Detecting anomalies in fibre systems using
3-dimensional image data

Denis Dresvyanskiy · Tatiana Karaseva · Vitalii Makogin · Sergei Mitrofanov · Claudia Redenbach · Evgeny Spodarev

Abstract We consider the problem of detecting anomalies in the directional distribution of fibre materials observed in 3D images. We divide the image into a set of scanning windows and classify them into two clusters: homogeneous material and anomaly. Based on a sample of estimated local fibre directions, for each scanning window we compute several classification attributes, namely the coordinate wise means of local fibre directions, the entropy of the directional distribution, and a combination of them. We also propose a new spatial modification of the Stochastic Approximation Expectation-Maximization (SAEM) algorithm. Besides the clustering we also consider testing the significance of anomalies. To this end, we apply a change point technique for random fields and derive the exact inequalities for tail probabilities of a test statistics. The proposed methodology is first validated on simulated images. Finally, it is applied to a 3D image of a fibre reinforced polymer.

Keywords Anomaly detection · classification · fibre composite · directional distribution · change-point problem · entropy · SAEM algorithm

This research was supported by the German Ministry of Education and Research (BMBF) through the project “AniS”, the DFG Research Training Group GRK 1932, DFG grant SP 971/10-1, as well as the DAAD scientific exchange program “Strategic Partnerships”.

D. Dresvyanskiy · T. Karaseva · S. Mitrofanov
Reshetnev Siberian State University of Science and Technology 31, Krasnoyarsky Rabochy Av., Krasnoyarsk, 660037, Russian Federation
E-mail: DresvyanskiyDenis@yandex.ru, tatyanakaraseva@yandex.ru, sergeimitrofanov95@gmail.com

V. Makogin · E. Spodarev
Institut für Stochastik, Universität Ulm, D-89069 Ulm
E-mail: vitalii.makogin@uni-ulm.de, E-mail: evgeny.spodarev@uni-ulm.de

C. Redenbach
Technische Universität Kaiserslautern, Fachbereich Mathematik, Postfach 3049, 67653 Kaiserslautern
E-mail: redenbach@mathematik.uni-kl.de

arXiv:1907.06988v1 [stat.ME] 16 Jul 2019
1 Introduction

Fibre composites, e.g., fibre reinforced polymers or high performance concrete, are an important class of functional materials. Physical properties of a fibre composite such as elasticity or crack propagation are influenced by its microstructure characteristics including the fibre volume fraction, the size or the direction distribution of the fibres. Therefore, an understanding of the relations between the fibre geometry and macroscopic properties is crucial for the optimisation of materials for certain applications. During the last years, micro computed tomography (CT) has proven to be a powerful tool for the analysis of the three-dimensional microstructure of materials.

In the compression moulding process of glass fibre reinforced polymers, the fibres order themselves inside the raw material as a result of mechanical pressure. During this process, deviations from the requested direction may occur, creating undesirable fibre clusters and/or deformations. These inhomogeneities are characterized by abrupt changes in the direction of the fibres, and their detection is studied in this paper.

The problem of detecting change points in random sequences, (multivariate) time series, panel and regression data has a long history, see the books [4,8,14,17,19,52]. Changes to be detected may concern the mean, variance, correlation, spectral density, etc. of the (stationary) sequence \( \{X_k, k \geq 0\} \). This kind of change detection has been considered by various authors starting with [11]. Sen and Srivastava [45] considered tests for a change in the mean of a Gaussian model. An overview can also be found in [6]. The CUSUM procedure, Bayesian approaches as well as maximum likelihood estimation are often used. Scan statistics come also into play naturally, see e.g. [7,8].

First approaches to change point analysis for random fields (or measures) have been developed in the papers [9,10,13,16,28,31,32,33,36,39,40,46,47,48], see also the review in [7, Section 2, D] and [8, Chapter 6]. The involved methods include M-estimation, minimax methods for risks, the geometric tube method, some non-parametric and Bayesian techniques. However, much is still to be done in this relatively new area of research.

In this paper, we develop a change-point test for \( m \)-dependent random fields. In the spirit of the book [8], it uses inequalities for tail probabilities of suitable test statistics. It is applied to the mean and the entropy of the local directional distribution of fibres observed in a 3D image of a fibre composite obtained by micro computed tomography. Characteristics are estimated in a moving scanning window that runs over the observed material sample, cf. [11,12]. Our main task is to detect areas with anomalous spreading of the fibres. Even though we focus on anomalies in fibres’ directions, our method will work with any local characteristic of fibres with values in a (compact) Riemannian manifold such as fibre length or mean curvature.

If an anomaly is present, its location is detected using a new spatial modification of the Stochastic Approximation Expectation Maximization (SAEM) algorithm (see [24] for a review of Expectation Maximization (EM) algorithms for the separation of components in a mixture of Gaussian distributions as well as a recent paper [37]). It allows for spatial clustering of the whole fibre material into a “normal” and an “anomaly” zones.
The paper is organized as follows. In Section 2, we introduce the stochastic model of a fibre process. In Section 3, we describe the procedure of generating the sample data, introduce the mean of local directions as well as their entropy. There, we compare two methods for entropy estimation: plug-in and nearest neighbor statistics. In Section 4, we consider the detection of anomalies as a change-point problem for the corresponding $m-$dependent random fields. In Section 5, we localize the anomalous region of fibres solving a clustering problem for multivariate random fields. For this purpose, we propose a new spatial modification of SAEM algorithm, which decreases the diffuseness of clusters. In Section 6, we apply our methods to 3D images of simulated (Section 6.1) and real (Section 6.2) fibre materials and compare their performance.

2 Problem setting

In this section, we give some basic definitions and results for fibre processes. For more details, see, for example, the book [18]. In 3-dimensional Euclidean space, a fibre $\gamma$ is a simple curve $\{\gamma(t) = (\gamma_1(t), \gamma_2(t), \gamma_3(t)), t \in [0, 1]\}$ of finite length satisfying the following assumptions:

- $\{\gamma(t), t \in [0, 1]\}$ is a $C^1$-smooth function.
- $\|\gamma'(t)\|_2^2 > 0$ for all $t \in [0, 1]$, where $\|\gamma'(t)\|_2^2 = |\gamma_1'(t)|^2 + |\gamma_2'(t)|^2 + |\gamma_3'(t)|^2$.
- A fibre does not intersect itself.

The collection of fibres forms a fibre system $\phi$ if it is a union of at most countably many fibres $\gamma^{(i)}$, such that any compact set is intersected by only a finite number of fibres, and $\gamma^{(i)}((0, 1)) \cap \gamma^{(j)}((0, 1)) = \emptyset$, if $i \neq j$, i.e., the distinct fibres may have only end-points in common. The length measure corresponding to the fibre system $\phi$ (and denoted by the same symbol) is defined by

$$\phi(B) = \sum_{\gamma^{(i)} \in \phi} h(\gamma^{(i)} \cap B)$$

for bounded Borel sets $B \in \mathcal{B}(\mathbb{R}^3)$, where $h(\gamma \cap B) = \int_0^1 \mathbb{1}[\gamma(t) \in B] \sqrt{|\gamma'(t)|^2} dt$ is the length of fibre $\gamma$ in window $B$. Then $\phi(B)$ is the total length of fibre pieces in the window $B$.

**Definition 1** A fibre process $\Phi$ is a random element with values in the set $\mathbb{D}$ of all fibre systems $\phi$ with $\sigma$-algebra $\mathcal{D}$ generated by sets of the form $\{\phi \in \mathbb{D} : \phi(B) < x\}$ for all bounded Borel sets $B$ and real numbers $x$. The distribution $P$ of a fibre process is a probability measure on $[\mathbb{D}, \mathcal{D}]$. The fibre process $\Phi$ is said to be stationary if it has the same distribution as the translated fibre process $\Phi_x = \Phi + x$ for all $x \in \mathbb{R}^3$.

For classification needs we consider an abstract fibre characteristic $w$. Let $(E, \mathcal{E}, \sigma)$ be a measurable space where $E$ is a (compact) Riemannian manifold equipped with a metric $\rho$. Let $w(x) \in E$ be some characteristic of a fibre at point $x \in \mathbb{R}^3$, assuming that exactly one fibre of $\Phi$ passes through $x$. Then a weighted random measure $\Psi$ can be defined by

$$\Psi(B \times L) = \int_B \mathbb{1}[w(x) \in L] \Phi(dx)$$
for bounded $B \in \mathcal{B}(\mathbb{R}^3)$ and $L \in \mathcal{E}$. Thus, $\Psi(B \times L)$ is the total length of all fibre pieces of $\Phi$ in $B$ such that their characteristic $w$ lies in range $L$.

As classifying characteristics $w$ we can for instance choose the fibre's local direction (with $E$ being the sphere $S^2$), their length or curvature (both with $E = \mathbb{R}$). In this article we focus on local directions of fibres, but the results can easily be applied to other choices of $w$.

If the fibre process $\Phi$ is stationary then the intensity measure of $\Psi$ can be written as $\mathbb{E}\Psi(B \times L) = \lambda |B| f(L)$, where $\lambda > 0$ is called the intensity of $\Psi$, $|\cdot|$ is the Lebesgue measure in $\mathbb{R}^3$ and $f$ is a probability measure on $S^2$ which is called the directional distribution of fibres. The distribution $f$ is the fibre direction distribution in the typical fibre point, hence length-weighted. In what follows, $|A|$ is either the cardinality of a finite set $A$ or the Lebesgue measure of $A$, if $A$ is uncountable and measurable.

If $H_1$ holds true, the region $A$ is called an anomaly region. In the following, we discuss how to test the hypothesis $H_0$ and how to detect the anomaly region $A$.

3 Data and clustering criteria

We assume that the dilated fibre system $\Xi \cap W$ is observed as a 3D greyscale image. Several methods for estimating the local fibre direction $w(x)$ in each fibre pixel $x \in \Xi$ are discussed in [51]. We use the approach based on the Hessian matrix that is implemented in the 3D image analysis software tool MAVI [26]. The smoothing parameter $\sigma$ required by the method is chosen as $\sigma = \hat{r}$, where $\hat{r}$ is an estimate of the (constant) thickness of the typical fibre. In the simulated samples, it is known. For the real data, it is obtained manually from the images.

We divide the observation window $W$ into small cubes $\tilde{W}_i$ (see Fig. [1] of the same size, whose edge length $\Delta$ equals three times the fibre diameter. The principal axis $\hat{w}_i$ of local directions (e.g. [51]) in each $\tilde{W}_i$, here referred to as “average local direction”, is then computed using the function $SubfieldFibreDirections$ in MAVI.
Let \( J_T = \{ i = (i_1, i_2, i_3), i_1 = 1, n_1, i_2 = 1, n_2, i_3 = 1, n_3 \} \) be the regular grid of cubes \( W_i \). Some of the cubes \( W_i \) may not contain enough fibre voxels to obtain a reliable estimate of the local fibre direction \( \hat{w}_i \). Let \( J \subset J_T \) be the subset of indices of cubes which allow for such an estimation. For each point \( i \in J \) denote by \( X_i = (x_i, y_i, z_i)^T \) the average local direction estimated from \( W_i \). We assume that our fibres are non-oriented and can be then transformed such that \( X_i \in S^2_+ \).
where $S^2_+ = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1, z \in [0, 1]\}$ is a hemisphere. The size of this sample is $N = |J| \leq n_1 n_2 n_3$.

