Stochastic 3D Modeling of Fiber–Based Materials

Gerd Gaiselmann[a,*], Ralf Thiedmann[a], Ingo Manke[b], Werner Lehnert[c], Volker Schmidt[a]

[a] *Institute of Stochastics, Ulm University, 89069 Ulm, Germany*
[b] *Institute of Applied Materials, Helmholtz Centre Berlin for Materials and Energy (HZB), 14109 Berlin, Germany*
[c] *Institute of Energy and Climate Research (IEK-3: Fuel Cells), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany*

Abstract

A stochastic multi–layer model is developed describing the microstructure of materials which are built up of strongly curved, but almost horizontally oriented fibers. This fully parametrized model is based on ideas from stochastic geometry and multivariate time series analysis. It consists of independent layers which are stacked together, where each single layer is described by a 2D germ–grain model dilated in 3D. The germs form a Poisson point process and the grains are given by random polygonal tracks describing single fibers in terms of multivariate time series. Exemplarily, on the basis of 2D data from SEM images, the parameters of the multi–layer model are fitted to the microstructure of a non–woven material which is used for gas–diffusion layers in PEM fuel cells. Therefore, an algorithm is presented which automatically extracts typical fiber courses from SEM images. Finally, the multi–layer model is validated by comparing structural characteristics computed for 3D data gained by synchrotron tomography from the same material, and for realizations drawn from the multi–layer model.

Keywords: porous material, curved fibers, non–woven, germ–grain model, multivariate time series, vectorial autoregression, SEM, synchrotron tomography

1. Introduction

Fiber–based materials find increasing interest in various applications, e.g. in thermal insulations, the housings of aircraft and vehicles and not to forget in fuel cell technology. The morphology of these materials has a deep impact on their macroscopic properties, therefore detailed structural investigations are needed. The aim of the present paper is the development of a structural simulation model and its exemplary application to gas diffusion layers (GDL) being

*Corresponding author: phone: +49 731 50 23590, fax: +49 731 50 23649, e-mail: gerd.gaiselmann@uni-ulm.de

a key component of polymer electrolyte membrane (PEM) fuel cells. The GDL material is built up of carbon fibers, which are almost horizontally oriented. In particular, the specific example which we consider in the present paper is the non–woven GDL H2315 produced by the company Freudenberg FCCT KG. It consists of strongly curved fibers.

Note that the microstructure of the GDL has a significant influence on the transportation of gases and water along the percolation pathways through the pores. Moreover, the carbon fibers manage the transport of electrons. For further details concerning the functionality of GDL in fuel cells we refer e.g. to [7, 12, 14].

In order to improve the understanding of these physical processes, we develop a stochastic microstructure model which is fully parameterized. It can be used for virtual scenario analyses with the general aim to construct microstructures with improved physical properties. By systematic modifications of model parameters and in combination with numerical transportation models, virtual 3D morphologies and virtual transportation processes in these morphologies can be simulated. In this way, the identification of optimized morphologies is possible.

We propose a multi–layer approach to describe horizontally oriented fiber systems, i.e., our 3D model consists of several (stochastically) independent layers stacked on top of each other. Each single layer is given by a so–called germ–grain model which is the set union of curved fibers running in horizontal direction where, for simplicity, the fibers are allowed to interpenetrate mutually. This single–layer model is constructed by combining tools of stochastic geometry and multivariate time series analysis (see e.g. [9, 20] and [6, 13], respectively), where we consider a germ–grain model such that the germs form a homogeneous Poisson point process in 2D and the grains are random planar polygons which are spherically dilated in 3D.

The resulting multi–layer model, i.e. the stack of modeled single layers, is able to adequately reproduce the 3D microstructure of fiber systems with strongly curved fibers running mainly in horizontal direction. Moreover, the multi–layer approach enables us to fit model parameters on the basis of structural information extracted from 2D images, where we use a similar algorithm like in [21] for the extraction of single fibers from the top layer.

Note that for industrial applications it is important to gain the necessary structural information of a material as much as possible from 2D images since they are relatively easy to obtain in contrast to 3D image data.

The imaging technique, used in this paper for model fitting, is scanning-electron microscopy (SEM) which delivers 2D images for various kinds of materials. From the 2D SEM images the typical tracks of fibers located near the top of the material can be detected. These fibers, represented as 2D polygonal tracks, are the main data basis in order to construct a stochastic model for single fibers. In particular, the 2D polygonal tracks are modeled by sequences of random vectors, where these tracks extracted from the SEM images are considered as realizations of multivariate time series. Finally, for model validation, we use 3D data gained by synchrotron tomography, where structural characteristics of real data and simulated data sampled from the fitted model are compared to

each other.

Recently, several approaches to stochastic modeling of fiber–based materials have been studied in the literature, see e.g. [1, 10, 16, 17, 19, 21]. In most of these papers materials consisting of straight fibers are considered, whereas in [1] a stochastic model is developed which can describe materials with curved fibers. However, in distinction to the modeling approach proposed in the present paper, 3D data are needed in [1] for model fitting.

Our paper is organized as follows. Section 2 describes the detection of fibers located on top of 2D SEM images. In Section 3, the stochastic model for single fibers is introduced using tools of multivariate time series analysis. Section 4 establishes the multi–layer model as a stack of independent copies of a single–layer model which is given in terms of a germ–grain model from stochastic geometry. Section 5 summarizes the results and provides a short outlook regarding possible future research.

2. Detection of Typical Fiber Courses

We first state an algorithm to detect the typical courses of fibers from SEM images, see Figure 1 (left). These fiber courses, represented as polygonal tracks, will be the main data basis for fitting the single–fiber model introduced in Section 3. Note that our algorithm is similar to a procedure which has been developed in [21] for fiber–based materials with straight fibers.

2.1. Preprocessing of SEM Images

A closer look at the 2D SEM image of Figure 1 (left) indicates that the observed data contains noise that shall be removed in the first step of image processing. A global smoothing using e.g. a Gaussian filter [3, 11, 15] has the disadvantage that also the edges indicating the fibers get blurred. To avoid this undesirable effect, the method of anisotropic diffusion is used.

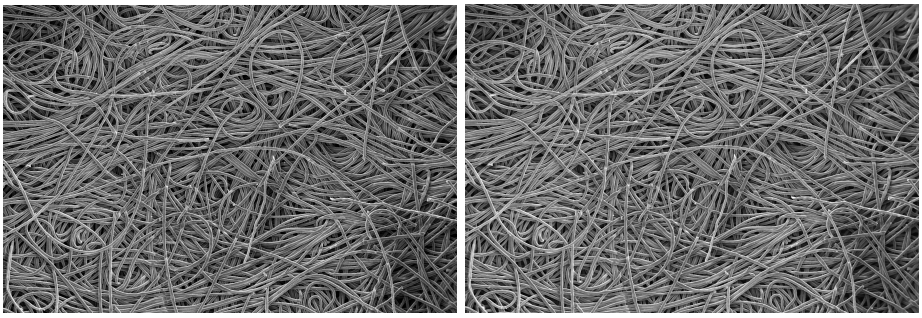

Figure 1: Left: 2D SEM image of non–woven GDL, right: Smoothed SEM image using anisotropic diffusion.

2.1.1. Smoothing, Linear Filtering, and Binarization

The principle idea of anisotropic diffusion is quite simple. First, the edges to be preserved are detected, and then the smoothing is not performed across them, see e.g. [18] for details. The advantage of this smoothing technique is that the loss of structural information is minimized, i.e., the edges remain sharp but noise is reduced, see Figure 1 (right).

In the next step, the edges of fibers are further emphasized by convoluting the smoothed image with a linear filter [3], see Figure 2 (left).

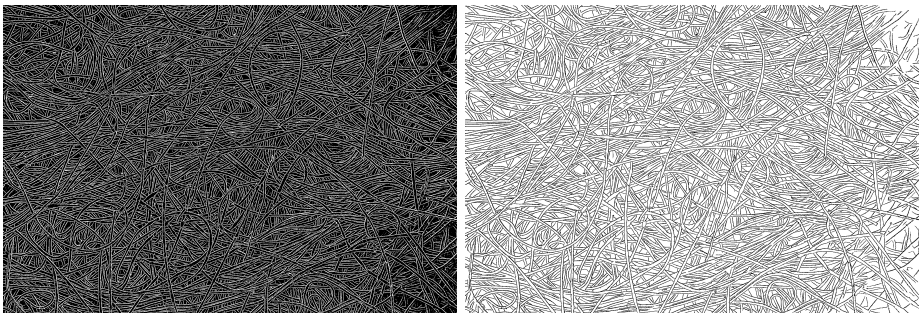

Figure 2: Left: 2D SEM image after convolution for edge enhancement, right: SEM image after binarization.

The smoothed and convoluted image is subsequently thresholded where we apply the isodata algorithm [3] to determine the threshold value non-interactively, see Figure 2 (right). Notice that in order to enhance the clarity of presentation, the images beginning from the right part of Figure 2 are inverted, i.e., black and white pixels are interchanged.

The binarized image contains small foreground clusters which obviously do not contribute to those fibers located on top of the SEM image. To remove them, an algorithm for the detection of isolated clusters [8] is used and small clusters with a size below a certain threshold are removed where this threshold has been set to 50 pixel. Furthermore, the connectivity of the remaining foreground phase is improved by dilation using a small disc as structuring element [18].

