb{R}^d,
\]

(1.2)
is a function of the vector difference $h \in \mathbb{R}^d$ with $h = y - x$. We thus write

$$Cov_{kl}(h) = Cov(\xi_k(o), \xi_l(h)),$$

where $o \in \mathbb{R}^d$ denotes the origin. Furthermore, suppose that the volume fraction

$$p_k = P(o \in \Xi_k)$$

of the RACS $\Xi_k$ is (hypothetically) given for each $k = 1, \ldots, r$, where $0 < p_k < 1$. Then, we have

$$Cov_{kl}(h) = P(o \in \Xi_k, h \in \Xi_l) - p_k p_l, \quad 1 \leq k, l \leq r.$$

Notice that $Cov_{kl}(-h) = Cov_{lk}(h)$. In the following, the function

$$C_{kl}(h) = P(o \in \Xi_k, h \in \Xi_l), \quad h \in \mathbb{R}^d,$$

will be called the cross-covariance of $\Xi_k$ and $\Xi_l$. For $k = l$ we get in particular

$$C_k(h) = P(o \in \Xi_k, h \in \Xi_k),$$

where $C_k(h)$ is known in the literature as the covariance of the stationary RACS $\Xi_k$. Similar to (1.5), the centered covariance function $Cov_k(h), h \in \mathbb{R}^d$, of $\Xi_k$ is given by

$$Cov_k(h) = C_k(h) - p_k^2, \quad 1 \leq k \leq r.$$ Notice that $Cov_k(h), h \in \mathbb{R}^d$, coincides with the covariance function of the stationary random field $\{\xi_k(x), x \in \mathbb{R}^d\}$, i.e.,

$$Cov_k(h) = Cov(\xi_k(o), \xi_k(h)).$$

Using the stationarity of $\Xi_k, 1 \leq k \leq r$, we obtain for a given sampling window $W \subset \mathbb{R}^d$ and for each $h \in \mathbb{R}^d$ with $0 < |W \cap (W + h)| < \infty$ that

$$C_k(h) = \frac{\mathbb{E} |\Xi_k \cap (\Xi_k + h) \cap W \cap (W + h)|}{|W \cap (W + h)|},$$

where $|W|$ denotes the Lebesgue measure of the Borel set $W$. Accordingly, since $\Xi_k$ and $\Xi_l$ are jointly stationary RACS for each $1 \leq k, l \leq r$, we have

$$C_{kl}(h) = \frac{\mathbb{E} |\Xi_l \cap (\Xi_k + h) \cap W \cap (W + h)|}{|W \cap (W + h)|}.$$ (1.10)

## 2 Motivation from image analysis

In this section, the consistent estimation of the covariances $Cov_{kl}(h), 1 \leq k, l \leq r$, given in (1.5) will be motivated by considering an asymptotic test for the $r$-dimensional vector $(p_1, \ldots, p_r)$ of (hypothetically given) volume fractions. Suppose that an image with $r$ different phases is observed within a given sampling window $W$ with $0 < |W| < \infty$, where the phases are visualized as different tones of a grayscale image. Furthermore, we assume that the $r$ different phases can be described by realizations of
the stationary RACS $\Xi_k$, $1 \leq k \leq r$, which are observed within the sampling window $W$. Then, the empirical volume fraction $\hat{p}_{W,k}$ is given by

$$\hat{p}_{W,k} = \frac{|\Xi_k \cap W|}{|W|}, \quad 1 \leq k \leq r. \quad (2.1)$$

The covariances $\text{Cov}_{kl}(h)$, $1 \leq k, l \leq r$, are of particular interest if we want to realize an asymptotic test to check for an observed image with $r$ different grayscales if the $r$-dimensional vector $(p_1, \ldots, p_r)^T$ is reinforced by the image or not. Such an asymptotic test can be performed using a multivariate central limit theorem for the $r$-dimensional random vectors $(\hat{p}_{W,1}, \ldots, \hat{p}_{W,r})^T$, where $W_n$, $n \in \mathbb{N}$, denotes an unboundedly increasing sequence of sampling windows. Indeed, under some moment and mixing conditions on the stationary RACS $\Xi_k$, $1 \leq k \leq r$, the following central limit theorem can be proved:

$$Y_{n,r} = \begin{pmatrix} \sqrt{|W_n|} (\hat{p}_{W,1} - p_1) \\ \vdots \\ \sqrt{|W_n|} (\hat{p}_{W,r} - p_r) \end{pmatrix} \xrightarrow{d} N(0, \Sigma_r), \quad (2.2)$$

see, for example, Heinrich (2003) and Mase (1982). Related central limit theorems for random fields on $d$-dimensional lattices can be found e.g. in Guyon (1995).

The asymptotic covariance matrix $\Sigma_r$ can be determined from the covariance matrix $\Sigma_{n,r}$ of the random vector $Y_{n,r}$ by $\Sigma_r = \lim_{n \to \infty} \Sigma_{n,r}$, where $\Sigma_{n,r}$ is given by

$$\Sigma_{n,r} = |W_n| \begin{pmatrix} \text{Var} \hat{p}_{W,1} & \cdots & \text{Cov}(\hat{p}_{W,1}, \hat{p}_{W,r}) \\ \vdots & \ddots & \vdots \\ \text{Cov}(\hat{p}_{W,r}, \hat{p}_{W,1}) & \cdots & \text{Var} \hat{p}_{W,r} \end{pmatrix}.$$  

Then, assuming that $\int_{\mathbb{R}^d} |C_{kl}(h) - p_k p_l| \, dh < \infty$ for $1 \leq k, l \leq r$, the asymptotic covariance matrix $\Sigma_r$ equals

$$\Sigma_r = \begin{pmatrix} \int_{\mathbb{R}^d} (C_1(h) - p_1^2) \, dh & \cdots & \int_{\mathbb{R}^d} (C_r(h) - p_r^2) \, dh \\ \vdots & \ddots & \vdots \\ \int_{\mathbb{R}^d} (C_{r1}(h) - p_r p_1) \, dh & \cdots & \int_{\mathbb{R}^d} (C_{rr}(h) - p_r^2) \, dh \end{pmatrix},$$

since the entries of the covariance matrix $\Sigma_n$ can be written as

$$\text{Var} \hat{p}_{W,k} = \frac{1}{|W_n|^2} \int_{\mathbb{R}^d} \gamma_{W_n}(h)(C_k(h) - p_k^2) \, dh, \quad 1 \leq k \leq r, \quad (2.3)$$

and

$$\text{Cov}(\hat{p}_{W,k}, \hat{p}_{W,l}) = \frac{1}{|W_n|^2} \int_{\mathbb{R}^d} \gamma_{W_n}(h)(C_{kl}(h) - p_k p_l) \, dh, \quad 1 \leq k, l \leq r, \quad (2.4)$$
where $\gamma_{W_n}(h) = |W_n \cap (W_n - h)|$ denotes the set-covariance function of the sampling window $W_n$. Usually the covariance $C_k(h)$ of the RACS $\Xi_k$ appearing in (2.3) and the cross-covariance $C_{kl}(h)$ of $\Xi_k$ and $\Xi_l$ in (2.4) are unknown, which prevents an exact evaluation of the entries of the covariance matrix $\Sigma_n$. Therefore, we need a consistent estimation for the asymptotic covariance matrix $\Sigma$. Then, Slutsky-type arguments can be employed in order to realize an asymptotic test for the $r$-dimensional vector $(p_1, \ldots, p_r)^T$. Such consistent estimators can be obtained by considering the spectral representations of $C_k(h)$ and $C_{kl}(h)$ and evaluating a consistent kernel estimator of the corresponding spectral density from the observed image; see Böhm, Heinrich, & Schmidt (2004). On the other hand, unbiased estimators for $C_k(h)$ and $C_{kl}(h)$, which are given by

$$\hat{C}_{W_n,k}(h) = \frac{|\Xi_k \cap (\Xi_k + h) \cap W_n \cap (W_n + h)|}{|W_n \cap (W_n + h)|}, \quad 1 \leq k \leq r,$$

and

$$\hat{C}_{W_n,kl}(h) = \frac{|\Xi_l \cap (\Xi_k + h) \cap W_n \cap (W_n + h)|}{|W_n \cap (W_n + h)|}, \quad 1 \leq k, l \leq r,$$

can be used in order to get a consistent estimation for $\Sigma$.