Our main task is to determine the anomaly regions or, in other words, to classify the set of points $J$ into two clusters corresponding to the “homogeneous material” and the “anomaly” (one of these clusters can be empty). To do so, we combine $M^3$ of the small cubes $\tilde{W}_i$ (having edge length $\Delta$) to a larger cube $W_1$, such that the 3D image $W$ is divided into larger non-empty cubic observation windows (see Figure 1). The larger cubes have side length $M \Delta$ and the corresponding grid of the larger cubes is denoted by $J_W = \{(l_1, l_2, l_3) : l_1 = \bar{l}_m, l_2 = \bar{l}_m, l_3 = \bar{l}_m\}$. The set of indexes of non-empty cubes $\tilde{W}_i$ within $W_1$ is denoted by $S_l = \{(i_1, i_2, i_3) \in J, \tilde{W}_i(\tilde{i}_1, \tilde{i}_2, \tilde{i}_3) \subset W_1\}$. For each window $W_l, l \in J_W$, we estimate the entropy and the mean of the local directions, based on the estimates $\{\hat{w}_i : i \in S_l\}$ as described below.

### 3.1 Mean of local fibres direction

The vector $M_l = (M_{x,l}, M_{y,l}, M_{z,l})^T$ is calculated for each window $W_l, l \in J_W$ as the coordinate-wise sample mean of local directions (MLD):

$$M_{x,l} = \frac{1}{|S_l|} \sum_{i \in S_l} \hat{w}_{i1}, \quad M_{y,l} = \frac{1}{|S_l|} \sum_{i \in S_l} \hat{w}_{i2}, \quad M_{z,l} = \frac{1}{|S_l|} \sum_{i \in S_l} \hat{w}_{i3}.$$  

Note that $|S_l| \leq M^3$ and normally $|S_l| \approx M^3$.

### 3.2 The entropy of the directional distribution

The entropy of a random variable is a certain measure of the diversity/concentration of its range. Let $P$ be a probability distribution of a random element $X$ on an abstract measurable phase-space $(\mathcal{X}, \sigma)$. The value

$$E_X = -\int_{\Omega} \ln(P(X(\omega))) P(d\omega)$$  \hspace{1cm} (1)

is called the Shannon (differential) entropy of $X$. If $P$ is absolutely continuous with respect to some measure $\sigma$ then there exists the Radon-Nikodym derivative (or density) $f = \frac{dP}{d\sigma}$, and the entropy of $X$ has the following form

$$E_f := E_X = -\int_{\mathcal{X}} \ln(f(x)) f(x) \sigma(dx).$$  \hspace{1cm} (2)

In what follows, we assume that the random variable $X$ is absolutely continuous with density $f : \mathcal{X} \rightarrow \mathbb{R}_+$. In our problem setting, $X$ corresponds to the local fibre direction, $\mathcal{X}$ is the sphere $S^2$ and $\sigma$ is the spherical surface area measure on $S^2$. Since our fibres are non-oriented ($X \in S^2_+$), we may consider even local direction densities $f$ on the whole sphere $S^2$ where appropriate, instead of a density $f$ defined on $S^2_+$. However, choosing another classifying characteristic $w$ will lead to a different measurable space $(\mathcal{X}, \sigma)$.
3.3 Entropy estimation

In the literature, there is a large number of papers devoted to non-parametric entropy estimation for i.i.d. random vectors in $\mathbb{R}^D$, see e.g. the review in [5] and references in [12]. We will dwell upon two important estimates: the plug-in and the nearest neighbor ones.

#### 3.3.1 Plug-in estimator of entropy

For simplicity, define the plug-in estimator for directional distributions on the sphere $S^2$ with even densities $f$: $f(x) = f(-x), x \in S^2$. Its general definition on compact manifolds $X$ can be found e.g. in [2].

For a directional distribution density $f$, take the kernel estimator $\hat{f}_B(\cdot)$ on a window $B \subseteq J_T$ of the form

$$\hat{f}_B(y) = \frac{1}{|B|} \sum_{i \in B \cap J} \frac{\sin \rho(y, X_i)}{h^2 \rho(y, X_i)} K\left(\frac{\rho(y, X_i)}{h}\right),$$

where $h > 0$ denotes the bandwidth, $K : \mathbb{R}^+ \rightarrow \mathbb{R}$ is a kernel function and $\rho : S^2 \times S^2 \rightarrow \mathbb{R}^+$ is a geodesic metric given by $\rho(x, y) = \arccos\langle x, y \rangle, x, y \in S^2$, where $\langle \cdot, \cdot \rangle$ is the Euclidean scalar product in $\mathbb{R}^3$.

Then the plug-in estimator of $E_f$ in the window $W_l$ is given by

$$\hat{E}_{f,l} = -\frac{1}{|S_l|} \sum_{i \in S_l} \ln \hat{f}_{B+i}(X_i),$$

where $B \subseteq J_T$ is the sub-window and $B+i := \{j+i, j \in B\}$ denotes the translation of $B$.

For homogeneous marked Poisson point processes, the plug-in estimator $\hat{E}_f$ as above is considered in [2]. See also [1] for the context of Boolean models of line segments. We also made an attempt to apply this method to our 3D image data. But we met difficulties which basically come from the relatively small amount of data available. Namely, $\hat{E}_f$ needs a large number of points in sub-windows $B'$ during the estimation of $f$ together with a large number of such sub-windows. Let us illustrate these difficulties on a simple example.

**Example 1** Consider the uniform distribution on the sphere $S^2$, i.e., the density is $f(x) = \frac{1}{4\pi}, x \in S^2$. We generate a sample from this distribution and estimate its entropy $E_f$ using the plug-in estimator (4) with $|B| = |S^2|^{1/9}$ (as in [1]). The results are presented in Table 1. Moreover, we run the procedure 100 times and compare the obtained values with the exact value of the entropy

$$E_f = -\int_{S^2} \frac{1}{4\pi} \ln \frac{1}{4\pi} \sigma(dx) = \ln 4\pi \approx 2.53.$$
### Table 1 The plug-in entropy estimator for the uniform directional distribution on $S^2$

| Sample size | Mean     | Variance | MSE        |
|-------------|----------|----------|------------|
| 500000      | 2.476    | $2.848 \times 10^{-6}$ | 0.003      |
| 375000      | 2.456    | $5.034 \times 10^{-6}$ | 0.006      |
| 250000      | 2.418    | $6.293 \times 10^{-6}$ | 0.013      |
| 125000      | 2.309    | $7.245 \times 10^{-6}$ | 0.049      |
| 62500       | 2.099    | $2.739 \times 10^{-5}$ | 0.187      |
| 12500       | 0.981    | $8.917 \times 10^{-5}$ | 2.403      |
| 6250        | 0.359    | $1.251 \times 10^{-4}$ | 4.718      |
| 1250        | 0.008    | $4.695 \times 10^{-4}$ | 6.36       |

Subdivide the images only into 4 non-intersecting regions with more than 100000 cubes $W_i$. In other words, in order to test the hypotheses $H_0$ vs. $H_1$ with test statistics based on estimated entropy (4) we have a sample of size 4, which is too small, compare Section 4, inequality (24). There, for $m = 1$ the minimal sample size $|W|$ must be 1000.

#### 3.3.2 Nearest neighbor estimator of entropy

In order to overcome the above difficulties we apply another estimator of $E_f$ introduced in the paper by Kozachenko and Leonenko [35]. We call this estimator “Dobrushin estimator” because its main idea is due to Dobrushin [20]. The estimator from [35] cannot be applied directly, because it is designed for random vectors in a $d$-dimensional Euclidean space which is flat. In our setting, the phase space $(S^2, \sigma)$ is a manifold of positive constant curvature with geodesic metric $\rho$. Therefore, we take a version of Dobrushin estimator for the case of an $d-$dimensional compact Riemannian manifold $X$ with geodesic metric $\rho$ and Hausdorff measure $\sigma$.

For defining this estimator, the following results will be useful. Denote by $B_\delta(x)$ the ball in $(X, \rho)$ with radius $\delta > 0$ and center $x$, i.e., $B_\delta(x) = \{y \in X : \rho(x, y) \leq \delta\}$. Since a $\rho-$ball and a Euclidean $d$-dimensional ball are bi-Lipschitz equivalent, $d$ coincides with the Hausdorff dimension of the manifold $X$, see [24, Corollary 2.4]. Furthermore, for $\sigma-$almost all points $x \in X$ the Lebesgue density theorem is true, i.e.,

$$\sigma(B_\delta(x)) \sim c\delta^d, \quad \text{as} \quad \delta \to 0^+, \quad (5)$$

where $d = \text{dim}_H X$ is the Hausdorff dimension of $X$ and $c = 2^dD_X > 0$, where $D_X$ is the Hausdorff density of $X$, see [24, Proposition 4.1,5.1] and [44, Theorem 30].

**Definition 2** Let $(\xi_1, \ldots, \xi_N)$ be a sample of i.i.d. $X-$valued random elements with continuous density function $f : X \to \mathbb{R}^+$. Denote by $\rho_i$ the distance to the nearest neighbor of $\xi_i$, $i = 1, N$, i.e., $\rho_i = \min_{j=1, N, j \neq i} \{\rho(\xi_i, \xi_j)\}$. Define the statistic

$$\hat{E} = \frac{d}{N} \sum_{i=1}^N \ln \rho_i + \ln(c(N-1)) + \gamma, \quad (6)$$

where $c$ and $d$ are defined by (5) and $\gamma = -\int_0^\infty (\ln y) e^{-y} dy \approx 0.5772$ is Euler’s constant. The statistic $\hat{E}$ is called nearest neighbor (Dobrushin) estimator of the entropy.
Detecting anomalies in fibre systems using 3-dimensional image data

Table 2 The Dobrushin estimator (6) for the uniform directional distribution on the sphere.

| Sample size | Mean | Variance | MSE |
|-------------|------|----------|-----|
| 125         | 2.51 | 0.02     | 0.02|
| 64          | 2.50 | 0.03     | 0.02|

It coincides with the nearest-neighbour entropy estimate given in [22, p. 2169] with the only difference that in [22] Euclidean distances between \( \xi \) are used instead of geodesic distances \( \rho \). The \( L_2 \)-consistency of \( \hat{E} \) is proven in [22, Theorem 2.4] for i.i.d. samples \( (\xi_1, \ldots, \xi_N) \) as above with bounded density \( f \) of compact support.