2.1.2. Skeletonization; Transformation into Vector Data

The next processing step is the skeletonization of the binarized SEM image. This means that foreground pixel, i.e., those belonging to the objects we are interested in, are changed to background pixel in a way that the remaining (still pixel–given) paths have a thickness of one pixel, where the connectivity has to be preserved [3, 18]. The result of skeletonizing the binary image on the right–hand side of Figure 2 is shown in Figure 3 (left).

In order to transform the skeletonized image into vector data, the foreground pixel are subdivided into three groups, where the number of neighboring foreground pixel is counted with respect to the 8–neighborhood. A pixel with exactly one neighbor is considered as endpoint, a pixel with exactly two neighbors is

classified as linepoint, and all other pixel are interpreted as crosspoints. Then, all pixels classified as crosspoints are removed, see Figure 4. This seems to be particularly reasonable for the detection of curved fibers, since in Section 2.2 an algorithm is introduced which appropriately connects parts of the fiber courses to be detected. Subsequently, the remaining set of foreground pixel is transformed into polygonal tracks which consist of relatively short line segments.

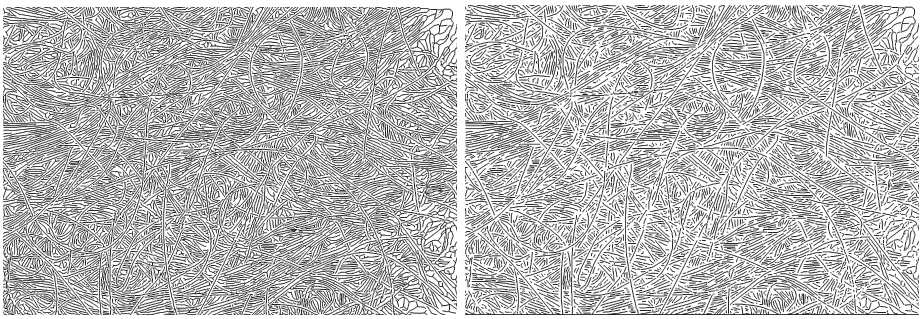

Figure 3: Left: Skeletonized image, right: Representation by vector data

Figure 4: Left: Skeletonized data; right: Removement of crosspoints

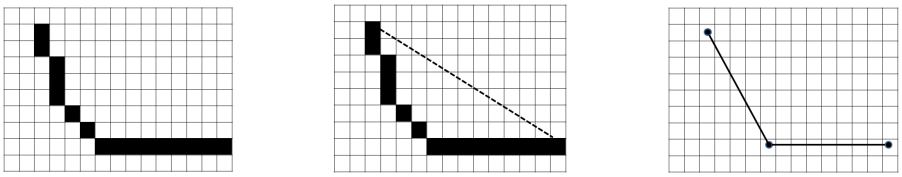

Figure 5: Transformation of pixel–given paths into polygonal tracks

The transformation of the connected clusters of foreground pixel into polygonal tracks is done in the following way. For each pixel cluster, starting from one of its endpoints, the (Euclidean) line segment between the current pixel and the endpoint is compared to the corresponding pixel–given path. If the discrepancy between the pixel–given path and the line segment is too large, i.e., if the

pixel-given path is too curved, the previous pixel is said to be the endpoint of the current line segment and, simultaneously, the starting point of the next line segment, see Figure 5. This procedure is repeated until the other endpoint of the pixel cluster is reached.

The final result of this transformation is a family of polygonal tracks, which can be interpreted as a graph structure, representing the pixel–given structure of the skeletonized SEM image, see Figure 3.

2.2. Reconstruction of Fibers

In this section we discuss a stochastic algorithm to connect the polygonal tracks, whose extraction from SEM images has been described in Section 2.1, in such a way that the curvature properties of the fibers on top of the considered SEM image are represented correctly.

2.2.1. Connection of Polygonal Tracks

The family of polygonal tracks $p_1, \ldots, p_n$, $n \geq 1$, obtained in Section 2.1.2 from the skeletonized SEM image, contains relatively short parts of the fiber courses to be detected, see Figure 1 and Figure 3 (right). Hence, for an adequate reconstruction of the fibers, the polygonal tracks considered in Section 2.1.2 have to be appropriately connected.

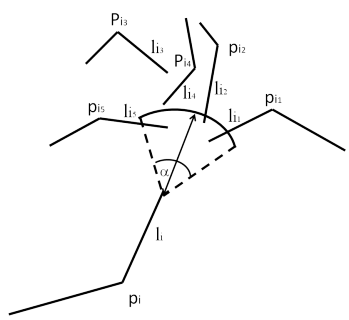

Figure 6: Connection of polygonal tracks

The basic idea of our algorithm is to look for sequences of polygonal tracks representing the courses of single fibers. More precisely, let p_i and p_j be two candidates of polygonal tracks to be possibly connected via their 'end–segments' l_i and l_j. Then we connect the polygonal tracks p_j and p_i only if the connecting line segment l_{ij} has nearly the same direction as l_i and l_j and if it is not too long. If so, the union $p_i \cup l_{ij} \cup p_j$ of segments is considered as one single polygonal track. At the end point of this new polygonal track, a next polygonal track is searched in the same way as before. This procedure is repeated until no further pairs of polygonal tracks can be found to be connected.

Since, for a given polygonal track, there can be several possibilities of connecting other polygonal tracks with it, a decision rule has to be established which

chooses the next polygonal track. In the following, some criteria are described which are used for this selection.

For connecting a polygonal track p_i to another one, we are looking for all end–segments $l_{i_1}, \ldots, l_{i_m} \in \{l_1, \ldots, l_n\} \setminus \{l_i\}$ with an endpoint belonging to a certain sector around the considered endpoint of l_i, see Figure 6. In case of the exemplary material (for non–woven GDL) studied in the present paper, this sector has been chosen with a length $|l_{max}|$ of 85 pixel and an angle $\alpha = \sphericalangle_{max}$ of 0.9 because these values provide very good visual results. This means that a connection of the line segment l_i to one of the line segments $l_{i_1}, \ldots, l_{i_m}$ is possible. Notice that the line segment l_i is only connected to exactly one line segment contained in the set $l_{i_1}, \ldots, l_{i_m}$. The probability that a polygonal track p_i is connected to a polygonal track $p_j \in \{p_{i_1}, \ldots, p_{i_m}\}$ via a line segment l_{ij} depends on

1) the angle between the direction of the line segment l_{ij} and the directions of the end–segments l_i and l_j, where the probability that p_i and p_j get connected increases if the differences between the directions of l_i, l_j and l_{ij} get smaller, and
2) the length $|l_{ij}|$ of the line segment l_{ij}, where the probability is smaller for longer line segments l_{ij}.

Thus, a weight ω_{ij} is considered for each line segment l_{ij} possibly connecting p_i and p_j, which is given by

$$\omega_{ij} = \exp\left(-\frac{1}{2}\left(\frac{|l_{ij}|}{|l_{max}|}\right)^2\right) \cdot \exp\left(-\frac{1}{2}\left(\frac{\sphericalangle_{l_{ij}}}{\sphericalangle_{max}}\right)^2\right),$$

where $|l_{ij}|$ denotes the length of the line segment l_{ij} connecting l_i and l_j, and $|l_{max}|$ is the maximum possible length of a connecting line segment. By $\sphericalangle_{l_{ij}}$ we denote the mean of the absolute values of the angles $\sphericalangle(l_i, l_{ij})$ and $\sphericalangle(l_{ij}, l_j)$, i.e., the differences of directions between the line segments l_i and l_{ij} and the line segments l_{ij} and l_j, whereas $\sphericalangle_{max}$ is the maximum possible difference of directions in accordance with the definition of the selection sector introduced above. Note that $|l_{max}|$ and $\sphericalangle_{max}$ are the parameters of our random connection algorithm.

After having determined the weights ω_{ij}, they are normalized such that the sum of the normalized weights $\widetilde{\omega}_{ij}$ equals 1. The polygonal track p_i is then connected with p_{i_j} with probability $\widetilde{\omega}_{ii_j}$.

2.2.2. Evaluation of Connected Tracks

Since polygonal tracks are connected according to a random selection rule, it is obvious that the result of one single run cannot be assured to be optimal. To find nearly optimal solutions, the simulation of connections, i.e., the random construction of elongated polygonal tracks, is repeated several times until results of similar good quality are found for a certain number of subsequent runs.