### 3 Consistent estimation of the covariance matrix

In the following we will specify conditions on the $r$-dimensional vector field given in Section 1, such that the matrix

$$\hat{\Sigma}_{n,r} = \left( \frac{1}{|W_n|} \int_{V_n} \gamma_{W_n}(h) (\hat{C}_{W_n,kl}(h) - \hat{\rho}_{W_n,k} \hat{\rho}_{W_n,l}) dh \right)_{k,l=1}^r, \quad V_n \subset W_n$$

is an asymptotically unbiased and mean-square consistent estimator for $\Sigma_r$, i.e.,

$$\lim_{n \to \infty} \mathbb{E} \hat{\Sigma}_{n,r} = \Sigma_r \quad \text{or equivalently} \quad \lim_{n \to \infty} \left| \mathbb{E} \hat{\Sigma}_{n,r} - \Sigma_r \right| = 0 \quad (3.2)$$

and

$$\lim_{n \to \infty} \mathbb{E} \left| \hat{\Sigma}_{n,r} - \Sigma_r \right|^2 = 0 \quad (3.3)$$

where $|A| = \left( \sum_{k,l=1}^r a_{kl}^2 \right)^{1/2}$ for some matrix $A = (a_{kl})_{k,l=1}^r$.

To state our result we consider the mixed cumulant density $\Gamma_k(Z_1, \ldots, Z_k)$ of order $k$ of the random vector $(Z_1, \ldots, Z_k)^T \in \mathbb{R}^k$, which is given by

$$\Gamma_k(Z_1, \ldots, Z_k) = \frac{1}{t_k} \frac{dt_1 \ldots dt_k}{dt_k} \log \mathbb{E} e^{i(t_1 Z_1 + \cdots + t_k Z_k)} \bigg|_{t_1=\cdots=t_k=0}$$

if $\mathbb{E} |Z_i|^k < \infty$ for each $1 \leq i \leq k$. Using $\bar{Z}_i = Z_i - \mathbb{E} Z_i$ we obtain for $k = 2, 3, 4$ that

$$\Gamma_2(Z_1, Z_2) = \mathbb{E} \bar{Z}_1 \bar{Z}_2 = \text{Cov}(Z_1, Z_2),$$

$$\Gamma_3(Z_1, Z_2, Z_3) = \mathbb{E} \bar{Z}_1 \bar{Z}_2 \bar{Z}_3,$$

$$\Gamma_4(Z_1, Z_2, Z_3, Z_4) = \mathbb{E} \bar{Z}_1 \bar{Z}_2 \bar{Z}_3 \bar{Z}_4 - \mathbb{E} \bar{Z}_1 \bar{Z}_2 \mathbb{E} \bar{Z}_3 \bar{Z}_4 - \mathbb{E} \bar{Z}_1 \bar{Z}_3 \mathbb{E} \bar{Z}_2 \bar{Z}_4 - \mathbb{E} \bar{Z}_1 \bar{Z}_4 \mathbb{E} \bar{Z}_2 \bar{Z}_3.$$
Furthermore, we have
\[
\Gamma_k(Z_{\pi(1)}, \ldots, Z_{\pi(k)}) = \Gamma_k(Z_1, \ldots, Z_k)
\]
for any permutation \(\pi\) of the set \(\{1, \ldots, k\}\). We will use the following notation for any \(k, l \in \{1, \ldots, r\}\).
\[
c_{kl}^{(1,1)}(x) = \Gamma_2(\xi_k(o), \xi_l(x)) = \text{Cov}(\xi_k(o), \xi_l(x)),
\]
\[
c_{kl}^{(1,2)}(x, y) = \Gamma_3(\xi_k(o), \xi_l(x), \xi_l(y)),
\]
\[
c_{kl}^{(2,2)}(x, y, z) = \Gamma_4(\xi_k(o), \xi_k(x), \xi_l(y), \xi_l(z)).
\]

**Theorem 3.1** Let \((\Xi_1, \ldots, \Xi_r)^\top\) be a vector of stationarily connected RACS in \(\mathbb{R}^d\) and \(W_n, n \in \mathbb{N}\), an isotope sequence of convex and compact sampling windows in \(\mathbb{R}^d\) with \(\lim_{n \to \infty} \rho(W_n) = \infty\), where \(\rho(W_n) = \sup\{r \geq 0 : b(x, r) \subseteq W_n, x \in W_n\}\) and \(b(x, r)\) denotes the disk with radius \(r\) that is centered at \(x \in \mathbb{R}^d\). Furthermore, let \(V_n = b(o, \varepsilon_n \sqrt{\rho(W_n)})\) and \(\varepsilon_n \downarrow 0\) a sequence such that \(\lim_{n \to \infty} \varepsilon_n^2 \rho(W_n) = \infty\). Assume that for any \(k, l \in \{1, \ldots, r\}\) and \(n \in \{1, 2, \ldots\}\)
\[
\frac{1}{|V_n|^2} \int_{V_n} \int_{V_n} \int_{\mathbb{R}^d} \left| \mathbb{P}(o \in \Xi_l \cap (\Xi_k + x), z \in \Xi_l \cap (\Xi_k + y)) - \mathbb{P}(o \in \Xi_k \cap (\Xi_l + x)) \mathbb{P}(o \in \Xi_l \cap (\Xi_k + y)) \right| dz \, dy \, dx \leq c_0 < \infty, \quad (3.4)
\]
\[
\int_{\mathbb{R}^d} |c_{kk}^{(1,1)}(x)| \, dx < \infty, \quad \text{and} \quad \int_{\mathbb{R}^d} |c_{kl}^{(1,1)}(x)| \, dx < \infty. \quad (3.5)
\]
Then, (3.2) and (3.3) hold, i.e., \(\hat{\Sigma}_{n,r}\) is an asymptotically unbiased and mean-square consistent estimator for \(\Sigma_r\).

**Proof** The definition of \(\rho(W_n)\) and \(V_n\) yields that for each \(h \in V_n\)
\[
|W_n \cap (W_n + h)| \geq |b(o, \rho(W_n)) \cap b(h, \rho(W_n))| \\
\geq |b(o, \rho(W_n)) - |h||| \\
\geq |b(o, \rho(W_n)) - \varepsilon_n \sqrt{\rho(W_n)}|.
\]
Thus, for any \(\varepsilon > 0\) there is a \(n_0 \in \mathbb{N}\) such that for each \(n \geq n_0\) and \(h \in V_n\)
\[
\left| 1 - \frac{\gamma_{W_n}(h)}{|W_n|} \right| \leq \varepsilon.
\]
Therefore, in order to prove (3.2), we show that for each \(1 \leq k, l \leq r\)
\[
\lim_{n \to \infty} \mathbb{E} \int_{V_n} (\hat{C}_{W_n,kl}(h) - \hat{p}_{W_n,k} \hat{p}_{W_n,l}) \, dh = \int_{\mathbb{R}^d} (C_{kl}(h) - p_k p_l) \, dh. \quad (3.6)
\]
Since \(\mathbb{E} \hat{C}_{W_n,kl}(h) = C_{kl}(h)\) and \(\mathbb{E} \hat{p}_{W_n,k} = p_k\) we have
\[
\mathbb{E} \int_{V_n} (\hat{C}_{W_n,kl}(h) - \hat{p}_{W_n,k} \hat{p}_{W_n,l}) \, dh = \int_{V_n} (C_{kl}(h) - p_k p_l) \, dh - |V_n| \mathbb{E} (\hat{p}_{W_n,k} - p_k) (\hat{p}_{W_n,l} - p_l).
\]
Using
\[
\left| \int_{V_n} (C_{kl}(h) - p_k p_l) \, dh \right| \leq \int_{V_n} |c_{kl}^{(1,1)}(x)| \, dx \to 0,
\]

\[
|V_n| \mathbb{E} (\hat{p}_{W_{n,k}} - p_k)(\hat{p}_{W_{n,l}} - p_l) = \frac{|V_n|}{|W_n|^2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} 1_{W_n}(x) 1_{W_n}(y) \Gamma_2(\xi_k(x), \xi_l(y)) \, dx \, dy
\]

\[
= \frac{|V_n|}{|W_n|^2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} 1_{W_n}(x) 1_{W_n}(z + x) c_{kl}^{(1,1)}(z) \, dx \, dz
\]

\[
\leq \frac{|V_n|}{|W_n|^2} \int_{\mathbb{R}^d} |c_{kl}^{(1,1)}(z)| \, dz,
\]

and \(|W_n| \geq |b(o, \rho(W_n))|\) we obtain (3.6). Now, we prove (3.3) showing that for each \(1 \leq k, l \leq r\)

\[
\lim_{n \to \infty} \mathbb{E} \left( \int_{V_n} (\hat{C}_{W_{n,kl}}(h) - \hat{p}_{W_{n,k}} \hat{p}_{W_{n,l}}) \, dh - \int_{\mathbb{R}^d} (C_{kl}(h) - p_k p_l) \, dh \right)^2 = 0. \tag{3.7}
\]