In fact, a large class of parametric distributions on a sphere, including the Fisher, the Watson or the Angular Central Gaussian distribution, has bounded densities with compact support.

**Remark 1** In many problems of probability theory, limit theorems for independent observations remain true for weakly dependent data. Since the fibre materials are weakly dependent (the fibers have a finite length), we can assume that the entropy and mean local directions are weakly dependent as well. The proof of consistency of (6) for weakly dependent \( \xi_i \) is non-trivial and goes beyond the scope of this paper.

**Remark 2** Our data sets consist of straight fibres which are longer than the edge length \( \Delta \) of small observation windows \( \tilde{W} \). Such fibres yield several almost equal values of average local directions \( X_i \). This leads to very small values of a distance to the nearest neighbor \( \rho_i \) and, consequently, to the large negative bias of \( \hat{E} \) which is computed using \( \log \rho_i \). Trying to eliminate this bias, we propose to use the following version of (6)

\[
\hat{E} = \frac{1}{\sum_{i=1}^{N} \mathbb{I}(\rho_i > \rho_0)} \sum_{i=1}^{N} \mathbb{I}(\rho_i > \rho_0) \ln \rho_i + \ln c \\
+ \ln \left( \sum_{i=1}^{N} \mathbb{I}(\rho_i > \rho_0) - 1 \right) + \gamma
\]

(7)

with penalty value \( \rho_0 = 0.01 \) found by computational tuning.

In order to test the accuracy of the Dobrushin estimator, we have generated 100 samples from the uniform directional distribution on \( S^2 \). We have computed the Dobrushin statistic and compared it with the exact entropy value \( \ln(4\pi) \approx 2.53 \).

The results are presented in Table 2.

Based on these results, we conclude that the Dobrushin estimator (6) is quite accurate for small sample sizes. Even for a sample with 64 entries the entropy is estimated much better than by the plug-in method.

4 Change point detection in random fields

To test the hypothesis \( H_0 \) against \( H_1 \), we check the existence of anomaly regions in a realization of an \( m \)-dependent geometric random field \( \{s_k, k \in W\} \). Here we
follow the ideas from [6], where change-point problems for mixing random fields on general parametric (disorder) regions were considered. The field \( \{s_k, k \in W\} \) will be chosen in a way such that the hypothesis \( H_0 \) implies that it is stationary, whereas \( H_1 \) means the presence of a region \( I_\theta \subset W \) with different mean value of \( s_k \). Later in our application to fibre materials in Section 4.3, we assume the anomaly region to be a box \([11]\).

4.1 Random fields with inhomogeneities in mean

Let \( \{\xi_k, k \in \mathbb{Z}^3\} \) be an integrable stationary real-valued random field with \( \mu = \mathbb{E}\xi_k \). Denote by \( \xi_k \) the centered field \( \xi_k = \xi_k - \mu, k = (k_1,k_2,k_3) \in \mathbb{Z}^3 \). Moreover, we assume that \( \{\xi_k, k \in \mathbb{Z}^3\} \) is \( m \)-dependent, i.e., \( \xi_k \) and \( \xi_l \) are independent if \( \max_{i=1,2,3}|k_i - l_i| > m, k = (k_1,k_2,k_3) \in \mathbb{Z}^3, l = (l_1,l_2,l_3) \in \mathbb{Z}^3 \). Let \( \Theta \) be a finite parameter space. For every \( \theta \in \Theta \) we define subsets of anomalies \( I_\theta \subset \mathbb{Z}^3 \) completely determined by a parameter \( \theta \). Then for some \( \theta_0 \in \Theta \) we observe

\[
s_k = \tilde{\xi}_k + h \mathbb{I}\{k \in I_{\theta_0}\}, \quad k \in W, \tag{8}
\]

where \( W = [1,M_1] \times [1,M_2] \times [1,M_3] \cap \mathbb{Z}^3 \), and \( h \in \mathbb{R} \). Assume that \( I_\theta \subset W \) for every \( \theta \in \Theta \). Denote \( I_{\theta}^c = W \setminus I_{\theta} \).

Let \( \Theta_0 \) correspond to the values of \( \theta \) for anomalies which we consider as significant, i.e., they are neither extremely small nor represent the majority of the data. Formally, for \( \gamma_0, \gamma_1 \in (0,1), \gamma_0 < \gamma_1 \), we let

\[
\Theta_0 = \{ \theta \in \Theta : |I_\theta| \leq |I_{\theta_0}| \leq \gamma_1 |I_\theta| \}.
\]

Then \( \Theta_1 = \Theta \setminus \Theta_0 \) corresponds to extremely small or large anomalies, i.e.,

\[
\Theta_1 = \{ \theta \in \Theta : |I_\theta| < \gamma_0 |I_\theta|, \text{ or } |I_{\theta}^c| < (1 - \gamma_1)|I_\theta| \}.
\]

4.2 Testing the change of expectation

Now we can formulate the change-point hypotheses for the random field \( \{s_k, k \in W\} \) with respect to its expectation as follows.

\( H_0' : \mathbb{E}s_k = \mu \) for every \( k \in W \) (i.e. \( h = 0 \)) vs.

\( H_1' : \exists \theta_0 \in \Theta_0 \) such that \( \mathbb{E}s_k = \mu + h, k \in I_{\theta_0}, \ h \neq 0, \) and \( \mathbb{E}s_k = \mu, k \in I_{\theta_0}^c \).

Consider the following change-in-mean statistics for the sample \( S = \{s_k, k \in W\} \):

\[
Z(\theta) = \frac{1}{|I_\theta|} \sum_{k \in I_\theta} s_k - \frac{1}{|I_{\theta}^c|} \sum_{k \in I_{\theta}^c} s_k \tag{9}
\]

\[
= \frac{1}{|I_\theta|} \sum_{k \in I_\theta} (\xi_k + \mu + h \mathbb{I}\{k \in I_{\theta_0}\})
\]

\[
- \frac{1}{|I_{\theta}^c|} \sum_{k \in I_{\theta}^c} (\xi_k + \mu + h \mathbb{I}\{k \in I_{\theta_0}\})
\]

\[
= \frac{1}{|I_\theta|} \sum_{k \in I_\theta} \xi_k - \frac{1}{|I_{\theta}^c|} \sum_{k \in I_{\theta}^c} \xi_k + h \left( \frac{|I_\theta \cap I_{\theta_0}|}{|I_\theta|} - \frac{|I_{\theta}^c \cap I_{\theta_0}|}{|I_{\theta}^c|} \right). \tag{10}
\]
Denote by
\[ \eta(\theta) := Z(\theta) - \mathbb{E}Z(\theta) \]
the centered field \( Z(\theta) \). In order to test \( H'_0 \) vs. \( H'_1 \) we use the test statistics
\[ T_W(S) = \max_{\theta \in \Theta_0} |Z(\theta)| \]
\[ = \max_{\theta \in \Theta_0} \left| \frac{1}{|I_\theta|} \sum_{k \in I_\theta} s_k - \frac{1}{|I_\theta|} \sum_{k \in I_c^\theta} s_k \right| \] (11)
We reject \( H'_0 \) if \( T_W(S) \) exceeds the critical value \( y_\alpha \). Let us find such \( y_\alpha > 0 \) via the probability of the 1st-type error \( P_{H'_0}(\max_{\theta \in \Theta_0} |Z(\theta)| \geq y_\alpha) \leq \alpha \). It holds
\[ P_{H'_0}(\max_{\theta \in \Theta_0} |Z(\theta)| \geq y_\alpha) \]
\[ = P_{H'_0} \left( \max_{\theta \in \Theta_0} \eta(\theta) + h \left( \frac{|I_\theta \cap I_{\theta_0}|}{|I_\theta|} - \frac{|I_c^\theta \cap I_{\theta_0}|}{|I_c^\theta|} \right) \geq y_\alpha \right) \]
\[ = P\left( \max_{\theta \in \Theta_0} |\eta(\theta)| \geq y_\alpha \right). \]
Thus, we find the bounds for tail probabilities of the random variable \( \max_{\theta \in \Theta_0} |\eta(\theta)| \). To do so, we use the ideas from [29] to get the following bounds for \( m \)-dependent random fields. For the sake of generality, our results are formulated in \( Z^D, D \in \mathbb{N} \).

**Theorem 1** Let \( \{\xi_k, k \in Z^D\} \) be a stationary real-valued \( m \)-dependent centered random field and \( \{b_k, k \in Z^D\} \) be real numbers. Assume that there exist \( H, \sigma > 0 \) such that
\[ \mathbb{E}|\xi_k|^p \leq \frac{p^p}{2} H^p \sigma^2, \quad p = 2, 3, \ldots \] (12)
Then for any \( W \subset Z^D, |W| < \infty \) we have
\[ P \left( \left| \sum_{k \in W} \xi_k b_k \right| \geq y \right) \leq 2 \exp \left( -\frac{y^2}{4m^D \sigma^2 \|b\|^2_2} \right) \]
\[ \times \mathbb{I} \left\{ 0 < y \leq \frac{\sigma^2 \|b\|^2_2}{H \|b\|_\infty} \right\} \]
\[ + 2 \exp \left( -\frac{y}{2H m^D \|b\|_\infty} + \frac{\sigma^2 \|b\|^2_2}{4H^2 m^D \|b\|^2_\infty} \right) \]
\[ \times \mathbb{I} \left\{ y > \frac{\sigma^2 \|b\|^2_2}{H \|b\|_\infty} \right\}, \]
where \( \|b\|_\infty = \max_{k \in W} |b_k| \) and \( \|b\|^2_\infty = \sum_{k \in W} b_k^2 \).
Proof Using the Markov inequality we have for any $u > 0$ that

$$
P \left( \sum_{k \in W} \xi_k b_k \geq y \right) = P \left( \sum_{k \in W} \xi_k b_k \geq y \right) + \frac{\mathbb{E} \exp(u \sum_{k \in W} \xi_k b_k)}{\exp(uy)} \exp(uy).$$  \tag{13}

Denote by $W(l) = \{l + mi \in W | i \in \mathbb{Z}^D\}$ for $l \in \{1, \ldots, m\}^D$. It follows from Hölder’s inequality and $m$–dependence that

$$
\mathbb{E} \exp \left( u \sum_{k \in W} \xi_k b_k \right) = \mathbb{E} \left( \prod_{l \in \{1, \ldots, m\}^D} \exp \left( u \sum_{k \in W(l)} \xi_k b_k \right) \right)^{1/m^D}
\leq \prod_{l \in \{1, \ldots, m\}^D} \left( \mathbb{E} \exp \left( m^D u \sum_{k \in W(l)} \xi_k b_k \right) \right)^{1/m^D}
= \prod_{l \in \{1, \ldots, m\}^D} \left( \prod_{k \in W(l)} \mathbb{E} e^{m^D u b_k \xi_k} \right)^{1/m^D}.
\tag{14}
$$