The evaluation is accomplished by means of the following criterion of goodness which is based on the smoothed 2D SEM image, given as gray value image,

i.e., each pixel of the image is assigned with a gray value between 0 and 255. For the detection of curved fibers as occurring in non–woven GDL, an empirical variance of gray values of polygonal tracks is considered, where each (elongated) polygonal track p is discretized and subsequently dilated by the unit disk $b(o,1) \subset \mathbb{R}^2$. Let $I(u_1,v_1),\ldots,I(u_{k(p)},v_{k(p)})$ be the gray values of the smoothed SEM image at the pixel $(u_1,v_1),\ldots,(u_{k(p)},v_{k(p)})$ of the dilated and discretized polygonal track p. The empirical variance of those pixel is computed only for whose gray values which are darker than the average gray value

$$\overline{I}_p = \frac{1}{k(p)} \sum_{i=1}^{k(p)} I(u_i,v_i)\,.$$

More precisely, for the polygonal track p, we consider the modified set of gray values $\widetilde{I}(u_1,v_1),\ldots,\widetilde{I}(u_{k(p)},v_{k(p)})$ given by $\widetilde{I}(u_i,v_i) = \min\{I(u_i,v_i),\overline{I}_p\}$ for $i = 1,\ldots,k(p)$, which are used to determine the empirical 'dark variance'

$$\widehat{\mathrm{Var}}_d(p) = \frac{1}{k(p)-1} \sum_{i=1}^{k(p)} \left(\widetilde{I}(u_i,v_i) - \overline{I}_p\right)^2 . \tag{1}$$

The idea of the dark variance introduced above is to look at locations, where the considered polygonal track crosses darker regions which are assumed to belong to the background. Hence, the smaller the dark variance $\widehat{\mathrm{Var}}_d(p)$ of p, the better the polygonal track p is assumed to be.

The above described criterion is used to identify the best track p for any initial (not yet elongated) polygonal track $p_i \in \{p_1,\ldots,p_n\}$ within the set of all (randomly) elongated polygonal tracks starting from p_i, i.e., the same random connection algorithm is performed for all initial tracks $\{p_1,\ldots,p_n\}$. The resulting best tracks are denoted by $p'_1,\ldots,p'_\ell$.

2.2.3. Ranking of Tracks

In order to identify the fibers on top of the considered material, a ranking of the best polygonal tracks $p'_1,\ldots,p'_\ell$ is introduced which is based on the gray values of the smoothed SEM image along the courses of $p'_1,\ldots,p'_\ell$. The idea for such a ranking is to indicate how close a polygonal track is to the surface. Hence, two characterizing 'rank factors' ω_i^1, $\omega_i^2 \in [0,1]$ are considered for each polygonal track p'_i; $i = 1,\ldots,\ell$.

The first factor ω_i^1 is related with the so–called 'overlay criterion'. This criterion is based on the assumption that fibers close to the surface lead to longer polygonal tracks compared to fibers in the background since the latter ones are overlaid more frequently by other fibers which are closer to the top. Thus, the rank factor ω_i^1 for p'_i is given by

$$\omega_i^1 = \frac{\text{coverage length of } p'_i}{\text{total length of } p'_i}\,,$$

where the coverage length of p'_i denotes the total length of those line segments of p'_i which coincide with line segments from the skeletonized image. The second

rank factor ω_i^2 is a normalized version of the empirical dark variance $\widehat{\mathrm{Var}}_d(p_i')$ of p_i' introduced in (1), where smaller variances are assumed to belong to fibers closer to the top.

Then, each polygonal track p_i' is equipped with a rank value rk_i which is given by $rk_i = (1-\omega_i^1)\cdot\omega_i^2$. This puts us in the position to rank the tracks $p_1', \ldots, p_\ell'$ according to their rank values, i.e., $rk_{\pi(1)} \leq \ldots \leq rk_{\pi(\ell)}$, where $(\pi(1), \ldots, \pi(\ell))$ is some permutation of $(1, \ldots, \ell)$.

From the manufacturer of the non–woven GDL considered in this paper, it is known that its porosity, i.e., the volume fraction of the pore phase, equals 75%. Hence, we can use this information in order to determine the set of fibers on top of the GDL, where the polygonal tracks from $p_{\pi(1)}', \ldots, p_{\pi(\ell)}'$ are successively added to the top layer according to their rankings, i.e., at first, we add $p_{\pi(1)}'$ then $p_{\pi(2)}'$, etc. This is repeated until the required volume fraction of 25% within the top layer (or the last track $p_{\pi(\ell)}'$) is reached. To determine the volume fraction of the current set of polygonal tracks, we dilate each polygonal track using the disk $b(o, 4.75)$ as structuring element which has a radius of 4.75 μm. The specific choice of this structuring element is motivated by the knowledge of the fiber diameter which is 9.5 μm.

Additionally, it is checked whether a higher ranked track covers the currently considered track p_j' too extensive. Therefore, the number of common line segments is counted and if it exceeds a certain threshold, in our case the level 5 has been chosen, the lower ranked track is rejected and the next polygonal track is considered. An exemplary output of our extraction algorithm is shown in Figure 7.

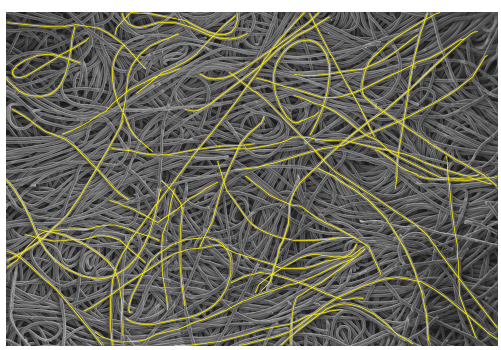

Figure 7: Set of fibers detected on top of the GDL

3. Modeling of Single Fibers

In this section we introduce a two–stage approach in order to model the typical courses of fibers, whose extraction from SEM images has been discussed

in Section 2. Both stages of our single–fiber model are based on vectorial autoregressive processes. First, a time series is considered which governs the decomposition of the random fiber into an alternating sequence of straight and curved parts of fibers, respectively. Then, these parts are separately modeled by means of vectorial autoregressive processes. Notice that all fibers have a fixed Euclidean length.

3.1. Model Description

We consider three different multivariate time series and denote them by $\{R_i,\, i \geq 1\}$, $\{C_i,\, i \geq 1\}$, and $\{S_i,\, i \geq 1\}$, respectively. The random sequence $\{R_i,\, i \geq 1\}$, which is constructed by means of a two–dimensional autoregressive process, describes the transitions from curved to straight parts of the random fiber, and vice versa. It distinguishes between left–curved and right–curved parts.

More precisely, the random vectors $R_i = (R_{i1}, R_{i2})$ take their values on the discrete grid $\{-1, 0, 1\} \times \{1, 2 \ldots\}$, where $R_{i1} = -1$ stands for a right–curved part of the fiber, $R_{i1} = 0$ indicates a straight part, and $R_{i1} = 1$ a left–curved part. Note that each left–curved part can be represented as a right–curved part if the sequence of its segments is considered in reverse order. The second component R_{i2} of R_i denotes the number of line segments of the i-th part of the fiber.

The times series $\{C_i,\, i \geq 1\}$ and $\{S_i,\, i \geq 1\}$ are also constructed by means of two–dimensional autoregressive processes. They describe the curved and straight parts of fibers themselves.

For each $n \geq 1$, let $\{C_i^{(n)},\, i \geq 0\}$ and $\{S_i^{(n)},\, i \geq 0\}$ be independent copies of $\{C_i,\, i \geq 1\}$ and $\{S_i,\, i \geq 1\}$, respectively, which are independent of $\{R_i,\, i \geq 1\}$. Furthermore, let $\{X_i,\, i \geq 1\}$ be the multivariate time series which describes the (complete) random fiber consisting of an alternating sequence of curved and straight parts. It is given in the following way. Suppose that $i \in \left(\sum_{j=1}^{k-1} R_{j2}, \sum_{j=1}^{k} R_{j2}\right]$ for some $k \geq 1$. Then, we put

$$X_i = \begin{cases} C^{(k)}_{i-\sum_{j=1}^{k-1} R_{j2}} & \text{if } R_{k1} = 1, \\ S^{(k)}_{i-\sum_{j=1}^{k-1} R_{j2}} & \text{if } R_{k1} = 0, \\ C^{(k)}_{\sum_{j=1}^{k} R_{j2}-(i-1)} & \text{if } R_{k1} = -1. \end{cases} \tag{2}$$

We assume that the time series $\{R_i,\, i \geq 1\}$, $\{C_i,\, i \geq 1\}$, and $\{S_i,\, i \geq 1\}$ are stationary and, therefore, can be extended to stationary random sequences $\{R_i,\, i \in \mathbb{Z}\}$, $\{C_i,\, i \in \mathbb{Z}\}$, and $\{S_i,\, i \in \mathbb{Z}\}$ indexed by the set $\mathbb{Z} = \{\ldots, -1, 0, 1, \ldots\}$ of all integers. Furthermore, we assume that these time series can be represented by means of some stationary autoregressive processes $\{Y_i^{(1)},\, i \in \mathbb{Z}\}$, $\{Y_i^{(2)},\, i \in \mathbb{Z}\}$, and $\{Y_i^{(3)},\, i \in \mathbb{Z}\}$ such that

$$R_i = \Phi(Y_i^{(1)}), \quad C_i = \Psi(Y_i^{(2)}), \quad S_i = \Psi(Y_i^{(3)}) \quad \text{for each } i \in \mathbb{Z}, \tag{3}$$

where the functions $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ and $\Psi : \mathbb{R}^2 \to \mathbb{R}^2$ are given by

$$\Phi(s,t) = \begin{cases} (-1, \max\{q+1, \lfloor t \rfloor\}) & \text{if } s < -0.5, \\ (0, \max\{q+1, \lfloor t \rfloor\}) & \text{if } -0.5 \le s \le 0.5, \\ (1, \max\{q+1, \lfloor t \rfloor\}) & \text{if } s > 0.5, \end{cases} \tag{4}$$

where $q > 0$ is some integer and

$$\Psi(s,t) = (s - 2k\pi, \max\{0, t\}) \qquad \text{if } (2k-1)\pi \le s < (2k+1)\pi \tag{5}$$

for some $k \in \mathbb{Z}$, and $\lfloor t \rfloor \in \mathbb{Z}$ denotes the integer closest to t, i.e. $-0.5 \le \lfloor t \rfloor - t \le 0.5$. The specific form of Φ and Ψ is related with the incremental representation of polygonal tracks introduced in Section 3.2.1 below.