This is satisfied whenever

\[
\lim_{n \to \infty} \mathbb{E} \left( \int_{V_n} (\hat{C}_{W_{n,kl}}(h) - C_{kl}(h)) \, dh \right)^2 = 0,
\]

since

\[
\left| \int_{V_n} (C_{kl}(h) - p_k p_l) \, dh \right| \to 0 \quad \text{and} \quad |V_n|^2 \mathbb{E} (\hat{p}_{W_{n,k}} \hat{p}_{W_{n,l}} - p_k p_l)^2 \to 0,
\]

where the latter relation is valid because of

\[
\frac{|V_n|^2}{|W_n|^2} = \frac{|b(o, 1)|^2 \varepsilon_n^2 \rho(W_n)^d}{|W_n|} \leq |b(o, 1)| \varepsilon_n^2 \to 0
\]

and

\[
\mathbb{E} (\hat{p}_{W_{n,k}} \hat{p}_{W_{n,l}} - p_k p_l)^2 = \mathbb{E} \left( (\hat{p}_{W_{n,k}} - p_k)\hat{p}_{W_{n,l}} + p_k (\hat{p}_{W_{n,l}} - p_l) \right)^2
\]

\[
\leq 2 \mathbb{E} (\hat{p}_{W_{n,k}} - p_k)^2 + 2 \mathbb{E} (\hat{p}_{W_{n,l}} - p_l)^2
\]

\[
\leq \frac{2}{|W_n|} \int_{\mathbb{R}^d} |c_{kk}^{(1,1)}(x)| \, dx + \frac{2}{|W_n|} \int_{\mathbb{R}^d} |c_{kl}^{(1,1)}(x)| \, dx.
\]

Using the definition of \(\hat{C}_{W_{n,kl}}(h)\) and \(C_{kl}(h)\) we obtain that

\[
\hat{C}_{W_{n,kl}}(h) - C_{kl}(h) = \frac{1}{|W_n \cap (W_n + h)|} \int_{W_n \cap (W_n + h)} \left( 1_{\mathbb{Z} \cap (\mathbb{Z} + h)}(u) - \mathbb{E} 1_{\mathbb{Z} \cap (\mathbb{Z} + h)}(u) \right) \, du,
\]

which results in
\[
\mathbb{E}\left( \int_{V_n} (\tilde{C}_{W_n,kl}(h) - C_{kl}(h)) \, dh \right)^2
\]

\[
= \int_{V_n} \int_{V_n} \mathbb{E} (\tilde{C}_{W_n,kl}(x) - C_{kl}(x))(\tilde{C}_{W_n,kl}(y) - C_{kl}(y)) \, dx \, dy
\]

\[
= \int_{V_n} \int_{V_n} t_{W_n \cap (W_n + x)} \cdot \frac{1}{|W_n \cap (W_n + x)|} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{1}_{W_n \cap (W_n + x)}(u) \mathbb{1}_{W_n \cap (W_n + y)}(v)

\times \text{Cov}\left( \mathbb{1}_{z \cap (z + x)}(u), \mathbb{1}_{z \cap (z + y)}(v) \right) \, du \, dv
\]

\[
= \int_{V_n} \int_{V_n} \left| \frac{|W_n \cap (W_n + y) \cap (W_n + z) \cap (W_n + x + z)|}{|W_n \cap (W_n + x)|} \right|

\times \text{Cov}\left( \mathbb{1}_{z \cap (z + x)}(o), \mathbb{1}_{z \cap (z + y)}(z) \right) \, dz \, dy .
\]

Using (3.4) we have

\[
\mathbb{E}\left( \int_{V_n} (\tilde{C}_{W_n,kl}(h) - C_{kl}(h)) \, dh \right)^2 \leq c_0 \frac{|V_n|^2}{\inf_{x \in V_n} |W_n \cap (W_n + x)|}
\]

and consequently

\[
\frac{|V_n|^2}{\inf_{x \in V_n} |W_n \cap (W_n + x)|} \leq |b(0, 1)| \left( \frac{\varepsilon_n^2}{1 - \varepsilon_n / \sqrt{\rho(W_n)}} \right)^d \xrightarrow{n \to \infty} 0 ,
\]

which completes the proof. \(\square\)

Notice that assumption (3.4) can be replaced by certain conditions on \(c_{kl}^{(2,2)}\), \(c_{kl}^{(1,2)}\), and \(c_{kl}^{(1,2)}\) if the following decomposition is considered.

\[
\text{Cov}\left( \mathbb{1}_{z \cap (z + x)} (o), \mathbb{1}_{z \cap (z + y)} (z) \right)
\]

\[
= \text{Cov}\left( \left( \mathbb{1}_{z} (o) - p_k \right) \left( \mathbb{1}_{z + x} (o) - p_k \right), \left( \mathbb{1}_{z} (z) - p_k \right) \left( \mathbb{1}_{z + y} (z) - p_k \right) \right)
\]

\[
+ \mathbb{E} \left( p_k \left( \mathbb{1}_{z} (o) - p_k \right) + p_k \left( \mathbb{1}_{z + x} (o) - p_k \right) \right) \left( \mathbb{1}_{z} (z) - p_k \right) \left( \mathbb{1}_{z + y} (z) - p_k \right)
\]

\[
+ \mathbb{E} \left( \left( \mathbb{1}_{z} (o) - p_k \right) \left( \mathbb{1}_{z + x} (o) - p_k \right) \right) \left( p_k \left( \mathbb{1}_{z} (z) - p_k \right) + p_k \left( \mathbb{1}_{z + y} (z) - p_k \right) \right)
\]

\[
+ \mathbb{E} \left( p_k \left( \mathbb{1}_{z} (o) - p_k \right) + p_k \left( \mathbb{1}_{z + x} (o) - p_k \right) \right) \left( p_k \left( \mathbb{1}_{z} (z) - p_k \right) + p_k \left( \mathbb{1}_{z + y} (z) - p_k \right) \right).
\]

Using the stationarity of the random field \((\xi_k(x), \xi_l(x))^\top, x \in \mathbb{R}^d\), we get

\[
\text{Cov}\left( \mathbb{1}_{z \cap (z + x)} (o), \mathbb{1}_{z \cap (z + y)} (z) \right)
\]

\[
= c_{kl}^{(2,2)} (z + x - y, z + x, x) + c_{kk}^{(1,1)} (z + x - y, z + x, x) + c_{kl}^{(1,1)} (y - z, y - z, y)
\]

\[
+ p_k c_{kl}^{(1,2)} (y, y - z) + p_k c_{kl}^{(1,2)} (x, y + z) + p_k c_{kl}^{(1,2)} (-y, -z, x) + p_k c_{kl}^{(1,2)} (-x, z - y)
\]

\[
+ p_k c_{kl}^{(1,1)} (z) + p_k^2 c_{kl}^{(1,1)} (z + x - y) + p_k p_l c_{kl}^{(1,1)} (z - y) + p_k p_l c_{kl}^{(1,1)} (z + x).
\]