From Taylor’s expansion we have for $u \in \left[ 0, \frac{1}{2m^D \sigma^2} \right]$

$$
\mathbb{E} e^{m^D u b_k \xi_k} = 1 + m^D u b_k \mathbb{E} \xi_k + \frac{1}{2} m^2 u^2 b_k^2 \mathbb{E} \xi_k^2
+ m^{2d} u^2 b_k^2 \sum_{p \geq 3} \frac{1}{p!} m^{d(p-2)} u^{p-2} b_k^p \mathbb{E} \xi_k^p
\leq 1 + \frac{1}{2} m^{2d} u^2 b_k^2 \sigma^2 + \frac{1}{2} m^{2d} u^2 b_k^2 \sigma^2 \sum_{p \geq 3} (Hm^D u |b_k|)^{p-2}
= 1 + \frac{1}{2} m^{2d} u^2 b_k^2 \sigma^2
\leq \exp(m^{2d} u^2 b_k^2 \sigma^2).
\tag{15}
$$

Combining (14) and (15) we continue for the first term in (13) with the following bound for $0 \leq u \leq \frac{1}{2m^D \sigma^2}$. \hfill \square
Detecting anomalies in fibre systems using 3-dimensional image data

\[
(2Hm^D \max_{k \in W} |b_k|)^{-1} : \\
e^{-uy} \prod_{k \in W} \left( \mathbb{E} e^{m^D u b_k \xi_k} \right)^{1/m^D} \\
\leq \exp \left( -uy + m^D u^2 \sigma^2 \sum_{k \in W} b_k^2 \right). 
\]

(16)

The minimum of (16) is achieved for \( u = y/(2m^D \sigma^2 \|b\|_\infty^2) \). Moreover, bound (16) is valid for the second term in (13), too. Therefore, for \( 0 \leq y \leq \sigma^2 \|b\|_\infty^2 \) we have

\[
P \left( \left| \sum_{k \in W} \xi_k b_k \right| \geq y \right) \leq 2 \exp \left( -\frac{y^2}{4m^D \sigma^2 \|b\|_\infty^2} \right).
\]

For \( y > \sigma^2 \|b\|_\infty^2 \) we put \( u = (2Hm^D \|b\|_\infty^2)^{-1} \) in (16) and obtain

\[
P \left( \left| \sum_{k \in W} \xi_k b_k \right| \geq y \right) \leq 2 \exp \left( -\frac{y}{2Hm^D \|b\|_\infty} + \frac{1}{4H^2 m^D \|b\|_\infty^2 \sigma^2 \|b\|_\infty^2} \right).
\]

This completes the proof.

We apply Theorem 1 to \( \eta(\theta), \theta \in \Theta_0, D = 3 \). First, we rewrite \( \eta(\theta) \) as

\[
\eta(\theta) = \frac{1}{|I_\theta|} \sum_{k \in W} \xi_k \mathbb{1}\{k \in I_\theta\} - \frac{1}{|I_\theta|} \sum_{k \in W} \xi_k \mathbb{1}\{k \in I_\theta^c\} \\
= \sum_{k \in W} \xi_k b_k(\theta), \theta \in \Theta_0,
\]

where

\[
b_k(\theta) = \frac{\mathbb{1}\{k \in I_\theta\}}{|I_\theta|} - \frac{\mathbb{1}\{k \in I_\theta^c\}}{|I_\theta^c|}.
\]

**Corollary 1.** Let \( |I_\theta| \leq |I_\theta^c| \) for \( \theta \in \Theta_0 \) and \( |\xi_k| \leq M_0, k \in W \) a.s., then under the conditions of Theorem 1 we have that

\[
P \left( |\eta(\theta)| \geq y \right) \\
\leq 2 \exp \left( -\frac{y^2}{4m^3 \sigma^2 |W|} |I_\theta| \right) \mathbb{1}\left\{ 0 < y \leq \frac{\sigma^2 |W|}{M_0 |I_\theta^c|} \right\} \\
+ 2 \exp \left( -\frac{y}{2M_0 m^3 |I_\theta|} + \frac{\sigma^2 |W|}{4M_0^2 m^3 |I_\theta^c|} |I_\theta| \right) \\
\times \mathbb{1}\left\{ y > \frac{\sigma^2 |W|}{M_0 |I_\theta^c|} \right\}.
\]
Proof From the definition of \(b_k(\theta)\) we have
\[
\|b(\theta)\|^2 = \sum_{k \in W} \left( \frac{\mathbb{I}\{k \in I_0\}}{|I_0|} + \frac{1}{|I_0|} \right)^2
\]
\[
= \frac{1}{|I_0|} + \frac{1}{|I_0|} = \left| W \right| \frac{1}{|I_0|},
\]
and
\[
\|b(\theta)\| = \max_{k \in W} |b_k(\theta)| = \max \left( \frac{1}{|I_0|}, \frac{1}{|I_0|} \right) = \frac{1}{|I_0|}.
\]
Since \(|\xi_k| \leq M_0\) then \(E[\xi_k^p] \leq \mathbb{E}(\xi_k^2)_{|\xi_k|^{p-2}} \leq M_0^{p-2} \mathbb{E}\xi_k^2 \leq M_0^{p-2} \sigma^2\). Therefore, we put \(H = M_0\) in (12) and obtain the statement of the corollary.

Since \(P(\max_{\theta \in \Theta_0} |\eta(\theta)| \geq y) \leq \sum_{\theta \in \Theta_0} P(|\eta(\theta)| \geq y)\), we have the following corollary.

Corollary 2 Let \(0 < \gamma_0 < \gamma_1 < 1/2\) and assume that the conditions of Theorem 4 and Corollary 4 are satisfied. Then
\[
P\left( \max_{\theta \in \Theta_0} |\eta(\theta)| \geq y \right)
\]
\[
\leq \sum_{\theta \in \Theta_0: |I_0| \leq \frac{\sigma^2|W|}{yM_0}} 2e^{-\frac{\sigma^2}{4yM_0^2} |I_0| |I_0| |W|} \]
\[
+ \sum_{\theta \in \Theta_0: |I_0| > \frac{\sigma^2|W|}{yM_0}} 2e^{-\frac{\sigma^2}{2yM_0} |I_0| + \frac{\sigma^2}{yM_0} |I_0| |I_0|} \leq \alpha. \tag{18}
\]
Hence, we reject \(H_0\) if the test statistic \(T_W(S) \geq y_\alpha\), where critical value \(y_\alpha\) is the minimum positive number such that
\[
\sum_{\theta \in \Theta_0: |I_0| \leq \frac{\sigma^2|W|}{yM_0}} 2e^{-\frac{\sigma^2}{4yM_0^2} |I_0| |I_0| |W|} \]
\[
+ \sum_{\theta \in \Theta_0: |I_0| > \frac{\sigma^2|W|}{yM_0}} 2e^{-\frac{\sigma^2}{2yM_0} |I_0| + \frac{\sigma^2}{yM_0} |I_0| |I_0|} \leq \alpha. \tag{19}
\]

Remark 3 In the case of Gaussian random field \(\{\xi_k, k \in \mathbb{Z}^D\}\), we can obtain the same statements of Corollaries 1 and 2 with \(M_0 = \sigma\). Indeed, \(E[\xi_0]^p \leq \sigma^{p-2} \sigma^2 |\zeta|^{p-2}, p \geq 2\), where \(\zeta \sim N(0, \sigma^2)\). So we put \(H = M_0\) in (12).

Simplifying the result of Corollary 2 we get that (17) and (18) is bounded by
\[
2 \left\{ \theta \in \Theta_0: |I_0| \leq \frac{\sigma^2|W|}{yM_0} \right\} e^{-\frac{\sigma^2}{4yM_0^2} |W| \gamma_0 (1-\gamma_0)}
\]
\[
+ 2 \left\{ \theta \in \Theta_0: |I_0| > \frac{\sigma^2|W|}{yM_0} \right\} e^{-\frac{\sigma^2}{2yM_0} |W| \gamma_0}.
\]
Particularly, if \( y < \frac{\sigma^2}{M_0(1-\gamma_0)} \) then
\[
P\left( \max_{\theta \in \Theta_0} |\eta(\theta)| \geq y \right) \leq 2 |\Theta_0| e^{-\frac{\sigma^2}{4M_0 m^3 W} |\gamma_0(1-\gamma_0)|}, \tag{20}
\]
and if \( y > \frac{\sigma^2}{M_0(1-\gamma_1)} \) then
\[
P\left( \max_{\theta \in \Theta_0} |\eta(\theta)| \geq y \right) \leq 2 |\Theta_0| e^{-\frac{\sigma^2}{4M_0 m^3 \gamma_0 W}}. \tag{21}
\]

4.3 Change-point detection in simulated random fields

In this section, we study the behaviour of the test statistics \( T_W(S) \) given in [11] and probabilities of 1st-type error \( P_{H_0}(T_W(S) \geq y_0) \) with respect to different values of \( \sigma^2 \) and \( m \).

The form of \( T_W \) allows to test the existence of the anomaly regions of arbitrary form and arbitrary number of connected components. On the other hand, we need to decrease the value \( |\Theta| \) up to a feasible quantity for computational reasons (see bounds (20) and (21)). Let \( W = [1, M_1] \times [1, M_2] \times [1, M_3] \cap \mathbb{N}^3 \). We fix \( \gamma_0 = 0.05 \) and \( \gamma_1 = 0.5 \), as the anomaly should not cover the majority of the window. In this paper we restrict \( I_0 \) to be a single rectangular parallelepiped of the form \( [1 + \Delta_0 i_1, 1 + \Delta_0 i_1 + \Delta_1 l_1] \times [1 + \Delta_0 i_2, 1 + \Delta_0 i_2 + \Delta_1 l_2] \times [1 + \Delta_0 i_3, 1 + \Delta_0 i_3 + \Delta_1 l_3] \).

Then the parametric set of significant anomaly regions is given by
\[
\Theta_0 = \{(1 + \Delta_0 i_1, 1 + \Delta_0 i_2, 1 + \Delta_0 i_3, \Delta_1 l_1, \Delta_1 l_2, \Delta_1 l_3) \}
\]
where \( (i_1, i_2, i_3, l_1, l_2, l_3) \in \mathbb{N}_0^6 \),
\[
1 + \Delta_0 i_j + \Delta_1 l_j \leq M_j, l_j \geq L_M, 1 \leq j \leq 3, \tag{22}
\]
\[
\gamma_0 \leq \frac{|I_0 \cap J|}{M_1 M_2 M_3} \leq \gamma_1.
\]

The offset parameters \( \Delta_0 \) and \( \Delta_1 \) as well as the parameter \( L_M \) controlling the minimal edge length of the cuboids have to be chosen by the user.