Thus, an important building stone of our stochastic modeling approach to strongly curved, but almost horizontally oriented fibers, are stationary autoregressive processes $\{Y_i,\, i \in \mathbb{Z}\}$ of some order $q \ge 0$, where

$$Y_i = \mu + A_1\, Y_{i-1} + \ldots + A_q\, Y_{i-q} + \varepsilon_i \qquad \text{for each } i \in \mathbb{Z}. \tag{6}$$

Note that $q > 0$, $\mu \in \mathbb{R}^2$, and the coefficient matrices $A_1, \ldots, A_q \in \mathbb{R}^{2\times 2}$ are considered as model parameters. The 'residuals' $\{\varepsilon_i\,, i \ge 1\}$ are assumed to form a sequence of two–dimensional random vectors which are independent and identically distributed with vanishing mean vector $\mathbb{E}\,\varepsilon_i = o$ and some (non–singular) covariance matrix $\Sigma = \mathbb{E}\,(\varepsilon_i \varepsilon_i^\top)$ where Σ is a further model parameter, see e.g. [6, 13].

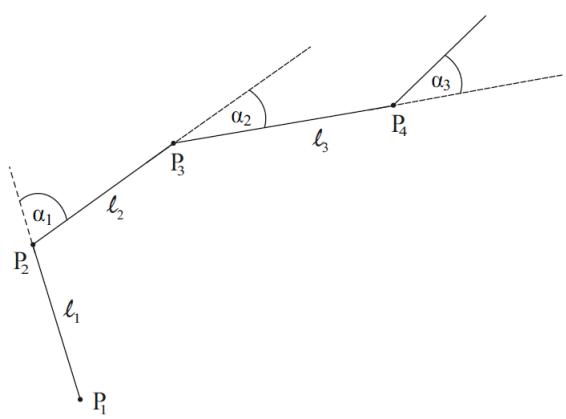

Figure 8: Incremental representation of polygonal tracks

3.2. Model Fitting

3.2.1. Incremental Representation of Polygonal Tracks

In order to fit the single–fiber model $\{X_i,\, i \ge 1\}$ introduced in Section 3.1 to track data extracted from SEM images, we consider the following incremental

representation of polygonal tracks. Instead of describing a polygonal track $p = (p_1, \ldots, p_n)$ by the endpoints $a_i, b_i \in \mathbb{R}^2$ of its line segments p_i like in Section 2, we now consider an angle–length representation, where we assume that the first segment has one endpoint at the origin and some direction $\alpha_0 \in [0, 2\pi]$ (with respect to the x–axis). The further course of the track p can then be described by the lengths $l_1, l_2, \ldots$ of the consecutive line segments and the angles $\alpha_1, \alpha_2, \ldots$, where α_i denotes the change of direction from the i-th to the $(i+1)$-th segment. Thus, altogether, polygonal tracks beginning at the origin can be characterized by the initial angle α_0 and the angle–length increments $(\alpha_1, l_1)^\top, (\alpha_2, l_2)^\top, \ldots$, see Figure 8.

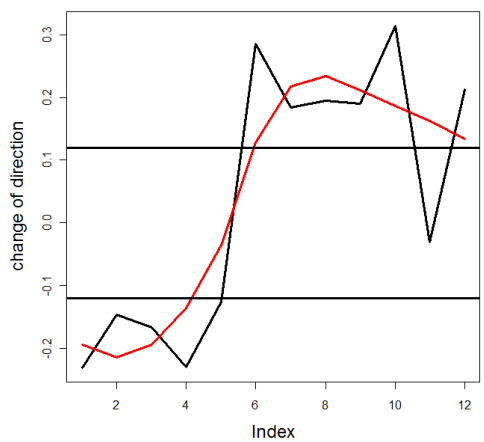

Figure 9: Changes of direction before (black) and after (red) smoothing; the horizontal black lines indicate the thresholds ρ and $-\rho$

Our next goal is to decompose the polygonal tracks into curved and straight parts. The decision at which positions the polygonal track is split depends on the angles $\alpha_1, \alpha_2, \ldots$. If there is a sequence $\alpha_i, \ldots, \alpha_{i+k}$ of consecutive angles with the same algebraic sign which are significantly different from zero, then the corresponding line segments are considered to form a curved part, otherwise a straight part. Since outliers and variability of the α_i can influence the splitting in an inappropriate way, a local smoothing is applied to the sequence $\alpha_1, \alpha_2, \ldots$, leading to a smoothed sequence $\widetilde{\alpha}_1, \widetilde{\alpha}_2, \ldots$ of angles, see e.g. [2, 4]. If $k \geq 3$ and $\widetilde{\alpha}_i, \ldots, \widetilde{\alpha}_{i+k} > \rho$ for a certain threshold $\rho > 0$, then the corresponding sequence of line segments is considered to form a left–curved part of the polygonal track. Vice versa, $\widetilde{\alpha}_i, \ldots, \widetilde{\alpha}_{i+k} < -\rho$ indicates a right–curved part, and $-\rho \leq \widetilde{\alpha}_i, \ldots, \widetilde{\alpha}_{i+k} \leq \rho$ a straight part, see Figure 9.

Visual comparison of decompositions obtained for different values of ρ have shown that putting $\rho = 0.12$ provides very good results for the polygonal tracks which have been extracted in Section 2 from the SEM image of the GDL material considered in the present paper, see Figure 10.

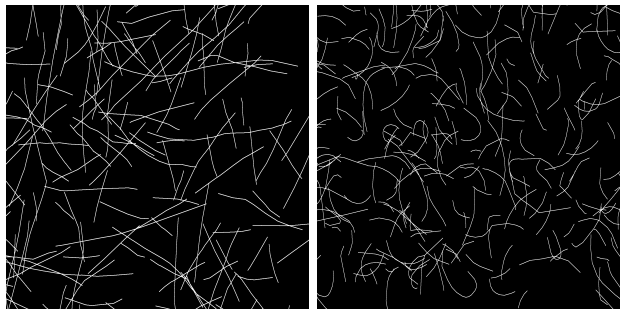

Figure 10: Straight (left) and curved (right) parts of the polygonal tracks

3.2.2. Estimation of Parameters

We now explain how the parameters of the single–fiber model $\{X_i,\, i \geq 1\}$ can be fitted to the angle–length data described in Section 3.2.1. This model fitting will be performed separately for each of the stationary autoregressive processes $\{Y_i^{(1)},\, i \in \mathbb{Z}\}$, $\{Y_i^{(2)},\, i \in \mathbb{Z}\}$, and $\{Y_i^{(3)},\, i \in \mathbb{Z}\}$ introduced in Section 3.1, i.e., we will determine three specifications of the parameters $q > 0$, $\mu \in \mathbb{R}^2$, $A_1, \ldots, A_q \in \mathbb{R}^{2\times 2}$, and $\Sigma \in \mathbb{R}^{2\times 2}$ of the autoregressive model $\{Y_i,\, i \in \mathbb{Z}\}$ given in (6).

As data basis, we use a large number of (short) trajectories of the time series $\{R_i,\, i \in \mathbb{Z}\}$, $\{C_i,\, i \in \mathbb{Z}\}$, and $\{S_i,\, i \in \mathbb{Z}\}$, respectively, obtained from the incremental representation and splitting of polygonal tracks described in Section 3.2.1. Actually, to get trajectories of the autoregressive processes $\{Y_i^{(1)},\, i \in \mathbb{Z}\}$, $\{Y_i^{(2)},\, i \in \mathbb{Z}\}$, and $\{Y_i^{(3)},\, i \in \mathbb{Z}\}$, these data should be transformed by some random functions $\widetilde{\Phi}, \widetilde{\Psi} : \Omega \times \mathbb{R}^2 \to \mathbb{R}^2$, which in a sense are 'inverse' to Φ and Ψ introduced in Section 3.1. However, for simplicity, we will consider the (non–transformed) data directly as trajectories of $\{Y_i^{(1)},\, i \in \mathbb{Z}\}$, $\{Y_i^{(2)},\, i \in \mathbb{Z}\}$, and $\{Y_i^{(3)},\, i \in \mathbb{Z}\}$. Thus, in all three cases, we assume that a large number k of (short) trajectories $\{y_{11}, \ldots, y_{1n_1}\}$, ..., $\{y_{k1}, \ldots, y_{kn_k}\}$ is given which are independently sampled from a stationary autoregressive process $\{Y_i,\, i \in \mathbb{Z}\}$, where $n_i \geq 3$ for each $i = 1, \ldots, k$.

Since a considerable fraction of the curved and straight parts of polygonal tracks which have been determined by the method stated in Section 3.2.1 just consists of three or four segments, we put $q = 2$ in all three cases.