This immediately leads to the following result.
Lemma 3.1 Assuming that $\int_{\mathbb{R}^d} |c_{kk}^{(1)}(x)| \, dx < \infty$ and $\int_{\mathbb{R}^d} |c_{kl}^{(1)}(x)| \, dx < \infty$ for any $k, l \in \{1, \ldots, r\}$, condition (3.4) is satisfied whenever

$$
\frac{1}{|V_n|^2} \int_{V_n} \int_{V_n} \int_{\mathbb{R}^d} |c_{kl}^{(2)}(z, z + y, x)| \, dz \, dy \, dx \leq c_1 < \infty, \\
\frac{1}{|V_n|} \int_{V_n} \int_{\mathbb{R}^d} |c_{kl}^{(1)}(z, x)| \, dz \, dx \leq c_2 < \infty.
$$

(3.8)

Furthermore, notice that condition (3.8) is satisfied whenever

$$
\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |c_{kl}^{(2)}(z, y, x)| \, dz \, dy \, dx < \infty, \\
\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |c_{kl}^{(1)}(z, x)| \, dz \, dx < \infty,
$$

(3.9)

or

$$
\sup_{x, y \in \mathbb{R}^d} \int_{\mathbb{R}^d} |c_{kl}^{(2)}(z, y + z, x)| \, dz < \infty, \\
\sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} |c_{kl}^{(1)}(z, x)| \, dz < \infty.
$$

(3.10)

4 Evaluation of cross-covariances

In the following we consider stationary RACS $\Xi_1, \Xi_2, \ldots$, which are deduced from the Boolean model in $\mathbb{R}^d$ such that the different sets indicate the frequency of coverage by the single grains of the underlying Boolean model. Notice that these RACS are not Boolean models anymore. However, for these RACS representation formulas of the quantities $p_k, C_k(h)$, and $C_{kl}(h)$ can be given. For certain grain distributions even explicit formulae can be obtained for these quantities. Therefore, the definition of the Boolean model will be briefly recalled in the following. Let

$$\Psi = \sum_{n \geq 1} \delta_{[X_n, M_n]} \sim P_{\lambda, Q}$$

be an independently marked stationary Poisson process in $\mathbb{R}^d$ with finite and positive intensity $\lambda$ of the points $\{X_n, n \geq 1\}$ and with marks $\{M_n, n \geq 1\}$ being a sequence of independent copies of a non-empty and compact RACS $M_0$ in $\mathbb{R}^d$ (called typical grain), where $Q$ denotes the distribution of the marks $\{M_n, n \geq 1\}$ and $P_{\lambda, Q}$ the distribution of the marked Poisson process $\Psi$. Furthermore, we assume that

$$\mathbb{E} |M_0| = \int_{\mathcal{K}} |K| \, Q(dK) < \infty, \tag{4.1}$$

where $\mathcal{K}$ denotes the space of all compact sets in $\mathbb{R}^d$. Then, the union

$$\Xi = \bigcup_{n \geq 1} (M_n + X_n) \tag{4.2}$$
is called a stationary Boolean (germ-grain) model in $\mathbb{R}^d$, where the $X_n$ are called germs, and the $M_n$ grains. Notice that in general condition (4.1) does not ensure the property of $\Xi$ being a closed set. This is the case if and only if $\mathbb{E}|M_0 \oplus b(o, \varepsilon)| < \infty$ for any $\varepsilon > 0$. Now we consider the random field

$$
\Phi(x) = \sum_{n \geq 1} \mathbf{1}_{(M_n + X_n)}(x), \quad x \in \mathbb{R}^d,
$$

and the $r$-dimensional random vector $\xi(x) = (\xi_1(x), \ldots, \xi_r(x))^\top, x \in \mathbb{R}^d$, with components $\xi_k(x) = \mathbf{1}_{(\Phi(x) \geq k)}, 1 \leq k \leq r$, which is a stationary $r$-dimensional vector field in $\mathbb{R}^d$. Then, for each $1 \leq k \leq r$,

$$
\Xi_k = \{x \in \Xi : \xi_k(x) = 1\}
$$

is a stationary RACS in $\mathbb{R}^d$, where $\Xi_k$ contains those areas of the Boolean model $\Xi$, which are covered by at least $k$ of the shifted grains $M_n + X_n, n \geq 1$. Conversely, according to (1.1), we have the relation

$$
\xi_k(x) = \mathbf{1}_{\xi_k}(x), \quad 1 \leq k \leq r.
$$

Notice that $\Xi_1 \supseteq \Xi_2 \supseteq \ldots \supseteq \Xi_r$, where $\Xi_1 = \Xi$. This decomposition of the Boolean model $\Xi$ into the sequence of stationary RACS $\Xi_1, \Xi_2, \ldots, \Xi_r$ will be called a multiphase model deduced from the Boolean model $\Xi$ with $r$ different phases $\Xi_k, 1 \leq k \leq r$.

The decomposition of the Boolean model $\Xi$ into the infinite sequence of stationary RACS $\Xi_1, \Xi_2, \ldots$ will just be called a multiphase model deduced from the Boolean model $\Xi$. In addition, for this model let us denote by the stationary RACS

$$
\Xi^{(i)} = \Xi_1 \setminus \text{int}(\Xi_{i+1}), \quad i \geq 1,
$$

those areas of the Boolean model $\Xi$, which are covered by exactly $i$ of the shifted grains $M_n + X_n, n \geq 1$, where $\text{int}(\Xi_{i+1})$ denotes the set of all interior points of $\Xi_{i+1}$.

In Figure 1, the realization of a multiphase model deduced from the Boolean model $\Xi$ in $\mathbb{R}^2$ with two different phases $\Xi_1$ and $\Xi_2$ is given within a sampling window of $800 \times 800$ pixel points. The underlying Boolean model $\Xi$ has intensity $\lambda = 8.8254 \cdot 10^{-5}$ and volume fraction $p = 0.5$, where the typical grain is a disc with uniformly distributed radius between 40 and 60 pixel points. Here, the multiphase model is visualized by a grayscale image, where the gray phase represents the RACS $\Xi^{(1)}$ and the black phase the RACS $\Xi_2$. This means that the gray phase indicates those areas of the Boolean model $\Xi$, which are covered by exactly one of the shifted grains $M_n + X_n, n \geq 1$, and the black phase those areas, which are covered by at least two of the shifted grains.

For the above defined multiphase model deduced from the Boolean model $\Xi$, the quantities $p_k$ and $C_{kl}(h)$ introduced in (1.4) and (1.6), respectively, can be determined as follows. Since $\Xi^{(1)}, \Xi^{(2)}, \ldots$ forms a sequence of pairwise disjoint stationary RACS in $\mathbb{R}^d$, we have

$$
p_k = 1 - \sum_{i = 0}^{k-1} \mathbb{P}(o \in \Xi^{(i)}),
$$

and

$$
C_{kl}(h) = \sum_{i \geq k} \sum_{j \geq l} \mathbb{P}(o \in \Xi^{(i)}, h \in \Xi^{(j)}),
$$

for any $1 \leq k \leq r, 1 \leq l \leq r$. In Figure 1, the realization of a multiphase model deduced from the Boolean model $\Xi$ in $\mathbb{R}^2$ with two different phases $\Xi_1$ and $\Xi_2$ is given within a sampling window of $800 \times 800$ pixel points. The underlying Boolean model $\Xi$ has intensity $\lambda = 8.8254 \cdot 10^{-5}$ and volume fraction $p = 0.5$, where the typical grain is a disc with uniformly distributed radius between 40 and 60 pixel points. Here, the multiphase model is visualized by a grayscale image, where the gray phase represents the RACS $\Xi^{(1)}$ and the black phase the RACS $\Xi_2$. This means that the gray phase indicates those areas of the Boolean model $\Xi$, which are covered by exactly one of the shifted grains $M_n + X_n, n \geq 1$, and the black phase those areas, which are covered by at least two of the shifted grains.
where $\Xi^{(0)} = \mathbb{R}^d \setminus \text{int}(\Xi)$. Since the probability $\mathbb{P}(o \in \Xi^{(i)})$ can be deduced from $\mathbb{P}(o \in \Xi^{(i)}, h \in \Xi^{(j)})$, only the evaluation of $\mathbb{P}(o \in \Xi^{(i)}, h \in \Xi^{(j)})$ is required, which is the probability that the origin $o \in \mathbb{R}^d$ is covered by exactly $i$ of the shifted grains $M_n + X_n$, $n \geq 1$, and some point $h \in \mathbb{R}^d$ is covered by exactly $j$ of the shifted grains.