Assuming the \( m \)-dependence for our observations, we do not know the exact value of \( m \). Hence, we need to impose some restrictions on the field \( \xi \). First, if we know a-priori the maximum length of a typical fiber we can immediately obtain the bound for \( m \). Second, we can estimate the covariance function of the random field \( \{s_k, k \in W\} \) and assess \( m \) as the range when this empirical covariance is sufficiently close to zero. From relation (21) with \( \frac{\sigma^2}{M_0(1-\gamma_0)} < y < M \) we obtain the following approximate bound for an admissible \( m \) :
\[
2 |\Theta_0| \exp \left(-\frac{y}{4M m^3} \gamma_0 W\right) \leq \alpha \Rightarrow m^3 \leq \frac{\gamma_0}{4 \log(2|\Theta_0|/\alpha)} |W|. \tag{23}
\]

For example, for \(|\Theta_0| = 10^4\), \( \alpha = 0.05, \gamma_0 = 0.05 \), one gets
\[
m \leq \frac{1}{10} |W|^{1/3} \text{ or } |W| \geq 10^3 m^3. \tag{24}
\]
Let us now compare the empirical probability of the error of the 1-st type with the bounds (19) for $P_{H_0}(T_W(\cdot) \geq y_\alpha)$. We generate 300 realizations of a Gaussian centered $m$-dependent $(m = 10)$ random field $\{Y_k, k \in W\}$ with $W = [1, 80] \cap \mathbb{N}^3$ (which is matched to the considered data sets) and $Y_k \sim N(0, 1)$. The dependence is modelled as follows: random variables $Y_{1+mk}, k \geq 0$ are independent, and $Y_{1+mk} = Y_{l+mk}, k \in \mathbb{N}^3$ for all $l \in \{1, \ldots, m\}$. We take $\Delta_0 = \Delta_1 = 8$, $\gamma_0 = 0.05$, and $\gamma_1 = 0.5$. In this case, $|\Theta_0| = 11954$.

Based on the simulated sample of values of the test statistics $T_W(Y)$ we compute the empirical critical value $\hat{y}_\alpha = 0.6396$ for $\alpha = 0.05$. From comparison of $\hat{y}_\alpha$ with critical values $y_\alpha$ based on inequality (19) with $M_0 = \sigma$ (presented in Table 3), we see that even under the exact value of $\sigma^2 = 1$, critical values $y_\alpha$ are quite conservative. For example, $y_\alpha = 0.7198$ for $m = 7$ is still greater than $\hat{y}_\alpha = 0.6396$ generated for $m = 10$.

Therefore, we can use critical values from inequality (19) with $m$ smaller than its real value.

Table 3 Critical values $y_{0.05}$ based on inequality (19) for different values of $m$ and $\sigma$.

| $\sigma^2$ | $m = 10$ | $m = 9$ | $m = 8$ | $m = 7$ | $m = 6$ | $m = 5$ | $m = 4$ | $m = 3$ | $m = 2$ |
|-----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| 1         | 1.0757   | 1.0492   | 0.8793   | 0.7198   | 0.5711   | 0.4345   | 0.3109   | 0.2019   | 0.1099   |
| 4         | 2.1513   | 2.0985   | 1.7587   | 1.4394   | 1.1423   | 0.8890   | 0.6218   | 0.4039   | 0.2198   |
| 8         | 3.0424   | 2.9677   | 2.4871   | 2.0357   | 1.6154   | 1.2289   | 0.8793   | 0.5711   | 0.3109   |

5 Cluster based anomaly detection

For processes $\Xi$ of thick fibres introduced in Section 2, the evidence of an anomaly is tested by applying the test of Section 4 to the random field $\{s_k, k \in W\}$ of estimated local mean or entropy of the chosen fibre characteristic $w$. In this paper, $w(x)$ is the average direction vector of the fibres of $\Xi$ at $x \in W$ or one of its coordinates (introduced in Section 3 as $\hat{w}_i$).

Assume that the anomaly test presented in Section 4 rejected the hypothesis $H'_0$ (and hence $H_0$), i.e., we have an evidence of an anomalous fibre behaviour in the rectangular subregion $I_{\theta_0}$ of our image data. Now we are interested in a more accurate estimate of the geometry of this anomaly. The search for an anomaly region in a 3D image can be interpreted as a problem of splitting the volume of the image into two disjoint clusters: **homogeneous material** and **anomaly**.

In our problem setting, the volume under investigation, $\bigcup_{i \in J_W} W_i$, is a union of scanning windows $W_1$ with meaningful local direction information. Each of them yields the clustering attributes mean of local directions (MLD) and entropy. We need to classify all the windows $W_1$ as either belonging to the homogeneous material or the anomaly. For this purpose, a spatial version of the Stochastic Expectation Maximization algorithm is used.
5.1 Spatial modification of a Stochastic Approximation Expectation
Maximization (SAEM) algorithm

We assume that under the alternative \( H_1 \) (see page 2), fibres in the material
may have two different distributions \( f_0 \) and \( f_1 \) of local directions. Therefore, the
distribution of clustering attributes is a mixture of the distributions \( f_0 \) and \( f_1 \).

The Expectation-Maximization algorithm (EM) is commonly used to separate
modes in a finite mixture of distributions, cf. [34] for a review. It is an iterative
procedure consisting of two steps: Expectation (Estimation) and Maximization.
In general, one assumes that the probability law under study is a mixture of
\( k \) distributions from the same parametric family. In the first step, the hidden
parameters of the sample distribution, i.e., the weights of the mixture components,
are estimated, while in the second step the resulting parameters are updated by
maximizing the likelihood function.

Since the EM algorithm belongs to the so-called “greedy” algorithms, that is,
it converges to the first local optimum that has been found, a modification that
compensates this deficiency should be used. One way out is a random “shaking”
of observations in each iteration. This method is the basis of the Stochastic EM
(SEM) algorithm (cf. [27,34]).

The SEM algorithm works relatively fast in comparison with other methods,
and its results are non-sensitive to an initial approximation. Random perturbations
on the parameter space in the S-step guarantee the convergence to the global
maximum of the likelihood function and help to avoid unstable local maxima. On
the other hand, the outputs of the SEM algorithm form a Markov chain and the
final solution is its stationary distribution. To avoid this additional problem we use
a modification called SAEM (Stochastic Approximation of EM) algorithm which
brings together advantages of both EM and SEM approaches, e.g. [15].

Assume that the observable distribution has a density of the form
\[
\varphi(x) = \beta \varphi(x, \delta_1) + (1 - \beta) \varphi(x, \delta_2), \quad x \in \mathbb{R}^d,
\]
where \( \varphi(x, \delta) \) is a multivariate Gaussian density with unknown parameter \( \delta = (\mu_i, \Sigma_i) \), \( i = 1, 2 \) and \( \beta \in [0, 1] \). Here \( \mu_i \) is the mean and \( \Sigma_i \) is the covariance
matrix of Gaussian component \( i = 1, 2 \). The combined unknown parameter is
\( \delta = (\beta, \delta_1, \delta_2) \). We call \( \varphi(\cdot, \delta_1) \) and \( \varphi(\cdot, \delta_2) \) the first and the second component
of the mixture, respectively. For each observation \( x, i \in J_W \), we define a new
variable \( y_i = I\{x_i \text{ belongs to the first component}\} \). Therefore, we have two samples:
observable \( \mathbf{x} = \{x_i, i \in J_W\} \) and unobservable \( \mathbf{y} = \{y_i, i \in J_W\} \). Then the log-
likelihood function equals
\[
\ln L(\delta, \mathbf{x}, \mathbf{y}) = \sum_{i \in J_W} [y_i \ln(\beta \varphi(x_i, \delta_1)) + (1 - y_i) \ln((1 - \beta) \varphi(x_i, \delta_2))] \\
= \nu_1 \ln \beta + \nu_2 \ln(1 - \beta) + \sum_{i \in J_W : y_i = 1} \ln \varphi(x_i, \delta_1) \\
+ \sum_{i \in J_W : y_i = 0} \ln \varphi(x_i, \delta_2),
\]
where \( \nu_1 = \sum_{l \in J_1} y_l \) denotes the number of observations belonging to the first mixture component and \( \nu_2 = m_3 \nu_1 - \nu_1 \) observations belong to the second one.

Assume that we know the a posteriori probability \( q_1^{(k-1)}, l \in J_1, \) that \( x_1 \) belongs to the first component \( \varphi(\cdot, \delta_1), \) where \( k - 1 \) is the iteration number. Let us describe the EM part. During the M-Step we obtain new estimates of the parameters \( \delta_1^{(k)} = (\mu_1^{(k)}, \Sigma_1^{(k)}), \delta_2^{(k)} = (\mu_2^{(k)}, \Sigma_2^{(k)}), \beta^{(k)} \) by

\[
\begin{align*}
\mu_1^{(k)} &= \frac{\sum_{l \in J_1} q_1^{(k-1)} x_l}{\sum_{l \in J_1} q_1^{(k-1)}}, \\
\Sigma_1^{(k)} &= \frac{\sum_{l \in J_1} q_1^{(k-1)} (x_l - \mu_1^{(k)})(x_l - \mu_1^{(k)})^T}{\sum_{l \in J_1} q_1^{(k-1)}}, \\
\mu_2^{(k)} &= \frac{\sum_{l \in J_2} (1 - q_1^{(k-1)}) x_l}{\sum_{l \in J_2} (1 - q_1^{(k-1)})}, \\
\Sigma_2^{(k)} &= \frac{\sum_{l \in J_2} (1 - q_1^{(k-1)}) (x_l - \mu_2^{(k)})(x_l - \mu_2^{(k)})^T}{\sum_{l \in J_2} (1 - q_1^{(k-1)})}, \\
\beta^{(k)} &= \frac{1}{m_1 m_2 m_3} \sum_{l \in J_1} q_1^{(k-1)}.
\end{align*}
\]

In the E-step we compute the new probabilities based on (25)-(28) as

\[
q_1^{(k), EM} = \frac{\beta^{(k)} \varphi(x_1, \delta_1^{(k)})}{\beta^{(k)} \varphi(x_1, \delta_1^{(k)}) + (1 - \beta^{(k)}) \varphi(x_1, \delta_2^{(k)})}, \ l \in J_1.
\]

In the SEM-part we act in a different way. In the S-step we generate independent Bernoulli-distributed random variables \( y_1^{(k)} \in \{0, 1\} \) with probabilities \( P(y_1^{(k)} = 1) = q_1^{(k-1)}, l \in J_1. \)

During the M-Step we get \( \nu_1^{(k)} = \sum_{l \in J_1} y_l^{(k)} \) and the estimates \( \delta_1^{(k)} = (\mu_1^{(k)}, \Sigma_1^{(k)}), \delta_2^{(k)} = (\mu_2^{(k)}, \Sigma_2^{(k)}), \beta^{(k)} \) by

\[
\begin{align*}
\mu_1^{(k)} &= \frac{\sum_{l \in J_1} y_1^{(k)} = 1 x_l}{\nu_1^{(k)}}, \\
\Sigma_1^{(k)} &= \frac{\sum_{l \in J_1} y_1^{(k)} = 1 (x_l - \mu_1^{(k)})(x_l - \mu_1^{(k)})^T}{\nu_1^{(k)}}, \\
\mu_2^{(k)} &= \frac{\sum_{l \in J_2} y_1^{(k)} = 0 x_l}{\nu_2^{(k)}}, \\
\Sigma_2^{(k)} &= \frac{\sum_{l \in J_2} y_1^{(k)} = 0 (x_l - \mu_2^{(k)})(x_l - \mu_2^{(k)})^T}{\nu_2^{(k)}}, \\
\beta^{(k)} &= \frac{\nu_1^{(k)}}{m_1 m_2 m_3} = \frac{1}{m_1 m_2 m_3} \sum_{l \in J_1} y_l^{(k)}.
\end{align*}
\]
In the E-Step, we compute the updated probabilities \( q^{(k,SE)}_l \) based on \( q^{(k,EM)}_l \), \( \lambda_k \), \( \beta \), \( \alpha \), \( \Omega \), \( \Omega_0 \), and \( \Omega_1 \) by relation (30)-(32).