In order to fit the remaining model parameters, we consider estimators of maximum–likelihood type where we assume for simplicity that the residuals ε_i in (6) have a normal distribution with vanishing mean vector and some (non–singular) covariance matrix Σ, i.e., $\varepsilon_i \sim \mathrm{N}(o, \Sigma)$. This implies that also the $2n$–dimensional random vector $(Y_1, \ldots, Y_n)$ given in (6) is normally distributed for each $n \geq 1$. In particular, both components of Y_n are normally distributed random variables.

Statistical analysis shows that the normal distribution is not perfectly reflected by our data (see Figure 11, first row). It is well known that the applica-

tion of a Box–Cox transformation [4] can improve this situation, see Figure 11 (second row). However, because the fitting procedure via maximum–likelihood estimation stated below is quite robust with respect to deviations from normality, we decided to apply this method of model fitting using the original (non–transformed) data.

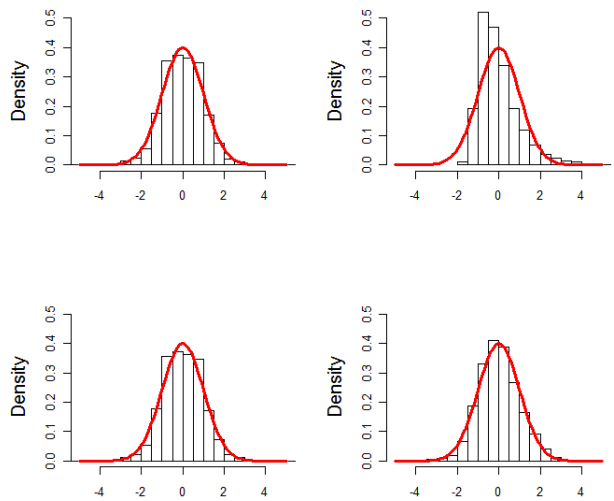

Figure 11: Standardized empirical densities of changes of direction (left) and segment lengths (right) for straight parts before (first row) and after (second row) Box–Cox transformation; plotted with the N(0,1)–density (red line)

It is useful to introduce some further notation. Let $a_{ij}^{(\ell)}$ be the (i,j)-th entry of the 2×2 matrix A_ℓ, i.e., $A_\ell = (a_{ij}^{(\ell)})$ for $\ell = 1, 2$. Furthermore, let $a^* = (a_{11}^{(1)}, a_{21}^{(1)}, a_{12}^{(1)}, a_{22}^{(1)}, a_{11}^{(2)}, a_{21}^{(2)}, a_{12}^{(2)}, a_{22}^{(2)})^\top$ denote the (column) vector consisting of the columns of the 2×2 matrices A_1 and A_2, i.e., the vector a^* is just another (equivalent) representation of the 2×4 matrix (A_1, A_2).

We preliminarily assume that the expectation $\eta = \mathbb{E}Y_i$ is known. In order to fit the parameter $\theta = (a^*, \Sigma)$ of the centered autoregressive process $\{Z_i,\, i \in \mathbb{Z}\}$ with $Z_i = Y_i - \eta$, where $Z_i = A_1\, Z_{i-1} + A_2\, Z_{i-2} + \varepsilon_i$ for each $i \in \mathbb{Z}$, to the centered data $\{z_{11}, \ldots, z_{1n_1}\}, \ldots, \{z_{k1}, \ldots, z_{kn_k}\}$ with $z_{ij} = y_{ij} - \eta$, we consider the likelihood function

$$L(z_{11}, \ldots, z_{kn_k}; \theta) = \prod_{i=1}^{k} f_{Z_3,\ldots,Z_{n_i}}(z_{i3}, \ldots, z_{in_i}; \theta)\,,$$

where $f_{Z_3,\ldots,Z_{n_i}}$ denotes the joint density of the random vector $(Z_3, \ldots, Z_{n_i})$ assuming that Z_1 and Z_2 are given fixed numbers, i.e., $Z_1 = z_{i1}$ and $Z_2 = z_{i2}$.

Then, the solution $\widetilde{\theta} = (\widetilde{a}^*, \widetilde{\Sigma})$ of the likelihood equation

$$\frac{d}{d\theta} L(z_{11}, \ldots, z_{1n_1}, \ldots, z_{k1}, \ldots, z_{kn_k}; \theta) = 0$$

is given by (see e.g. [6, 13])

$$\begin{aligned}
\widetilde{a}^* &= \left(\sum_{i=1}^{k} \widetilde{B}_i \widetilde{B}_i^\top \right)^{-1} \left(\sum_{i=1}^{k} \left(\widetilde{B}_i \otimes I_2 \right) \left((y_{i3} - \eta)^\top, \ldots, (y_{in_i} - \eta)^\top \right)^\top \right) , \\
\widetilde{\Sigma} &= \left(\sum_{i=1}^{k} n_i \right)^{-1} \sum_{i=1}^{k} \left(\widetilde{B}_i^0 - (\widetilde{A}_1, \widetilde{A}_2) \widetilde{B}_i \right) \left(\widetilde{B}_i^0 - (\widetilde{A}_1, \widetilde{A}_2) \widetilde{B}_i \right)^\top ,
\end{aligned}$$

where $\widetilde{B}_i = \left(\widetilde{B}_{i3}, \ldots, \widetilde{B}_{in_i} \right)$ and $\widetilde{B}_{ij} = \left((y_{i,j-1} - \eta)^\top, (y_{i,j-2} - \eta)^\top \right)^\top$, $\widetilde{B}_i^0 = \left(y_{i3} - \eta, \ldots, y_{in_i} - \eta \right)$, and $\widetilde{A}_\ell = (\widetilde{a}_{ij}^{(\ell)})$. Furthermore, $\otimes$ denotes the Kronecker product and I_2 the 2×2 unit matrix.

By $\widehat{A}_1, \widehat{A}_2$, and $\widehat{\Sigma}$ we denote the sample functions corresponding to $\widetilde{A}_1, \widetilde{A}_2$, and $\widetilde{\Sigma}$, where η is replaced by the sample mean

$$\widehat{\eta} = \frac{1}{\sum_{ti=1}^{k} n_i} \sum_{i=1}^{k} \sum_{j=1}^{n_i} y_{ij} , \tag{7}$$

everywhere in the above definitions of $\widetilde{A}_1, \widetilde{A}_2$, and $\widetilde{\Sigma}$. Finally, we put $\widehat{\mu} = (I_2 - \widehat{A}_1 - \widehat{A}_2)^{-1} \widehat{\eta}$. Then, $\widehat{\mu} \in \mathbb{R}^2$, $\widehat{A}_1, \widehat{A}_2 \in \mathbb{R}^{2\times 2}$, and $\widehat{\Sigma} \in \mathbb{R}^{2\times 2}$ are the parameters of the fitted autoregressive model $\{Y_i,\, i \in \mathbb{Z}\}$ of order $q = 2$.

In particular, for the autoregressive process $\{Y_i^{(1)},\, i \in \mathbb{Z}\}$ governing the transitions between curved and straight parts we get that

$$Y_i^{(1)} = \begin{pmatrix} -0.032 \\ 2.14 \end{pmatrix} + \begin{pmatrix} -0.105 & 0.001 \\ -0.005 & 0.042 \end{pmatrix} Y_{i-1}^{(1)} + \begin{pmatrix} -0.278 & -0.0004 \\ 0.121 & 0.105 \end{pmatrix} Y_{i-2}^{(1)} + \varepsilon_i ,$$

where $\varepsilon_i \sim \mathrm{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.325 & -0.039 \\ -0.039 & 2.251 \end{pmatrix} \right)$. The process $\{Y_i^{(2)},\, i \in \mathbb{Z}\}$ corresponding to the curved parts is given by

$$Y_i^{(2)} = \begin{pmatrix} 0.334 \\ 36.55 \end{pmatrix} + \begin{pmatrix} 0.126 & -0.001 \\ -2.346 & 0.193 \end{pmatrix} Y_{i-1}^{(2)} + \begin{pmatrix} -0.067 & -0.0003 \\ 6.78 & 0.081 \end{pmatrix} Y_{i-2}^{(2)} + \varepsilon_i ,$$

where $\varepsilon_i \sim \mathrm{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.018 & -0.658 \\ -0.658 & 494.5 \end{pmatrix} \right)$, and the process $\{Y_i^{(3)},\, i \in \mathbb{Z}\}$ corresponding to the straight parts by

$$Y_i^{(3)} = \begin{pmatrix} 0.005 \\ 64.38 \end{pmatrix} + \begin{pmatrix} -0.19 & -0.00002 \\ 10.346 & 0.098 \end{pmatrix} Y_{i-1}^{(3)} + \begin{pmatrix} -0.137 & -0.00004 \\ -1.568 & 0.064 \end{pmatrix} Y_{i-2}^{(3)} + \varepsilon_i ,$$

where $\varepsilon_i \sim \mathrm{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.01 & -0.03 \\ -0.03 & 1185 \end{pmatrix} \right)$.