**Theorem 4.1** Consider the Boolean model $\Xi$ in $\mathbb{R}^d$ with intensity $\lambda$ and typical grain $M_0$. Then, for each $h \in \mathbb{R}^d$ and $1 \leq i \leq j$,

$$\mathbb{P}(o \in \Xi^{(i)}, h \in \Xi^{(j)}) = \sum_{0 \leq k \leq i} \binom{j}{i-k} \frac{1}{k!j!} \lambda^{k+j} \left( \mathbb{E} |M_0 \cap (M_0 + h)| \right)^{i-k} \left( \mathbb{E} |M_0| - \mathbb{E} |M_0 \cap (M_0 + h)| \right)^{j-i+2k} \exp \left\{ -\lambda \left( 2 \mathbb{E} |M_0| - \mathbb{E} |M_0 \cap (M_0 + h)| \right) \right\}.$$  

**Proof** Let $m_1, \ldots, m_k, n_1, \ldots, n_j \geq 1$ and $q_1, \ldots, q_{i-k} \geq 1$ be pairwise different indices with $\{q_1, \ldots, q_{i-k}\} \subset \{n_1, \ldots, n_j\}$, where $i, j \geq 1$ with $i \leq j$ and $0 \leq k \leq i$. Then,

$$A_{q_1, \ldots, q_{i-k}; m_1, \ldots, m_k; n_1, \ldots, n_j} = \left\{ o \in M_{q_1} + X_{q_1}, \ldots, o \in M_{q_{i-k}} + X_{q_{i-k}}, o \in M_{m_1} + X_{m_1}, \ldots, o \in M_{m_k} + X_{m_k}, o \notin \bigcup_{l \neq q_1, \ldots, q_{i-k}, m_1, \ldots, m_k} (M_l + X_l), \right.$$ 

$$h \in M_{n_1} + X_{n_1}, \ldots, h \in M_{n_j} + X_{n_j}, h \notin \bigcup_{l \neq n_1, \ldots, n_j} (M_l + X_l) \right\}$$
denotes the event that the origin \( o \in \mathbb{R}^d \) is covered by the shifted grains \( M_{q_1} + X_{q_1}, \ldots, M_{q_{i-1}} + X_{q_{i-1}}, M_{q_i} + X_{q_i}, \ldots, M_{q_n} + X_{q_n} \), but not touched by any other grain \( M_i + X_i \), and that the vector \( h \in \mathbb{R}^d \) is covered by the grains \( M_{n_1} + X_{n_1}, \ldots, h \in M_{n_j} + X_{n_j} \), but not touched by any other grain \( M_i + X_i \), where \( M_{q_1} + X_{q_1}, \ldots, M_{q_{i-1}} + X_{q_{i-1}} \) represent those grains, which cover both the origin \( o \) and \( h \). Thus, the quantity \( \mathbb{P}(o \in \Xi^{(i)}, h \in \Xi^{(j)}) \) equals the probability of the event

\[
B_{i,j} = \bigcup_{0 \leq k \leq i} \bigcup_{\substack{(q_1, \ldots, q_{i-1}) \in \{n_1, \ldots, n_j\}^i \ \forall \ m \leq m_k \leq m_{i-1} \ \forall \ n \leq n_j}} A_{q_1, \ldots, q_{i-1} ; m_1, \ldots, m_k ; n_1, \ldots, n_j},
\]

where the following two properties (i) and (ii) are valid.

(i) \( A_{q_1, \ldots, q_{i-1} ; m_1, \ldots, m_k ; n_1, \ldots, n_j} = A_{q_{n(1)}, \ldots, q_{n(i-k)} ; m_{n(1)}, \ldots, m_{n(k)} ; n_{n(1)}, \ldots, n_{n(j)}} \),

(ii) \( A_{q_1, \ldots, q_{i-1} ; m_1, \ldots, m_k ; n_1, \ldots, n_j} \cap A_{p_1, \ldots, p_{i-k} ; m'_1, \ldots, m'_k ; n'_1, \ldots, n'_j} = \emptyset \),
if \( \{q_1, \ldots, q_{i-1}\} \times \{m_1, \ldots, m_k\} \times \{n_1, \ldots, n_j\} \neq \{p_1, \ldots, p_{i-k}\} \times \{m'_1, \ldots, m'_k\} \times \{n'_1, \ldots, n'_j\} \).

Using properties (i) and (ii) it follows that

\[
\mathbb{P}(o \in \Xi^{(i)}, h \in \Xi^{(j)}) = \mathbb{P}(B_{i,j}) = \mathbb{E} 1(B_{i,j})
\]

\[
= \mathbb{E} \sum_{0 \leq k \leq i} \sum_{\substack{(q_1, \ldots, q_{i-1}) \in \{n_1, \ldots, n_j\}^i \ \forall \ m \leq m_k \leq m_{i-1} \ \forall \ n \leq n_j}} \frac{1}{k! j!} \sum_{m_1, \ldots, m_k \geq 1} \sum_{m_1, \ldots, n_j \geq 1} ^* 1(A_{q_1, \ldots, q_{i-1} ; m_1, \ldots, m_k ; n_1, \ldots, n_j}),
\]

where the symbol \( \sum^* \) means the summation over pairwise different indices. Furthermore, we have

\[
1(A_{q_1, \ldots, q_{i-1} ; m_1, \ldots, m_k ; n_1, \ldots, n_j}) = 1_{\tilde{K}}(X_{q_1}) \cdot \ldots \cdot 1_{\tilde{K}}(X_{q_{i-1}}) \cdot 1_{\tilde{M}_{q_1}}(X_{m_1}) \cdot \ldots \cdot 1_{\tilde{M}_{q_k}}(X_{m_k})
\]

\[
\prod_{l \neq q_1, \ldots, q_{i-1}, m_1, \ldots, m_k} \left( 1 - 1_{\tilde{M}_{l}}(X_{l}) \right) \cdot 1_{\tilde{M}_{n_1} + h}(X_{n_1}) \cdot \ldots \cdot 1_{\tilde{M}_{n_j} + h}(X_{n_j})
\]

where \( \tilde{K} = \{-x : x \in K\} \) denotes the reflection at the origin \( o \in \mathbb{R}^d \). Thus, we obtain that

\[
\mathbb{E} \sum_{m_1, \ldots, m_k \geq 1 \ \forall \ n_1, \ldots, n_j \geq 1} ^* 1(A_{q_1, \ldots, q_{i-1} ; m_1, \ldots, m_k ; n_1, \ldots, n_j})
\]

\[
= \int_{N_{\tilde{K}}} \sum_{[x_{m_1}, K_{m_1}], \ldots, [x_{m_k}, K_{m_k}], [x_{n_1}, K_{n_1}], \ldots, [x_{n_j}, K_{n_j}] \in S_{\tilde{K}}} ^* 1_{K_{q_1}}(x_{q_1}) \cdot \ldots \cdot 1_{K_{q_{i-1}}}(x_{q_{i-1}})
\]
\[ \prod_{[y,L] \in \mathbb{S}_\psi \setminus \{[x_1, K_1], \ldots, [x_{m_k}, K_{m_k}], [x_{n_1}, K_{n_1}], \ldots, [x_{n_j}, K_{n_j}]\}} \left( 1 - \mathbf{1}_L(y) \right) \left( 1 - \mathbf{1}_{L+h}(y) \right) \prod_{[y,L] \in \{[x_{n_1}, K_{n_1}], \ldots, [x_{n_j}, K_{n_j}]\} \setminus \{[x_1, K_1], \ldots, [x_{m_k}, K_{m_k}]\}} \left( 1 - \mathbf{1}_L(y) \right) \mathbf{1}_{L+h}(y), \]

where \( N_\psi \) denotes the space of the marked counting measure \( \psi = \sum_{n \geq 1} \delta_{[x_n, K_n]} \) and \( S_\psi \) the support of the marked counting measure \( \psi \) with

\[ S_\psi = \{[x, K] : [x, K] \in \mathbb{R}^d \times \mathbb{K}, \psi([\{x, K\}]) > 0\}. \]