The essential idea of the SAEM algorithm is to mix \( q^{(k,EM)}_l \) and \( q^{(k,SE)}_l \) in iteration step \( k \) as

\[
q^{(k)}_l = \lambda_k q^{(k,SE)}_l + (1 - \lambda_k) q^{(k,EM)}_l, \quad l \in J_W, 
\]

which gives the a priori probabilities for the next \( (k + 1) \)-th iteration. Here in (33), \( \{\lambda_k, k \geq 1\} \) is a sequence of positive real numbers \( \lambda_k \in (0, 1) \) decreasing to zero. We stop the SAEM algorithm after the \( k \)-th iteration if \( \sum_{l \in J_W} |q^{(k-1)}_l - q^{(k)}_l| \leq \varepsilon \), where \( \varepsilon \) is some threshold. In the following, the choice of parameters is a result of experimental tuning to our image data yielding good practical results. Particularly, we use \( \lambda_k = \frac{50}{2k+1}, k \geq 1 \) and \( \varepsilon = 0.0001 \) in our computations.

When the SAEM algorithm stops in the \( k_0 \)-th iteration, we obtain the values \( \{q^{(k_0)}_l, l \in J_W\} \) which indicate that \( x_l \) belongs to the first component if \( q^{(k_0)}_l > 1/2 \) and to the second one, otherwise.

Applying the above SAEM algorithm to our image data yields diffuseness in the resulting clusters (see Figure 2). To avoid this, we propose a smoothing modification (Spatial SAEM), which takes the spatial location of the sample data into account. Let us describe the new Spatial step.

Let SAEM stop after \( k_0 \) iterations. For each sample entry \( x_l = (l_1, l_2, l_3) \in J_W \) we define the coordinate \( y_l = ((l_1 - 1)M\Delta, (l_2 - 1)M\Delta, (l_3 - 1)M\Delta) \), that is a vertex of the cube \( W_l \). In each further iteration, i.e. for \( k > k_0 \), Bernoulli random variables \( y^{(k)_l} \in \{0, 1\}, l \in J_W \) with success probability \( q^{(k_0)}_l \), \( l \in J_W \) are simulated. Now \( y^{(k)_l} \) classifies \( x_l \) in such a way that \( y^{(k)_l} = 0 \) indicates that \( x_l \) belongs to the first component and the second one, otherwise. Then, we compute the number \( a^{(k)}_r \) of neighbors of \( x_l \) belonging to the same cluster as \( x_l \). Neighborhood is defined in terms of the \( r \)-neighborhood of \( v_l \) such that

\[
a^{(k)}_r = \sum_{l \in J_W, l \neq 1} \mathbb{1}(y^{(k)}_l = y^{(k)}_1, \|v_l - v_1\|_\infty \leq r),
\]

\( r > 0 \).

If all \( a^{(k)}_r \) are greater than or equal to a certain threshold \( a \) (for our image data, \( a = 3 \) is used) we call the classification \( y^{(k)}_l, l \in J_W \), admissible and move to the next iteration. Otherwise, for sample entries \( x_l \) with \( a^{(k)}_r \) being less than \( a \), we change \( y^{(k)}_l \), hence, the class of \( x_l \). If the new set \( y^{(k)}_l, l \in J_W \), is admissible, then we pass to iteration \( k + 1 \). If no, we resimulate \( y^{(k)}_l, l \in J_W \), until \( a^{(k)}_r \geq a \) for all \( l \in J_W \).

The smoothing procedure stops when \( K \) admissible classifications have been generated (\( K = 1000 \) is used). The final a posteriori probabilities \( q \) in the space of all admissible classifications are computed over the sample of \( \{y^{(k)_l}, l \in J_W\}, k_0 \leq k \leq k_0 + K \) by

\[
q_l = \frac{1}{K} \sum_{k=k_0}^{k_0+K} y^{(k)_l}, \quad l \in J_W. 
\]

We also get estimates of components’ weights \( \hat{\beta}, 1 - \hat{\beta} \) by (28). If \( \hat{\beta} \geq 0.5 \) we say that the second component corresponds to the “anomaly”. The observation...
window $W_1$ thus belongs to the zone of homogeneous material if $q_l \geq 0.5$ and to the anomaly zone if $q_l < 0.5$. If $\beta < 0.5$ the roles of the components are swapped.

6 Application to 3D image data of fibre materials

![Image of simulated layered RSA fibre data, 2000 × 2000 × 2100 voxels]

**Fig. 3** Visualization of simulated layered RSA fibre data, 2000 × 2000 × 2100 voxels

6.1 Simulated data

First, we illustrate the use of the methods from Sections 4 and 5 on simulated 3D fibre images. We choose a random sequential absorption (RSA) model that randomly adds fibres to the existing material, such that they do not intersect each other, cf. [3]. Figure 3 shows simulated RSA fibre data in an image of 2000 × 2000 × 2100 voxels. The sample exhibits three layers, where fibres differ in their local directional distribution. Each layer has a thickness of 700 voxels and contains 82474 fibres with a constant radius of 4 voxels and length of 100 voxels. Fibre directions are distributed according to a special case of the Angular Central Gaussian distribution described in [25]. In the two outer layers, the preferred direction is the $x$-direction and the concentration parameter is $\beta = 0.1$ resulting in a high concentration of the fibres along the main direction. In the middle layer, which is considered the anomaly region, the preferred direction is the $y$-direction and the fibres are less concentrated ($\beta = 0.5$).
Detecting anomalies in fibre systems using 3-dimensional image data

Additionally, we investigate a homogeneous RSA data set where no anomalies should be detected. The data set consists of an image of $2000 \times 2000 \times 2100$ voxels. Here, the concentration parameter of the fibre direction distribution is $\beta = 0.1$ in the whole sample. The preferred direction is the $x$-direction. The fibre radius is 4 voxels, the fibre length is 100 voxels. A visualisation of a realisation of this model is shown in Figure 4.

Now we apply the change-point analysis of Section 4 to random fields of mean local directions and entropy estimates for the homogeneous and layered RSA data.

To do so, we transform the data of average local directions $X_i, i \in J$. In order to avoid cancelling effect of averaging, we build for each coordinate $x, y, z$, the samples $\tilde{x}_k, \tilde{y}_k, \tilde{z}_k$, such that their entries lie in the hemispheres $S^2_x = \{(x, y, z) \in S, x \geq 0\}, S^2_y = \{(x, y, z) \in S, y \geq 0\}, S^2_z = \{(x, y, z) \in S, z \geq 0\}$, respectively, i.e., $\tilde{x}_k = |x|_k, \tilde{y}_k = |y|_k, \tilde{z}_k = |z|_k$.

The sample of the estimated entropy values $\hat{E}_k$ of directional distribution of fibres $X_i, i \in J$ in the windows $W_k, k \in J_W$ is build by estimator (7) over the transformed directions $\tilde{X}_i \in S^2_+, i \in J$.

We apply the results of Section 4 consequently to the random fields $s_k = \tilde{x}_k, \tilde{y}_k, \tilde{z}_k$, and $\hat{E}_k$. Therefore, we have the following 4 pairs of hypotheses of $(H_0^0, H_1^0)$-type.

- $H_0^0 : \mathbb{E}\tilde{x}_k = \mu_x$ for every $k \in W$ vs.
- $H_1^0 : \exists \theta_0 \in \Theta_0$ such that $\mathbb{E}\tilde{x}_k = \mu_x + h_x, k \in I_{\theta_0}$ and $\mathbb{E}\tilde{x}_k = \mu_x, k \in I_{\theta_0^c}$, $h_x \neq 0$;
- $H_0^0 : \mathbb{E}\tilde{y}_k = \mu_y$ for every $k \in W$ vs.

**Fig. 4** Visualization of simulated homogeneous RSA fibre data, $2000 \times 2000 \times 2100$ voxels
Since we test only 4 hypotheses simultaneously, we stick to the classical Bonferroni method, e.g. we test each direction and entropy separately with significance level $\frac{1}{4}\alpha$.

Before running the algorithms we need to choose the right size of scanning windows. From the initial layered and homogeneous RSA images with $2000 \times 2000 \times 2100$ voxels we obtain $83 \times 83 \times 87$ small windows $W_i$ with $24 \times 24 \times 24$ voxels each, and 463537 and 460559 nonempty entries, respectively.

Due to the model parameters (fibre length of 100 voxels corresponds to 5 points in $W$), we can assume the random field $X_k$ to be $m$-dependent with $m = 5$ and $\sigma^2 = 0.2, M_0 = \frac{1}{2}$. For mean local directions $\bar{x}, \bar{y}, \bar{z}$, the parametric set $\Theta_0$ is constructed with $\Delta_0 = \Delta_1 = 8, \gamma_0 = 0.05, \gamma_1 = 0.5$, and $L_M = 22$ in [22], $|\Theta_0| = 39395$.

We point out that the samples $(\bar{x}, \bar{y}, \bar{z})$ and $\bar{E}$ have different sizes due to the construction described in Section 3. Therefore, the parameters $m, \sigma^2$ and the parameter set $\Theta_0$ in [22] for $\bar{E}$ differ from the ones for $\bar{x}, \bar{y}, \bar{z}$.

For the sample of estimated entropy $\bar{E}_k, k \in J_W$, we have $m = 1, \sigma^2 = 0.5$ and the parametric set $\Theta_0$ is constructed with $\Delta_0 = \Delta_1 = 2, \gamma_0 = 0.05, \gamma_1 = 0.5$, and $L_M = 4$ in [22], $|\Theta_0| = 16536$. Entropy values approximately have a normal distribution, so we put $M_0 = \sigma$ in [19].