The estimated parameters of the single fiber model can be interpreted as follows. The process means of $Y_i^{(1)}, Y_i^{(2)}, Y_i^{(2)}$, calculated by means of (7), are given by $\eta^{(1)} = (-0.01, 2.51)^\top$, $\eta^{(2)} = (0.26, 51.94)^\top$ and $\eta^{(3)} = (0.0003, 76.84)^\top$. It can be clearly seen that the mean value of the direction change is significantly larger for process $\{Y_i^{(2)}\}$ (curved parts) than for process $\{Y_i^{(1)}\}$ (straight parts). This is mainly caused by the larger first component of the corresponding intercept vector. Moreover, the mean length of line segments of the polygonal tracks representing the straight parts of the fibers is larger than the corresponding value of the curved parts.

The mean value of the first component for process $\{Y_i^{(1)}\}$ (transition process) is almost zero. Thus, we can conclude that left–curved parts and right–curved parts appear in the same frequency.

3.3. Model Validation

In this section the goodness–of–fit of the single–fiber model $\{X_i,\, i \in \mathbb{Z}\}$ introduced in (2) is investigated. In particular, we analyze how good the curvature of the polygonal tracks extracted from the SEM image is reflected by the fitted single–fiber model. Therefore, a curvature measure is defined which is then computed for each (real) polygonal track extracted from the SEM image and for simulated polygonal tracks drawn from the single–fiber model. In addition, the same procedure of model validation is performed for the models $\{C_i,\, i \geq 1\}$ and $\{S_i,\, i \geq\ 1\}$ introduced in (3) which describe the curved and straight parts of the polygonal tracks.

3.3.1. Simulation of Single Fibers

To simulate a single fiber with a given Euclidean length l, where in our case $l = 50\,000\,\mu m$ (see Section 4.2 below), we first generate a sufficiently long trajectory of the autoregressive process $\{Y_i^{(1)},\, i \in \mathbb{Z}\}$ governing the transitions between curved and straight parts. It turns out that trajectories $(y_{m+1}, \ldots, y_{m+n})$ of length $n = 200$ are long enough to achieve this goal, where we consider a so–called 'burn–in period' [13] of length $m = 150$ in order to capture the assumed stationarity of $\{Y_i^{(1)},\, i \in \mathbb{Z}\}$ sufficiently well.

Thus, using the estimated model parameters $\widehat{\mu}$, $\widehat{A}_1, \widehat{A}_2$, and $\widehat{\Sigma}$ which have been obtained in Section 3.2.2 for $\{Y_i^{(1)},\, i \in \mathbb{Z}\}$, we first simulate a sequence of independent residual vectors $\varepsilon_3, \ldots, \varepsilon_{m+n} \sim \mathrm{N}(o, \widehat{\Sigma})$. Then, putting $y_1 = y_2 = \widehat{\eta}$ where $\widehat{\eta}$ is the sample mean given in (7), we compute the values of $y_3, \ldots, y_{m+n}$ recursively by $y_i = \widehat{A}_1\, y_{i-1} + \widehat{A}_2\, y_{i-2} + \varepsilon_i$ for $i = 3, \ldots, m+n$.

Furthermore, for $i = 1, \ldots, n$, we compute the vectors $r_i = (r_{i1}, r_{i2})$, where $r_i = \Phi(y_{m+i})$ and Φ is given by (4). If $r_{11} = 0$ and $r_{12} = n_1$ for some $n_1 \geq 1$, then in the same way as described above we generate a trajectory $(s_{11}, \ldots, s_{1n_1})$ of length n_1 of the time series $\{S_i,\, i \geq 1\}$ which is the first (straight) part of the trajectory $(\alpha_1, l_1), (\alpha_2, l_2), \ldots$ of $\{X_i,\, i \geq 1\}$ to be simulated. Analogously, if $r_{11} \neq 0$ and $r_{12} = n_1$ for some $n_1 \geq 1$, then we generate a trajectory $(c_{11}, \ldots, c_{1n_1})$ of length n_1 of $\{C_i,\, i \geq 1\}$, and so on. Notice that $s_i = \Psi(y'_{m+i})$

and $c_i = \Psi(y''_{m+i})$, where Ψ is given by (5). The procedure terminates after k_0 steps, where $k_0 = \min\{k : \sum_{i=1}^{k} l_i \geq l\}$.

3.3.2. Curvature of Polygonal Tracks

We now analyze how good the curvature of the polygonal tracks extracted from the SEM image is reflected by the fitted single–fiber model, where the following curvature measure is considered. Let $\beta_0(p) = \beta(p)/\ell^2$, where $\beta(p)$ denotes the area circumscribed by the polygonal track $p = (p_1, \ldots, p_n)$ and ℓ is the length of the (straight) line segment connecting starting point of p_1 and the end point of p_n, see Figure 12.

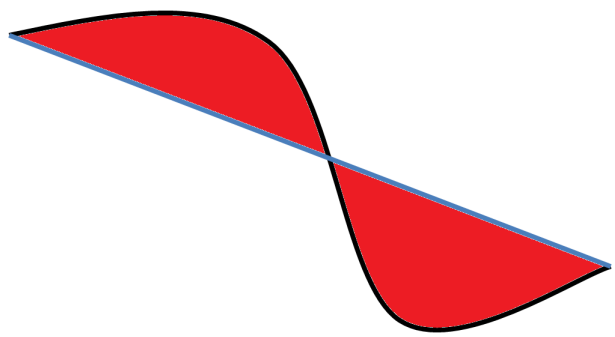

Figure 12: Curvature measure $\beta_0(p)$: Circumscribed area (red) $\beta(p)$ divided by the square length ℓ^2 of the blue line

Histograms of $\beta_0(p)$ have been computed for (real) polygonal tracks extracted from the SEM image and for simulated polygonal tracks drawn from the fitted single–fiber model $\{X_i,\, i \geq 1\}$. Figure 13 shows that these two histograms nicely coincide.

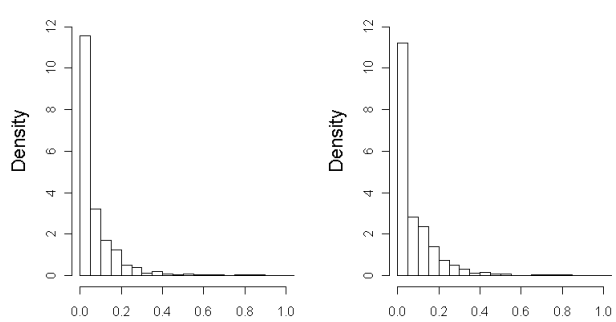

Figure 13: Histograms of $\beta_0(p)$ for real (left) and simulated (right) tracks

Subsequently, distributional characteristics of $\beta_0(p)$ have been computed separately for curved and straight parts of extracted and simulated fibers, see Figure 14. They show that also the components $\{C_i,\, i \geq 1\}$ and $\{S_i,\, i \geq 1\}$ of the single–fiber model $\{X_i,\, i \geq 1\}$ are fitted quite well.

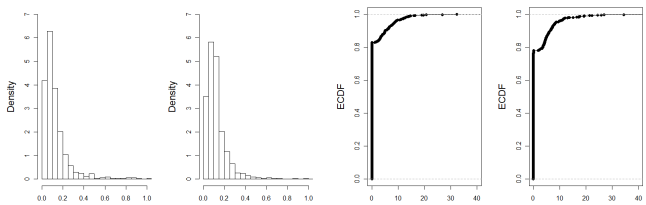

Figure 14: Histograms and cummulative distribution functions of $\beta_0(p)$ for real and simulated data of curved (left) and straight (right) parts

Finally, the goodness–of–fit is analyzed for the time series $\{R_i,\, i \geq 1\}$ governing the transitions between curved and straight parts. This is accomplished by comparing the histograms of the components R_{i1} and R_{i2} of the random vectors $R_i = (R_{i1}, R_{i2})$ which have been computed for (real) data extracted from the SEM image and for simulated data, respectively. Figure 15 shows that also these histograms are similar.

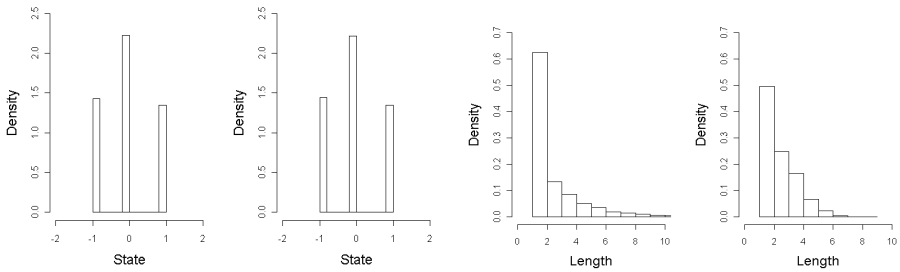

Figure 15: Histograms of real and simulated data for the first (left) and second (right) component of R_i

4. Multi-Layer Model

Using the single–fiber model introduced in Section 3, in the next modeling step we first construct a stochastic microstructure model for single (thin) layers of fibers. This will be a germ–grain model such that the germs form a homogeneous Poisson point process in 2D and the grains are random planar polygons dilated in 3D. Then, the multi–layer model is given by a stack of the modeled single layers. It turns out that the multi–layer model is able to adequately

reproduce the 3D microstructure of fiber systems with strongly curved fibers running mainly in horizontal direction.