Using the refined Campbell theorem with the reduced \((k+j)\) - fold Palm distribution \( P_{\lambda, Q}^1 : [x_1, K_1], \ldots, [x_{m_k}, K_{m_k}], [x_{n_1}, K_{n_1}], \ldots, [x_{n_j}, K_{n_j}] \) of the marked Poisson process \( \Psi \) (see, for example, Chapter 12 of Daley & Vere-Jones (1988)), we have

\[ \mathbb{E} \sum_{m_1, \ldots, m_k \geq 1} \sum_{n_1, \ldots, n_j \geq 1} \mathbf{1}(A_{q_1, \ldots, q_{k+j} : m_1, \ldots, m_k : n_1, \ldots, n_j}) = \int_{(\mathbb{R}^d \times \mathbb{K})^{k+j}} \prod_{[y,L] \in \{[x_{q_1}, K_{q_1}], \ldots, [x_{q_{k+j}}, K_{q_{k+j}}]\}} \mathbf{1}_L(y) \cdot \mathbf{1}_{L+h}(y) \prod_{[y,L] \in \{[x_{m_1}, K_{m_1}], \ldots, [x_{m_k}, K_{m_k}]\}} \left( 1 - \mathbf{1}_L(y) \right) \cdot \mathbf{1}_{L+h}(y) \prod_{[y,L] \in \{[x_{n_1}, K_{n_1}], \ldots, [x_{n_j}, K_{n_j}]\} \setminus \{[x_{m_1}, K_{m_1}], \ldots, [x_{m_k}, K_{m_k}]\}} \left( 1 - \mathbf{1}_L(y) \right) \cdot \mathbf{1}_{L+h}(y) \int_{N_\psi} \prod_{[y,L] \in \mathbb{S}_\psi} \left( 1 - \mathbf{1}_L(y) - \mathbf{1}_{L+h}(y) + \mathbf{1}_{L \cap (L+h)}(y) \right) P_{\lambda, Q}^1([x_{m_1}, K_{m_1}], \ldots, [x_{m_k}, K_{m_k}], [x_{n_1}, K_{n_1}], \ldots, [x_{n_j}, K_{n_j}]) (d\psi) \]

\[ \alpha_{k+j}^1 \left( d([x_1, K_1], \ldots, [x_{m_k}, K_{m_k}], [x_{n_1}, K_{n_1}], \ldots, [x_{n_j}, K_{n_j}]) \right), \]

where \( \alpha_{k+j}^1 \) is the \((k+j)\)th order factorial moment measure of \( \Psi \). Since \( \Psi \) is an independently marked stationary Poisson process in \( \mathbb{R}^d \) with intensity \( \lambda \), the factorial moment measure \( \alpha_{k+j}^1 \) can be written as

\[ \alpha_{k+j}^1 \left( d([x_1, K_1], \ldots, [x_{m_k}, K_{m_k}], [x_{n_1}, K_{n_1}], \ldots, [x_{n_j}, K_{n_j}]) \right) = \lambda^{k+j} \cdot dx_{m_1} \cdot dx_{m_k} \cdot dx_{n_1} \cdots dx_{n_j} \cdot Q(dK_{m_1}) \cdots Q(dK_{m_k}) \cdot Q(dK_{n_1}) \cdots Q(dK_{n_j}). \]
Furthermore, Slivnyak’s theorem (see e.g. Daley & Vere-Jones (1988)) yields
\[
P_{\lambda, Q; [x_{m_1}, K_{m_1}]; \ldots; [x_{m_k}, K_{m_k}]; [x_{n_1}, K_{n_1}]; \ldots; [x_{n_j}, K_{n_j}]} = P_{\lambda, Q}, \tag{4.11}
\]
and for any measurable function \(\nu([y, L])\) on \(\mathbb{R}^d \times \mathbb{K}\) with \(0 \leq \nu([y, L]) \leq 1\) for all \([y, L] \in \mathbb{R}^d \times \mathbb{K}\) the generating functional of \(\Psi\) becomes
\[
\int_{\mathbb{R}^d} \prod_{[y, L] \in S_\psi} \nu([y, L]) P_{\lambda, Q}(d\psi) = \exp \left\{ -\lambda \int_{\mathbb{R}^d \times \mathbb{K}} \left( 1 - \nu([y, L]) \right) dy Q(dL) \right\}. \tag{4.12}
\]
Using (4.10), (4.11), and (4.12) we obtain that
\[
\mathbb{E} \sum_{m_1, \ldots, m_k \geq 1}^{m_{k+1}} \sum_{n_1, \ldots, n_j \geq 1}^{n_{k+1}} \mathbb{1}(A_{q_1, \ldots, q_{k+1}; m_1, \ldots, m_k; n_1, \ldots, n_j})
= \lambda^{k+j} \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} \prod_{[y, L] \in \{ [x_{q_1}, K_{q_1}]; \ldots; [x_{q_{k+1}}, K_{q_{k+1}}] \}} \mathbb{1}_L(y) \cdot \mathbb{1}_{L+h}(y)
\prod_{[y, L] \in \{ [x_{n_1}, K_{n_1}]; \ldots; [x_{n_j}, K_{n_j}] \} \setminus \{ [x_{q_1}, K_{q_1}]; \ldots; [x_{q_{k+1}}, K_{q_{k+1}}] \}} \mathbb{1}_L(y) \cdot \left( 1 - \mathbb{1}_{L+h}(y) \right)
\exp \left\{ -\lambda \int_{\mathbb{R}^d \times \mathbb{K}} \left( \mathbb{1}_L(y) + \mathbb{1}_{L+h}(y) - \mathbb{1}_{L\cap(L+h)}(y) \right) dy Q(dL) \right\}
\left( \int_{\mathbb{K}} |L\cap(L+h)| Q(dL) \right)^{j+2k}\tag{4.9}
= \lambda^{k+j} \left( \int_{\mathbb{R}} |L\cap(L+h)| Q(dL) \right)^{j+2k}
\exp \left\{ -\lambda \int_{\mathbb{K}} \left( 2 |L| - |L\cap(L+h)| \right) Q(dL) \right\}
\left( \mathbb{E} |M_0 \cap (M_0 + h)| \right)^{j+2k}\tag{4.9}
= \lambda^{k+j} \left( \mathbb{E} |M_0 \cap (M_0 + h)| \right)^{j+2k}\tag{4.9}
\exp \left\{ -\lambda \left( 2 \mathbb{E} |M_0| - \mathbb{E} |M_0 \cap (M_0 + h)| \right) \right\}.\tag{4.9}
\]
This result combined with relation (4.9) completes the proof. \(\square\)

Notice that for \(1 \leq j \leq i\), a similar representation formula for \(\mathbb{P}(o \in \Xi^{(i)}, h \in \Xi^{(j)})\) can be determined using Theorem 4.1 and the fact that \(\mathbb{P}(o \in \Xi^{(i)}, h \in \Xi^{(j)}) = \mathbb{P}(o \in \Xi^{(j)}, -h \in \Xi^{(i)})\). Furthermore, as an immediate consequence of Theorem 4.1, we obtain the well-known result for the probability \(\mathbb{P}(o \in \Xi^{(i)})\), which is the volume fraction of that subset of the Boolean model \(\Xi\) that is covered by exactly \(i\) of the shifted grains \(M_n + X_n, n \geq 1\); see, for example, Chapter 4 in Hall (1988).
Corollary 4.2 Consider the Boolean model $\Xi$ in $\mathbb{R}^d$ with intensity $\lambda$ and typical grain $M_0$. Then, for each $i \geq 1$,

$$\mathbb{P}(o \in \Xi^{(i)}) = \frac{\left(\lambda \mathbb{E}[M_0]\right)^i}{i!} \exp \left\{ -\lambda \mathbb{E}[M_0] \right\}.$$ 

Remarks

1. For a multiphase model deduced from the Boolean model $\Xi$ with $r$ different phases $\Xi_i$, $1 \leq i \leq r$, we want to check conditions (3.4) and (3.5) of Theorem 3.1. In case of condition (3.4) (or alternatively in case of (3.8), (3.9) or (3.10)) this is not obvious since in the first place we have to find representation formulas for $\mathbb{P}(o \in \Xi_i \cap (\Xi_k + x), z \in \Xi_i \cap (\Xi_k + y))$ or the mixed cumulant densities $c_{kl}^{(1,2)}(x, y)$ and $c_{kl}^{(2,2)}(x, y, z)$ with $k, l \in \{1, \ldots, r\}$.