The computed statistics (given in [11]) $T_W(\bar{x}), T_W(\bar{y}), T_W(\bar{z}), T_W(\bar{E})$ and corresponding $p$-values from relation [19] are presented in Table 4 for the homogeneous and in Table 5 for the layered RSA data. Thus, there is no evidence to reject $H^0_\theta, H^0_\mu, H^0_\sigma, H^L_\theta$ in the homogeneous case. The described test allows to claim that there is an anomaly region in the layered RSA image data, because we reject $H^0_\theta, H^0_\mu$, and $H^L_\sigma$, but have no evidence to reject $H^L_\theta$.

| Attribute | Sample var. | Test statistics | $p$-value |
|-----------|------------|----------------|----------|
| $\bar{x}$ | 0.04360 | 0.0344 | 1.00 |
| $\bar{y}$ | 0.03749 | 0.0146 | 1.00 |
| $\bar{z}$ | 0.08984 | 0.0942 | 1.00 |

Table 4 Change-point test for mean local directions of homogeneous RSA data.

Therefore, the choice of the mean of local directions attribute for the change-point analysis in our problem with layered fibre image data is reasonable. Moreover, depending on the data (e.g. containing whirlpools of fibres) it may be better to choose entropy or other attributes to test for other types of anomalies.

It follows from [42, Theorem 2.4.] that the Dobrushin estimator of the entropy of i.i.d. vectors on a $C^1$--smooth manifold is asymptotically Gaussian. Although the RSA fibre data do not satisfy the i.i.d. assumption of mutual independence
Detecting anomalies in fibre systems using 3-dimensional image data

| Attribute | Sample var. | Test statistics | p-value |
|-----------|-------------|-----------------|---------|
| $\tilde{x}$ | 0.10592 | 0.44036 | $4.6 \times 10^{-30}$ |
| $\tilde{y}$ | 0.10948 | 0.43163 | $2.8 \times 10^{-23}$ |
| $\tilde{z}$ | 0.06151 | 0.18764 | 0.301 |
| $\tilde{E}$ | 0.3583 | 1.07030 | 0.00 |

Table 5 Change-point test for mean local directions of layered RSA data.

of fibre locations and directions, the estimated local directional entropy $\tilde{E}$ for the homogeneous data seems to have a unimodal distribution, see Figure 6. Assuming that the Gaussian distribution provides a reasonable approximation also in this case, we apply the $3\sigma$-rule with $\sigma^2$ being the sample variance of $\tilde{E}$, compare [1], to find anomaly regions in both Figures 3 and 4.

![Fig. 5](image)

Fig. 5 Histogram of frequencies of the local entropy of fibre directions and two separated Gaussian probability density functions found by the spatial SAEM algorithm. Layered RSA fibre data, $\sigma = 0.5986$.

![Fig. 6](image)

Fig. 6 Histogram of frequencies of the local entropy of fibre directions fitted Gaussian probability density. Homogeneous RSA fibre data, $\sigma = 0.2997$.
One can see that all centered entropy values lie in the interval $[-3\sigma; 3\sigma]$, so the $3\sigma$-method does not distinguish between homogeneous (Figure 4) and inhomogeneous (Figure 3) images. We conclude that the $3\sigma$-rule for anomaly detection does not work well if the anomaly regions are large enough to produce histograms of the clustering attribute with many modes.

The fact that the distribution of $\hat{E}$ seems to have two modes might indicate that it is a mixture of two Gaussian distributions. So we apply the Spatial SAEM algorithm from Section 5.1 to separate these modes. By an empirical study, a scanning window $W$ consisting of $5 \times 5 \times 5$ small windows $\tilde{W}_k$ was selected, i.e. $M = 5$. Additionally, we put $r = \Delta M$ in (34) by default.

First, let us consider clustering based on the attribute entropy. In the layered data, the Spatial SAEM algorithm finds two clusters and determines the distributional parameters for them. The clustering results are presented in Figures 7 and 8. Green labels denote the centers of scanning windows corresponding to the homogeneous fibre material and blue labels mark objects belonging to the anomalous region. As expected, the Spatial SAEM algorithm found no anomaly for homogeneous RSA fibre data.

![Fig. 7 Anomaly detection in layered RSA fibre data using the local entropy: SAEM algorithm.](image)

We also ran the Spatial SAEM algorithm with the attributes mean of local direction (MLD) and a vector combining entropy and MLD. The results for the layered data are presented in Figures 9 and 10. One can see that combination of both attributes gives a more reliable result.
Remark 4 The problem of clustering a fibre material into homogeneity and anomaly zones using vector-valued cluster attributes can be solved by a variety of other clustering methods, see the books \[23,30,49\] for an overview. In addition to the results reported here, we tried the recent AWC algorithm \[22\]. However, the spatial SAEM approach yields better results, cf. \[21\]. Moreover, it does not require a complex parameter tuning and operates fast.

We also tried to use a principal axis of fibre directions as a classification attribute in the described SAEM algorithm. But the results are worse than the ones for the MLD attribute. This effect can be explained by the fact that the distribution of principal axes has its support on a unit sphere and thus cannot be a mixture of Gaussian distributions in \(\mathbb{R}^3\). Therefore, the estimates in the M-step \(25-27\), \(30\), \(31\) have to be modified, cf. \[25\]. This task goes, however, beyond the scope of the present article.

6.2 Real glass fibre reinforced polymer

Now we apply our anomaly detection approach to a 3D-image of a glass fibre reinforced polymer. The images are provided by the Institute for Composite Materials (IVW) in Kaiserslautern, see Figure \[11\]. For a detailed description of the material we refer to \[50\].
We apply the change point analysis from Section 4 to real data with $970 \times 1469 \times 1217$ voxels and the estimated radius of 3 voxels. We obtain $64 \times 97 \times 80$ small windows $\tilde{W}_i$ with $15 \times 15 \times 15$ voxels.

For mean local directions, the parametric set $\Theta_0$ is constructed with $\Delta_0 = \Delta_1 = 8$, $\gamma_0 = 0.05$, $\gamma_1 = 0.5$ and $L_M = 22$ in (22), which gives $|\Theta_0| = 33004$. To choose the suitable value of $m$ for $m-$dependence we need additional investigation.

Under the hypotheses $H_x^g, H_y^g, H_z^g$ the random fields $\tilde{x}, \tilde{y},$ and $\tilde{z}$ are assumed to be stationary, so that we can estimate their covariance functions. We use the standard approach and estimate e.g. $\rho_x(h) = \text{Cov}(\tilde{x}_1, \tilde{x}_{1+h})$ as

$$\hat{\rho}(h) = \frac{1}{|K| - 1} \sum_{k \in K} (\tilde{x}_k - \bar{x}_0) (\tilde{x}_{k+h} - \bar{x}_h),$$

where $K = \{(k_1, k_2, k_3) \in \mathbb{N}^3, 1 \leq k_1 \leq M_1 - h_1, 1 \leq k_2 \leq M_2 - h_2, 1 \leq k_3 \leq M_3 - h_3\}$, and $\bar{x}_0$ and $\bar{x}_h$ are the sample means of $\tilde{x}$ over the index ranges $K$ and $K + h$ respectively. To visualize $\hat{\rho}_x$ we compute its maximum values in the following way:

$$\hat{\rho}_{x, \text{max}}(i) = \max_{1 \leq k_1, k_2, k_3 \leq i} (\hat{\rho}_x(i, k_2, k_3), \hat{\rho}_x(k_1, i, k_3), \hat{\rho}_x(k_1, k_2, i)), i \geq 1.$$

The values of $\hat{\rho}_{x, \text{max}}$ (black color), $\hat{\rho}_{y, \text{max}}$ (red color), $\hat{\rho}_{z, \text{max}}$ (blue color) are given in Figure 13. We choose the value of $m$ in such a way that $\hat{\rho}_{x, \text{max}}(i) \leq \varepsilon_0$,.
Detecting anomalies in fibre systems using 3-dimensional image data

Fig. 10 Anomaly detection with Spatial SAEM algorithm in the layered RSA fibre data using local entropy and mean of local directions attributes.

\[ \hat{\rho}_{y, \text{max}}(i) \leq \varepsilon_0, \hat{\rho}_{z, \text{max}}(i) \leq \varepsilon_0, \] for all \( i \geq m \), where \( \varepsilon_0 \) is a threshold. Here we use \( \varepsilon_0 = 0.04 \) and obtain \( m = 9 \). We have \( M_0 = 0.5 \) and assume that \( \sigma^2 = 0.2 \).

Moreover, due to simulation experiments in Section 4, we compute critical values for the change-point statistics \( T_W(\cdot) \) and \( p \)-values from inequality (19) with \( m = 7 \).

For the random field of estimated local entropies we have \( m = 1 \) and the parametric set \( \Theta_0 \) is constructed with \( \Delta_0 = \Delta_1 = 2, \gamma_0 = 0.05, \gamma_1 = 0.5 \) and \( L_M = 4 \) in (22), which gives \( |\Theta_0| = 12366 \). We assume that \( \sigma^2 = 0.5 \) and \( M_0 = \sigma \). The result of our change point analysis is presented in Table 6. Our change point

| Attribute | Sample var. | Test statistics | \( p \)-value |
|-----------|-------------|-----------------|--------------|
| \( \hat{x} \) | 0.04589 | 0.15995 | 1.00 |
| \( \hat{y} \) | 0.06795 | 0.44733 | \( 2.1 \times 10^{-10} \) |
| \( \hat{z} \) | 0.07982 | 0.43383 | \( 1.3 \times 10^{-6} \) |
| \( \hat{E} \) | 0.30126 | 0.46811 | \( 3.96 \times 10^{-8} \) |

Table 6 Change-point test for mean local directions of real data.

test detects the evidence of anomaly regions in real fibre data at significance level \( \alpha = 8.4 \times 10^{-10} \).

Similarly to the case of RSA data, the detection of anomalies by the 3\( \sigma \)–rule gives meaningless results. The Spatial SAEM algorithm works much better. Its
results are presented in Figures 14, 15, and 16, where the color labelling of points is the same as for the simulated data.

We conclude that the Spatial SAEM algorithm with attributes entropy, MLD, and a combination of both produces adequate results.
Detecting anomalies in fibre systems using 3-dimensional image data

Fig. 13 Values of empirical covariance $\hat{\rho}_{x,max}, \hat{\rho}_{y,max}, \hat{\rho}_{z,max}$.

Fig. 14 Anomaly detection in the fibre image (Fig. 11) using local entropy.

Since the images in Figure 15 and 16 look very similar at first glance, we investigate the results of the Spatial SAEM anomaly detection in more detail. We separate a part of the 3D image (Figure 12) into 9 layers and present the result of clustering for the 1st and 5th layer for the entropy in Figure 17, for MLD in Figure 18, and for the combination of both in Figure 19.