4.1. Model Description

Our basic modeling assumption is to suppose that the considered fiber–based material can be decomposed into stochastically independent thin layers with a thickness equal to one fiber diameter, i.e., we assume the fibers to be mutually interpenetrating curved tubes which are horizontally oriented within each thin layer. The vertical cross section visualized in Figure 16 shows that the latter assumption is nicely fulfilled for the non–woven GDL material considered in this paper.

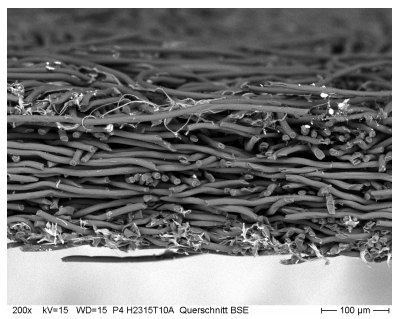

Figure 16: Cross section through the GDL material

This visual impression is approved by the result of a statistical analysis of edge directions drawn from a 3D graph. It represents the fiber courses extracted from a 3D image gained by synchrotron tomography from the same material. The 3D graph is obtained by skeletonization of the solid phase of the binarized 3D image and by subsequent transformation into vector data, where standard algorithms of 3D image processing have been applied, using the software system 'Avizo' [5]. In particular, the histogram $e : [0, 2\pi] \times [0, \frac{\pi}{2}] \to [0, \infty)$ of edge directions has been computed. Note that $e(\phi, \theta)$ is the frequency of edges having a normalized direction vector with spherical coordinates $(\phi, \theta, 1)$, where ϕ denotes the azimuthal angle, θ the polar angle, and 1 the (normalized) length.

Figure 17 shows that the edges proceed mainly horizontally. However, note that there is a small peak in the vertical direction which is not attended in our model. It can also be seen from Figure 17 that the edges seem to have no preferential directions with respect to their azimuthal angle. This motivates our choice of an isotropic germ–grain model for the thin layers.

More precisely, the single (thin) layers are described by a germ–grain model which is given by the set union

$$\Xi = \bigcup_{n=1}^{\infty} \left((X^{(n)} \oplus b(o, r)) + P_n \right), \tag{8}$$

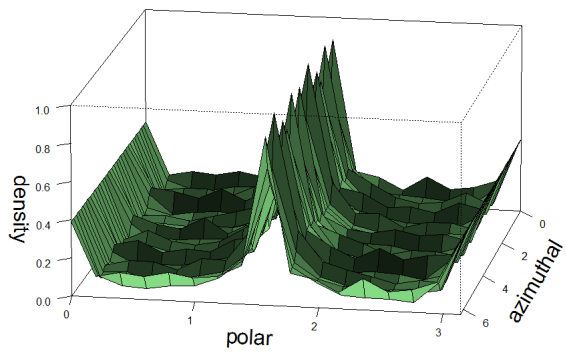

Figure 17: Histogram of edge directions in 3D

where the germs P_n form a Poisson point process $\{P_n\}$ in the x-y–plane $E = \mathbb{R}^2 \times \{0\} \subset \mathbb{R}^3$ with some intensity $\lambda > 0$, and the grains $X^{(1)} \oplus b(0,r), X^{(2)} \oplus b(o,r), \ldots$ are given by random planar polygons $X^{(1)}, X^{(2)}, \ldots$ of some (fixed) length $l > 0$ which are independent copies drawn from the single–fiber model $X = \{X_i, i \geq 1\}$ introduced in Section 3 and then spherically dilated in 3D by the ball $b(o,r) \subset \mathbb{R}^3$ with some (deterministic) radius $r > 0$. This means that the planar section $\Xi \cap E$ is a so–called 'Boolean model' in $\mathbb{R}^2$ [9, 20].

Thus, besides the parameters of the single–fiber model considered in Section 3, the single–layer model has three further parameters λ, r and l which have to be specified.

4.2. Model Fitting

For the fiber–based GDL material considered in this paper it is known from the manufacturer that the fibers have a diameter of $9.5\,\mu m$, i.e. $r = 4.75\,\mu m$, and a length of $50\,mm$, i.e. $l = 50\,000\,\mu m$. Furthermore, it is known that the volume fraction of the pore phase is 0.765 and, consequently, the volume fraction of the fiber system is 0.235. This information can be used to determine the intensity λ such that the volume fraction 0.235 of the solid phase is approximatively matched.

To achieve this goal we use the following relationships. Consider a single (straight) line segment $L \subset \mathbb{R}^3$ of length $l > 0$ which belongs to the (horizontally oriented) x-y–plane $E \subset \mathbb{R}^3$. It is clear that the three–dimensional volume $\nu_3(L \oplus b(o,r)) = \pi r^2 l$ of the spherically dilated segment $L \oplus b(o,r)$ can be approximated by

$$\nu_3(L \oplus b(o,r)) \approx \frac{\pi r}{2}\, \nu_2\big((L \oplus b(o,r)) \cap E\big)\,, \tag{9}$$

where $\nu_2\big((L \oplus b(o,r)) \cap E\big) = 2rl$ is the (two–dimensional) area of the planar section $(L \oplus b(o,r)) \cap E$.

Let $W = [-w, w]^2 \times [-r, r] \subset \mathbb{R}^3$ be a certain sampling window for some $w > 0$. Assuming that the individual line segments of the random planar polygons $X^{(1)}, X^{(2)}, \ldots$ considered in Section 4.1 are sufficiently long and the intensity λ of germs is not too high, then formula (9) can be used to justify the following (approximative) relationship:

$$0.235 \;=\; \frac{\mathbb{E}\nu_3(\Xi \cap W)}{\nu_3(W)} \;\approx\; \frac{\pi}{4}\,\frac{\mathbb{E}\nu_2(\Xi \cap W \cap E)}{\nu_2(W \cap E)} \;=\; \frac{\pi}{4}\,\zeta\,, \tag{10}$$

where $\zeta = \mathbb{E}(\Xi \cap ([0,1]^2 \times \{0\}))$ denotes the area fraction of the Boolean model $\Xi \cap E$. On the other hand, it is well known (see e.g. [9, 20]) that

$$\zeta = 1 - \exp\bigl(-\lambda\,\mathbb{E}\nu_2((X^{(1)} \oplus b(o,r)) \cap E)\bigr)\,. \tag{11}$$

Using for $\mathbb{E}\nu_2((X^{(1)} \oplus b(o,r))) \cap E)\bigr)$ the same kind of an approximation as in (10), i.e. $\mathbb{E}\nu_2((X^{(1)} \oplus b(o,r))) \cap E)\bigr) \approx 2rl$, (10) and (11) give that

$$\lambda \approx \;\frac{-\log(1-\pi^{-1})}{2rl}\;. \tag{12}$$

Thus, having in mind that $r = 4.75$ and $l = 50\,000$, we put $\lambda = 8.067 \cdot 10^{-7}$.

4.3. Simulation Algorithm; Plus Sampling

To avoid edge effects when simulating the germ–grain model Ξ introduced in (8), the sampling window $W = [-w, w]^2 \times [-r, r] \subset \mathbb{R}^3$ has to be increased, i.e., we consider so–called 'plus sampling' with respect to the enlarged window $\widetilde{W} = [-(w+l), w+l]^2 \times [-r, r] \subset \mathbb{R}^3$. Thus, in order to simulate a multi-layer model being a stack of $m > 1$ single–layer models, we can proceed as follows:

1. Simulate the random number N which has a Poisson distribution with mean value $4\lambda(w+l)^2$.
2. If $N = 0$, then repeat step 1.
3. If $N = n > 0$, then simulate n independent and uniformly distributed points $P_1, \ldots, P_n$ in $[-(w+l), w+l]^2$.
4. Simulate the random planar polygons $X^{(1)} \ldots, , X^{(n)}$ of length $l > 0$ by independent sampling from the single–fiber model $X = \{X_i, i \geq 1\}$, using the algorithm stated in Section 3.3.1.
5. Generate the dilated fiber systems $X^{(1)} \oplus b(0,r), \ldots, X^{(n)} \oplus b(o,r)$.
6. Determine the set $\Xi_W = \bigl(X^{(1)} \oplus b(0,r) \cup \ldots \cup X^{(n)} \oplus b(o,r)\bigr) \cap W$
7. Generate m independent copies $\Xi_W^{(1)}, \ldots, \Xi_W^{(m)}$ of Ξ_W.
8. Determine the set $\Xi_W^{(1)} \cup (\Xi_W^{(2)} + (0,0,2r)) \ldots \cup (\Xi_W^{(m)} + (0,0,2r(m-1)))$.

In Figure 18 (right) an examplary output of our simulation algorithm is shown. By visual comparison, one feels that the simulated image coincides quite nicely with the (real) synchrotron image which is shown on the left–hand side of Figure 18.