2. On the other hand, using Theorem 4.1 we get a representation formula for $c_{kl}^{(1,1)}(x)$, which shows that $\int_{\mathbb{R}^d} |c_{kl}^{(1,1)}(x)| \, dx < \infty$ holds whenever $\mathbb{E}[M_0]^2 < \infty$, where the fact is used that

$$\int_{\mathbb{R}^d} \left( e^{-\lambda \mathbb{E}[M_0]} - e^{-2\lambda \mathbb{E}[M_0]} \right) \, dx \leq \int_{\mathbb{R}^d} \left( 1 - e^{-\lambda \mathbb{E}[M_0 \cap (M_0 + x)]} \right) \, dx$$

$$\leq \lambda \int_{\mathbb{R}^d} \mathbb{E}[M_0 \cap (M_0 + x)] \, dx = \lambda \mathbb{E}[M_0]^2.$$

3. Furthermore, the considerations made in Heinrich (2003) show that for the Boolean model $\Xi$ the following inequality holds uniformly in $x, y \in \mathbb{R}^d$.

$$\int_{\mathbb{R}^d} |\mathbb{P}(o \in \Xi \cap (\Xi + x), z \in \Xi \cap (\Xi + y)) - \mathbb{P}(o \in \Xi \cap (\Xi + x)) \mathbb{P}(o \in \Xi \cap (\Xi + y))| \, dz \leq c_1(\lambda) + c_2(\lambda) \mathbb{E}[M_0]^2.$$ 

In addition, it can be shown that the assumption $\mathbb{E}[M_0]^2 < \infty$ implies that

$$\sup_{x, y \in \mathbb{R}^d} \int_{\mathbb{R}^d} |\Gamma_4(\xi(0), \xi(x), \xi(y + z), \xi(z))| \, dz < \infty,$$

$$\sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} |\Gamma_3(\xi(0), \xi(x), \xi(z))| \, dz < \infty,$$

where $\xi(x) = \mathbb{1}_{\Xi}(x)$. This is a consequence of Lemma 7 in Heinrich (2003), where it is proved that

$$\int_{\mathbb{R}^{(k-1)d}} |\Gamma_k(\xi(0), \xi(x_1), \ldots, \xi(x_{k-1}))| \, dx_1 \ldots dx_{k-1} < \infty$$

follows from $\mathbb{E}[M_0]^k < \infty$. The proof of Theorem 4.1 yields that in the representation formulas of $c_{kl}^{(2,2)}(x, y, z)$ or $\mathbb{P}(o \in \Xi_i \cap (\Xi_k + x), z \in \Xi_i \cap (\Xi_k + y))$ the appearing terms have the form $\exp\{-\lambda \mathbb{E}[M_0 \cup (M_0 + x) \cup (M_0 + y) \cup (M_0 + z)]\}$. Therefore, repeating the considerations made in Heinrich (2003) it can be shown that conditions (3.4) and (3.10) are valid if again $\mathbb{E}[M_0]^2 < \infty$. 

14
4. Moreover, it can be shown that $\mathbb{E} |M_0|^2 < \infty$ also implies the validity of the central limit theorem given in (2.2), where a central limit theorem for $m$-dependent random fields (see Heinrich (1988)) and the Cramér-Wold device is applied.

5. The idea of an alternative proof of Theorem 4.1 can be sketched as follows. Instead of the independently marked stationary Poisson process $\Psi = \sum_{n \geq 1} \delta_{[X_n, M_n]}$, we consider now the process

$$
\Psi_A = \sum_{n \geq 1} \delta_{[X_n, A_n]},
$$

which is a (stochastically independent) position-dependent marking of the Poisson points $\{X_n, n \geq 1\}$ in $\mathbb{R}^d$, where the marks $A_n \in A = \{0, 1, 2, 3\}$ are determined by

- $A_n = 0$ if $o \in M_n + X_n$ and $h \in M_n + X_n$,
- $A_n = 1$ if $o \in M_n + X_n$ and $h \notin M_n + X_n$,
- $A_n = 2$ if $h \in M_n + X_n$ and $o \notin M_n + X_n$,
- $A_n = 3$ if $o \notin M_n + X_n$ and $h \notin M_n + X_n$.

It can be shown that $\Psi_A$ is a non-stationary Poisson process in the product space $\mathbb{R}^d \times A$. For an unformal proof of this fact, see Proposition 4.10.1 in Resnick (1992), for example. This implies in particular that, for each $i \in A$,

$$
\Psi_i = \sum_{n \geq 1: A_n = i} \delta_{X_n}
$$

is a non-stationary Poisson process in $\mathbb{R}^d$ with finite intensity measure for $i = 0, 1, 2$. Furthermore, the processes $\Psi_i$ and $\Psi_j$ are independent if $i \neq j$. In the following, the (finite) random variables $Z_{o,h} = \Psi_0(\mathbb{R}^d)$, $Z_{o-h} = \Psi_1(\mathbb{R}^d)$, and $Z_{h-o} = \Psi_2(\mathbb{R}^d)$ denote the numbers of shifted grains $M_n + X_n$, $n \geq 1$, which cover $o$ and $h$ at the same time, $o$ but not $h$, and $h$ but not $o$, respectively. Then, the following decomposition of the probability $\mathbb{P}(o \in \Xi^{(i)}, h \in \Xi^{(j)})$ can be obtained. Namely,

$$
\mathbb{P}(o \in \Xi^{(i)}, h \in \Xi^{(j)}) = \sum_{0 \leq k \leq i} \mathbb{P}(Z_{o-h} = k, Z_{o,h} = i - k, Z_{h-o} = j - (i - k))
$$

$$
= \sum_{0 \leq k \leq i} \mathbb{P}(Z_{o-h} = k) \mathbb{P}(Z_{o,h} = i - k) \mathbb{P}(Z_{h-o} = j - (i - k)),
$$

where the last equation follows from the independence of the processes $\Psi_0$, $\Psi_1$, and $\Psi_2$. Associated with the Poisson processes $\Psi_i, i \in A$, the random variables $Z_{o-h}$ and $Z_{h-o}$ have each a Poisson distribution with parameter $\lambda(\mathbb{E} |M_0| - \mathbb{E} |M_0 \cap (M_0 + h)|)$ and $Z_{o,h}$ has a Poisson distribution with parameter $\lambda \mathbb{E} |M_0 \cap (M_0 + h)|$, where, for example, in the latter case the parameter $\lambda \mathbb{E} |M_0 \cap (M_0 + h)|$ represents the expected number of shifted grains that cover $o$ and $h$ at the same time, i.e.,

$$
\mathbb{E} \sum_{[y,L] \in S_L} \mathbb{I}_{L+y(o)} \mathbb{I}_{L+y(h)} = \lambda \int_{\mathbb{R}^d \times \mathcal{K}} \mathbb{I}_{L \cap (L+h)}(y) Q(dL) dy
$$

$$
= \lambda \int_{\mathcal{K}} |L \cap (L + h)| Q(dL)
$$

$$
= \lambda \mathbb{E} |M_0 \cap (M_0 + h)|.
$$
Thus, we obtain that
\[
\mathbb{P}(o \in \Xi^{(i)}, h \in \Xi^{(j)}) = \sum_{0 \leq k \leq i} \frac{\lambda^k}{k!} \left( \mathbb{E} |M_0| - \mathbb{E} |M_0 \cap (M_0 + h)| \right)^k \exp \left\{ -\lambda \left( \mathbb{E} |M_0| - \mathbb{E} |M_0 \cap (M_0 + h)| \right) \right\} 
\]
\[
\frac{\lambda^{i-k}}{(i-k)!} \left( \mathbb{E} |M_0 \cap (M_0 + h)| \right)^{i-k} \exp \left\{ -\lambda \mathbb{E} |M_0 \cap (M_0 + h)| \right\} \frac{\lambda^{j-i+k}}{(j-i+k)!} \left( \mathbb{E} |M_0| - \mathbb{E} |M_0 \cap (M_0 + h)| \right)^{j-i+k} \exp \left\{ -\lambda \left( \mathbb{E} |M_0| - \mathbb{E} |M_0 \cap (M_0 + h)| \right) \right\} 
\]
which confirms the assertion of Theorem 4.1.