One can observe that the Spatial SAEM algorithm with local entropy attribute detects vortices of fibers in the material as anomaly regions. This is natural since a vortex exhibits a large diversity of fibre directions, and the entropy is a measure of such diversity. The Spatial SAEM anomaly detection using the mean of local fibre directions identifies the central part of the image (cf. Figure 18) as an anomaly region, where the directions of fibres differ from the average throughout the image. Finally, the Spatial SAEM approach using both clustering attributes identifies both vortices of fibres and layers of fibres with principally different main direction, cf. Figure 19.
Fig. 15 Anomaly detection in the fibre image (Fig. 11) using mean of local direction.

Fig. 16 Anomaly detection in the fibre image (Fig. 11) using local entropy and mean of local direction.

Acknowledgements We are grateful to Dr. Stefanie Schwaar from Fraunhofer ITWM for valuable discussions and to Jan Niedermeyer for simulating RSA data.

References

1. P. Alonso-Ruiz and E. Spodarev. Entropy-based inhomogeneity detection in fiber materials. *Methodology and Computing in Applied Probability*, Nov. 2017. [https://doi.org/](https://doi.org/)
Detecting anomalies in fibre systems using 3-dimensional image data

Fig. 17 Spatial SAEM clustering according to entropy

Fig. 18 Spatial SAEM clustering according to mean of local directions

2. P. Alonso-Ruiz and E. Spodarev. Estimation of entropy for Poisson marked point processes. *Adv. in Appl. Probab.*, 49(1):258–278, 2017.

3. H. Andrä, M. Gurka, M. Kabel, S. Nissle, C. Redenbach, K. Schladitz, and O. Wirjadi. Geometric and mechanical modeling of fiber-reinforced composites. In *Proceedings of the 2nd International Congress on 3D Materials Science*, pages 35–40. Springer, 2014.

4. M. Basseville and I. Nikiforov. Detection of abrupt changes: theory and application. Prentice Hall Information and System Sciences Series. Prentice Hall, Inc., Eaglewood Cliffs, NJ, 1993.

5. J. Beirlant, E. J. Dudewicz, L. Györfi, and E. C. van der Meulen. Nonparametric entropy estimation: an overview. *Int. J. Math. Stat. Sci.*, 6(1):17–39, 1997.

6. B. E. Brodsky and B. S. Darkhovsky. *Nonparametric methods in change-point problems*, volume 243 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1993.

7. B. E. Brodsky and B. S. Darkhovsky. Problems and methods of probabilistic diagnostics. *Avtomat. i Telemekh.*, 60(8):3–30, 1999.

8. B. E. Brodsky and B. S. Darkhovsky. *Non-parametric statistical diagnosis*, volume 509 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 2000. Problems and methods.

9. B. Bucchia. Testing for epidemic changes in the mean of a multiparameter stochastic process. *J. Statist. Plann. Inference*, 150:124–141, 2014.

10. B. Bucchia and C. Heuser. Long-run variance estimation for spatial data under change-point alternatives. *J. Statist. Plann. Inference*, 165:104–126, 2015.
Fig. 19 Spatial SAEM clustering according to a combination of entropy and MLD

11. B. Bucchia and M. Wendler. Change-point detection and bootstrap for Hilbert space valued random fields. *Journal of Multivariate Analysis*, 155:344 – 368, 2017.

12. A. Bulinski and D. Dimitrov. Statistical estimation of the Shannon entropy. *Acta. Math. Sin.-English Ser.*, 2018. [https://doi.org/10.1007/s10114-018-7440-z](https://doi.org/10.1007/s10114-018-7440-z).

13. J. Cao and K. J. Worsley. The detection of local shape changes via the geometry of Hotelling’s $T^2$ fields. *Ann. Statist.*, 27(3):925–942, 1999.

14. E. Carlstein, H.-G. Müller, and D. Siegmund, editors. *Change-point problems*, volume 23 of *Institute of Mathematical Statistics Lecture Notes—Monograph Series*. Institute of Mathematical Statistics, Hayward, CA, 1994. Papers from the AMS-IMS-SIAM Summer Research Conference held at Mt. Holyoke College, South Hadley, MA, July 11–16, 1992.

15. G. Celeux and J. Diebolt. A stochastic approximation type EM algorithm for the mixture problem. *Stochastics Stochastics Rep.*, 41(1-2):119–134, 1992.

16. A. Chambaz. Detecting abrupt changes in random fields. *ESAIM Probab. Statist.*, 6:189–209, 2002. New directions in time series analysis (Luminy, 2001).

17. J. Chen and A. K. Gupta. *Parametric statistical change point analysis*. Birkhäuser/Springer, New York, second edition, 2012. With applications to genetics, medicine, and finance.

18. S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic geometry and its applications*. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester, third edition, 2013.

19. M. Csörgő and L. Horváth. *Limit theorems in change-point analysis*. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester, 1997.

20. R. L. Dobrushin. A simplified method of experimentally evaluating the entropy of a stationary sequence. *Theory of Probability & Its Applications*, 3(4):428–430, 1958.

21. D. Dresvyanskiy, T. Karaseva, S. Mitrofanov, C. Redenbach, S. Schwaar, V. Makogin, and E. Spodarev. Application of clustering methods to anomaly detection in fibrous media. arXiv preprint arXiv:1810.12401 [stat.AP], 2018.

22. K. Efimov, L. Adamyan, and V. Spokoiny. Adaptive nonparametric clustering. arXiv preprint arXiv:1709.09102, 2017.

23. B. S. Everitt, S. Landau, M. Leese, and D. Stahl. *Cluster analysis*. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester, fifth edition, 2011.

24. K. Falconer. *Fractal geometry*. John Wiley & Sons, Inc., Hoboken, NJ, second edition, 2003. Mathematical foundations and applications.

25. J. Franke, C. Redenbach, and N. Zhang. On a mixture model for directional data on the sphere. *Stand. J. Stat.*, 43(1):139–155, 2016.

26. Fraunhofer ITWM, Department of Image Processing. MAVI – modular algorithms for volume images. [http://www.mavi-3d.de](http://www.mavi-3d.de) 2005.

27. A. K. Gorshein, V. Y. Korolev, and A. M. Tursunbaev. Median modifications of the EM-algorithm for separation of mixtures of probability distributions and their applications to the decomposition of volatility of financial indexes. *J. Math. Sci. (N.Y.)*, 227(2):176–195, 2017.
28. T. Hahubia and R. Mnatsakanov. On the mode-change problem for random measures. *Georgian Math. J.*, 3(4):343–362, 1996.
29. L. Heinrich. Some bounds of cumulants of $m$-dependent random fields. *Math. Nachr.*, 149:303–317, 1990.
30. C. Hennig, M. Meila, F. Murtagh, and R. Rocci, editors. *Handbook of cluster analysis*. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. CRC Press, Boca Raton, FL, 2016.
31. D. Jarusková and V. I. Piterbarg. Log-likelihood ratio test for detecting transient change. *Statist. Probab. Lett.*, 81(5):552–559, 2011.
32. E. I. Kaplan. On the change-point problem for random fields. *Teor. Veroyatnost. i Primenen.*, 35(2):353–358, 1990.
33. E. I. Kaplan. Convergence of estimates for partitions in the change point problem for random fields. *Teor. Imovir. ta Mat. Statist.*, 47:34–39, 1992.
34. V. Y. Korolev. EM-algorithm, its modifications and their use in the problem of decomposing the mixtures of probability distributions. IPiRAS, Moscow, 2007.
35. L. F. Kozachenko and N. N. Leonenko. A statistical estimate for the entropy of a random vector. *Problemy Peredachi Informatsii*, 23(2):9–16, 1987.
36. T. L. Lai. Saddlepoint approximations and boundary crossing probabilities for random fields and their applications. In *Third International Congress of Chinese Mathematicians. Part 1, 2*, volume 2 of AMS/IP Stud. Adv. Math., 42, pt. 1, pages 29–39. Amer. Math. Soc., Providence, RI, 2008.
37. B. Laurent, C. Marteau, and C. Maugis-Rabusseau. Multidimensional two-component Gaussian mixtures detection. *Ann. Inst. Henri Poincaré Probab. Stat.*, 54(3):842–865, 2018.
38. W. D. Miller. *Quasi-Hefting algebras: A new class of lattices, and a foundation for nonclassical model theory with possible computational applications*. ProQuest LLC, Ann Arbor, MI, 1993. Thesis (Ph.D.)–Kansas State University.
39. H. G. Müller and K.-S. Song. Cube splitting in multidimensional edge estimation. In *Change-point problems (South Hadley, MA, 1992)*, volume 23 of IMS Lecture Notes Monogr. Ser., pages 210–223. Inst. Math. Statist., Hayward, CA, 1994.
40. Y. Ninomiya. Construction of conservative test for change-point problem in two-dimensional random fields. *J. Multivariate Anal.*, 89(2):219–242, 2004.
41. E. S. Page. Continuous inspection schemes. *Biometrika*, 41:100–115, 1954.
42. M. D. Penrose and J. E. Yukich. Limit theory for point processes in manifolds. *Ann. Appl. Probab.*, 23(6):2161–2211, 2013.
43. C. Redenbach and I. Vecchio. Statistical analysis and stochastic modelling of fibre composites. *Composites Science and Technology*, 71:107–112, 2011.
44. C. A. Rogers. *Hausdorff measures*. Cambridge University Press, 1998.
45. A. Sen and M. S. Srivastava. On tests for detecting change in mean. *Ann. Statist.*, 3:98–108, 1975.
46. O. Sharipov, J. Tewes, and M. Wendler. Sequential block bootstrap in a Hilbert space with application to change point analysis. *Canad. J. Statist.*, 44(3):300–322, 2016.
47. D. Siegmund and B. Yakir. Detecting the emergence of a signal in a noisy image. *Stat. Interface*, 1(1):3–12, 2008.
48. D. O. Siegmund and K. J. Worsley. Testing for a signal with unknown location and scale in a stationary Gaussian random field. *Ann. Statist.*, 23(2):608–639, 1995.
49. S. T. Wierzchon and M. A. Klopotek. *Modern algorithms of cluster analysis*, volume 34 of Studies in Big Data. Springer, Cham, 2018.
50. O. Wirjadi, M. Godedhardt, K. Schladitz, B. Wagner, A. Rack, M. Gurka, S. Nissele, and A. Noll. Characterization of multilayer structures of fiber reinforced polymer employing synchrotron and laboratory X-ray CT. *International Journal of Materials Research*, 105(7):649–654, 2014.
51. O. Wirjadi, K. Schladitz, P. Easwaran, and J. Olser. Estimating fibre direction distributions of reinforced composites from tomographic images. *Image Analysis & Stereology*, 35(3):167–179, 2016.
52. Y. Wu. *Inference for change-point and post-change means after a CUSUM test*, volume 180 of Lecture Notes in Statistics. Springer, New York, 2005.