5 Asymptotic test for volume fractions

In this section we presuppose that the multivariate CLT (2.2) holds and the asymptotic covariance matrix \( \Sigma_r \) has a strictly positive determinant \( \det(\Sigma_r) \). Further, assume that the symmetric, non-negative definite matrix \( \hat{\Sigma}_{n,r} \) defined by (3.1) is a (weakly) consistent estimator for \( \Sigma_r \) which is guaranteed by the conditions imposed on \((\Xi_1, \ldots, \Xi_r)^T\) in Theorem 3.1. By standard arguments one can show that \( ||\hat{\Sigma}_{n,r} - \Sigma_r|| \xrightarrow{P} 0 \) implies
\[
\det(\hat{\Sigma}_{n,r}) \xrightarrow{P_{n \to \infty}} \det(\Sigma_r) > 0,
\]
which in turn yields
\[
\mathbb{P}(\det(\hat{\Sigma}_{n,r}) > 0) \xrightarrow{P_{n \to \infty}} \hat{\Sigma}_{n,r}^{-1/2} \Sigma_r^{1/2} \xrightarrow{P_{n \to \infty}} I_r,
\]
where \( I_r \) stands for the \( r \)-dimensional unit matrix and \( \Sigma_r^{1/2} \) (resp. \( \hat{\Sigma}_{n,r}^{-1/2} \)) denotes the square root of \( \Sigma_r \) (resp. \( \hat{\Sigma}_{n,r}^{-1} \)). There are computationally efficient ways of calculating \( \hat{\Sigma}_{n,r}^{-1/2} \) based on the spectral decomposition
\[
\hat{\Sigma}_{n,r} = Q_n \text{diag}(\hat{\lambda}_{n,1}, \ldots, \hat{\lambda}_{n,r}) Q_n^T,
\]
where \( \hat{\lambda}_{n,1}, \ldots, \hat{\lambda}_{n,r} \) are the positive eigenvalues of \( \hat{\Sigma}_{n,r} \) (provided that \( \det(\hat{\Sigma}_{n,r}) > 0 \)) and \( Q_n \) is an orthogonal \( r \times r \)-matrix whose columns are eigenvectors of \( \hat{\Sigma}_{n,r} \). Thus,
\[
\hat{\Sigma}_{n,r}^{-1/2} = Q_n \text{diag}(\hat{\lambda}_{n,1}^{-1/2}, \ldots, \hat{\lambda}_{n,r}^{-1/2}) Q_n^T.
\]
is a possible choice for the square root of $\Sigma_r^{-1}$. There is an alternative approach - known as Cholesky decomposition method - to obtain the product representation $\tilde{\Sigma}_{n,r} = \tilde{\Sigma}_{n,r}^{1/2} (\tilde{\Sigma}_{n,r}^{1/2})^\top$, where $\tilde{\Sigma}_{n,r}^{1/2}$ can be chosen as a lower triangular $r \times r$-matrix. For details we refer the reader to Chapter 3 in Cressie (1991).

Summarizing the above facts and using Slutsky-type arguments we are in a position to state the following result: Under the hypothesis $H_0 : (p_1, \ldots, p_r)^\top = (p_0^1, \ldots, p_0^r)^\top$ we have

$$Z_{n,r} = \mathbb{1}(\det(\tilde{\Sigma}_{n,r}) > 0) \frac{\tilde{\Sigma}_{n,r}^{-1/2}}{\sqrt{|W_n|}} \begin{pmatrix} (\tilde{\Sigma}_{n,r}^{-1/2})(p_{n,1} - p_0^1) \\ \vdots \\ (\tilde{\Sigma}_{n,r}^{-1/2})(p_{n,r} - p_0^r) \end{pmatrix} \xrightarrow{d} N(0, I_r),$$

where $\tilde{\Sigma}_{n,r}^{-1/2}$ is a lower triangular matrix. For details we refer the reader to Chapter 3 in Cressie (1991).

Summarizing the above facts and using Slutsky-type arguments we are in a position to state the following result: Under the hypothesis $H_0 : (p_1, \ldots, p_r)^\top = (p_0^1, \ldots, p_0^r)^\top$ we have

$$Z_{n,r} = \mathbb{1}(\det(\tilde{\Sigma}_{n,r}) > 0) \frac{\tilde{\Sigma}_{n,r}^{-1/2}}{\sqrt{|W_n|}} \begin{pmatrix} (\tilde{\Sigma}_{n,r}^{-1/2})(p_{n,1} - p_0^1) \\ \vdots \\ (\tilde{\Sigma}_{n,r}^{-1/2})(p_{n,r} - p_0^r) \end{pmatrix} \xrightarrow{d} N(0, I_r),$$

which, by applying the continuous mapping theorem (see e.g. Daley & Vere-Jones (1988)), implies

$$||Z_{n,r}||^2 \xrightarrow{d} ||N(0, I_r)||^2,$$

where the random variable $||N(0, I_r)||^2$ is $\chi^2$-distributed with $r$ degrees of freedom.

Thus, for a given significance level $\alpha$, we will reject the hypothesis $H_0 : (p_1, \ldots, p_r)^\top = (p_0^1, \ldots, p_0^r)^\top$ if the test statistic $||Z_{n,r}||^2$ exceeds the critical value $\chi^2_{r,1-\alpha}$, which is determined by the equation $\mathbb{P}(||N(0, I_r)||^2 > \chi^2_{r,1-\alpha}) = \alpha$.

In the particular case $r = 2$ we get $\chi^2_{2,1-\alpha} = -2 \ln(\alpha)$ (since $||N(0, I_2)||^2$ is exponentially distributed with mean 2) so that the hypothesis $H_0 : (p_1, p_2)^\top = (p_0^1, p_0^2)^\top$ is rejected if $||Z_{n,2}|| > \sqrt{-2 \ln(\alpha)}$.

### 6 Numerical examples

To illustrate the fit of the estimator $\hat{C}_{W,kl}(h)$ given in (2.6) to the theoretical covariance $C_{kl}(h)$ in (1.6), we consider the realization of the multiphase model deduced from the Boolean model $\Xi$ in $\mathbb{R}^2$ with two different phases $\Xi_1$ and $\Xi_2$ within a sampling window $W$. If we additionally assume that $\Xi$ is isotropic, the covariance $C_{kl}(h)$ depends only on the radial coordinate $r = ||h||$. Notice that the estimation of $C_{kl}(h)$ from a two-dimensional image can only be done for vectors $h$, which take values on a two-dimensional lattice. Therefore, in the isotropic case it is convenient to estimate the covariance using the rotation average $\overline{C}_{W,kl}(r)$, where

$$\overline{C}_{W,kl}(r) = \frac{1}{\#\{h : ||h|| \approx r \}} \sum_{h : ||h|| \approx r} \hat{C}_{W,kl}(h).$$
Here, the set \( \{ h : \|h\| \approx r \} \) can be obtained using the so-called *midpoint circle algorithm*, which is a common algorithm in image analysis to detect a circle on a lattice; see, for example, Hearn & Baker (1997). This averaging method enables us to improve the accuracy of the estimation. An efficient way to compute the estimator \( \widehat{C}_{W,kl}(h) \) from the observed image is to use methods from spectral analysis; see, for example, Böhm, Heinrich, & Schmidt (2004), Ohser & Mücklich (2000), and Press, Flannery, Teukolsky, & Vetterling (1986). In particular, we compare \( C_{12}(h) \) using Theorem 4.1 with the corresponding estimator \( \widehat{C}_{W,12}(h) \) obtained from the realization of the multiphase model shown in Figure 1. In Figure 2 it can be seen that the estimator \( \overline{C}_{W,12}(r) \) fits the theoretical covariance \( C_{12}(r) \) quite good for the considered range of \( 0 \leq r \leq 120 \).

Figure 2: Comparison of \( C_{12}(h) \) (——) and \( \overline{C}_{W,12}(r) \) (- - -) for the multiphase model shown in Figure 1, where \( p_1 = 0.5, \ p_2 = 0.1534, \) and \( p_1 p_2 = 0.0767. \)
References