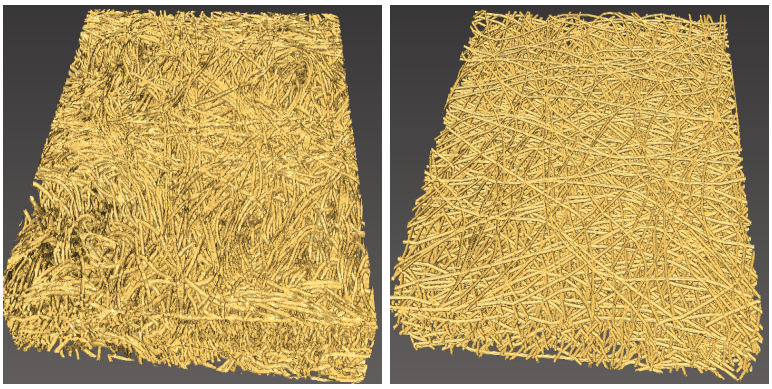

Figure 18: Binarized synchrotron image (left) and simulated 3D image drawn from the multi-layer model (right)

In Section 4.4, we check more formally if the multi–layer model describes the microstructure of the non–woven GDL sufficiently well. Therefore, the stochastic simulation model will be validated by comparing structural characteristics computed from the 3D image gained by synchrotron tomography, and from realizations of the multi–layer model, respectively. The goal is to show that transport–relevant structural properties of the stochastic model match those of the (real) synchrotron image quite well.

4.4. Model Validation

We now show how the fitted multi–layer model can be validated by comparing structural characteristics computed for 3D data gained by synchrotron tomography and for simulated 3D morphologies drawn from the multi–layer model. In particular, it turns out that the porosity of 0.235 is perfectly fitted by our model. Moreover, we consider two further structural characteristics which are relevant for the performance of fiber–based materials.

4.4.1. Spherical Contact Distribution Function

An example of such a characteristic is the cumulative distribution function $H : [0, \infty] \to [0, 1]$ of spherical contact distances of the pore phase, where $H(t)$ denotes the probability that the minimum distance from a randomly chosen location of the pore phase to the fiber phase is not larger than t.

The values of $H(t)$ which have been obtained for the 3D data gained by synchrotron tomography and for simulated 3D morphologies drawn from the multi–layer model, are very similar, see Figure 19. This means that the fit of the multi–layer model is quite good with respect to spherical contact distances of the pore phase.

4.4.2. Geometric Tortuosity

For the GDL material considered in the present paper, the geometrical properties of the percolation pathways through the pore phase play an important role

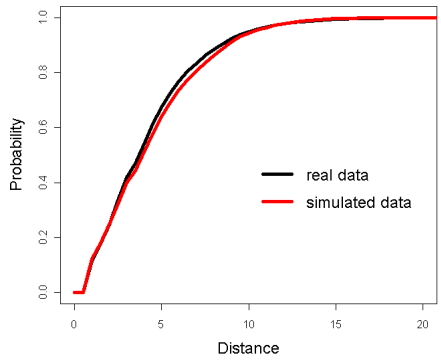

Figure 19: Distribution function H of real (black) and simulated (red) data

because the main function of GDL in PEM fuel cells is the transportation of gases towards the electrodes, and the drainage of water. It is well known that the quality of these transportation processes is closely related with the tortuosity of the pore phase which is often defined as the ratio of the mean effective path length through the pore space divided by the material thickness. As an alternative, the notion of 'geometric tortuosity' has been investigated in the literature, where the lengths of shortest paths along the edges of a certain geometric graph though the pore phase are considered instead of effective path lengths. Moreover, starting from a randomly chosen location on top of the porous material, its geometric tortuosity can be represented by a probability distribution instead of looking at the mean tortuosity only, see e.g. [22, 23].

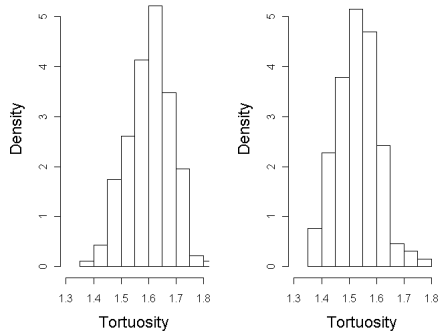

Figure 20: Histograms of tortuosity for real (left) and simulated data (right)

The results which have been obtained for the distribution of geometric tortuosity are given in Figure 20, see also Table 1. These results show that there is no

perfect matching of the distributions of geometric tortuosity computed for the 3D data gained by synchrotron tomography and for simulated 3D morphologies drawn from the multi–layer model. However, the histograms given in Figure 20 are similarly shaped and the difference between the two mean values is relatively small. Thus, also with respect to geometric tortuosity, the stochastic multi-layer model proposed in the present paper provides a reasonable fit to real data.

Table 1: Mean and standard deviation of geometric tortuosity

	mean value	standard deviation
real data	1.60	0.079
simulated data	1.52	0.076
Relative error	0.05	0.03

5. Conclusions and Outlook

We propose a stochastic multi–layer model describing the 3D microstructure of materials which are built up of strongly curved, but almost horizontally oriented fibers. This fully parametrized model is based on ideas from stochastic geometry and multivariate time series analysis. It consists of independent layers which are stacked together, where each single layer is described by a 2D germ–grain model dilated in 3D. The multi–layer approach enables us to fit model parameters on the basis of structural information extracted from 2D images. This is important for industrial applications since 2D images are relatively easy to obtain in contrast to 3D image data. The fitted multi–layer model has been validated by considering structural characteristics which are relevant for the performance of fiber–based materials, like the distributions of spherical contact distances and geometric tortuosity of the pore phase. These characteristics have been computed for 3D data gained by synchrotron tomography and for simulated 3D morphologies drawn from the multi–layer model, where the obtained results were quite similar in each case. Thus, the multi–layer model can be an appropriate tool for the examination of the microstructure of fiber–based materials.

In particular, our model can be used for virtual scenario analyses with the general aim to construct microstructures with improved physical properties. By systematic modifications of model parameters and in combination with numerical transportation models, virtual 3D morphologies and virtual transportation processes in these morphologies can be simulated. In this way, the identification of optimized morphologies is possible which will be the subject of a forthcoming paper.

Acknowledgement

This research has been supported by the German Federal Ministry of Education and Research (BMBF) in the framework of the priority program ‘Math-

ematics for Innovations in Industry and Services'.

References

[1] Altendorf, H. and Jeulin, D. (2011). Random walk based stochastic modeling of 3D fiber systems. *Physical Review E*, **83**, 041804.

[2] Brockwell, P.J. and Davis, R.A. (1991). Time Series: Theory and Methods. New York: Springer.

[3] Burger, W. and Burge, M.J. (2007). Digital Image Processing: An Algorithmic Introduction Using Java. Berlin: Springer.

[4] Chan, N.H. (2010). Time Series: Applications to Finance with R and S-Plus. New Jersey: J. Wiley & Sons.

[5] Fourard, C., Malandain, G., Prohaska, S. and Westerhoff, M. (2006). Blockwise processing applied to brain microvascular network study. *IEEE Transactions on Medical Imaging*, **156**, B1339–B1347.

[6] Fuller, W. A. (1996). Introduction to Statistical Time Series. 2. ed., New York: J. Wiley & Sons.

[7] Hartnig, C., Jörissen, L., Kerres, J., Lehnert, W. and Scholta, J. (2008). Polymer electrolyte membrane fuel cells (PEMFC). In: M. Gasik (ed.), *Materials for Fuel Cells*. Cambridge: Woodhead Publishing, 101–184.

[8] Hoshen, J. and Kopelman, R. (1976). Percolation and cluster distribution. I. Cluster multiple labeling technique and critical concentration algorithm. *Physical Review B*, **14**, 3438–3445.

[9] Illian, J., Penttinen, A., Stoyan, H. and Stoyan, D. (2008). Statistical Analysis and Modelling of Spatial Point Patterns. Chichester: J. Wiley & Sons.

[10] Inoue, G., Yoshimoto, T., Matsukuma, Y. and Minemoto, M. (2008). Development of simulated gas diffusion layer of polymer electrolyte fuel cells and evaluation of its structure. *Journal of Power Sources*, **175**, 145–158.

[11] Jähne, B. (2005). Digital Image Processing. 6th revised and extended ed., Berlin: Springer.

[12] Jia, K. and Li, X. (2011). Water transport in polymer electrolyte membrane fuel cells. *Progress in Energy and Combustion Science*, **37**, 221–291.

[13] Lütkepohl, H. (2006). New Introduction to Multiple Time Series Analysis. Berlin: Springer.

[14] Mathias, M.F., Roth, J., Fleming, J. and Lehnert, W. (2003). Diffusion Media materials and characterisation. In: W. Vielstich, A. Lamm and H. Gasteiger (eds.), *Handbook of Fuel Cells*. London: J. Wiley & Sons, 517-537.

[15] Ohser, J. and Schladitz, K. (2009). 3D Images of Materials Structures - Processing and Analysis. Weinheim: Wiley-VCH.

[16] Redenbach, C. and Vecchio, I. (2011). Statistical analysis and stochastic modeling of fibre composites. *Composites Science and Technology*, **71**, 107–112.

[17] Schulz, V.P., Becker, J., Wiegmann, A., Mukherjee, P.P. and Wang, C.Y. (2007). Modeling of two–phase behavior in the gas diffusion medium of PEFCs via full morphology approach. *Journal of the Electrochemical Society*, **154**, B419–B426.

[18] Soille, P. (2003). Morphological Image Analysis: Principles and Applications. 2nd ed., Berlin: Springer.

[19] Spiess, M. and Spodarev, E. (2010). Anisotropic Poisson processes of cylinders. *Methodology and Computing in Applied Probability*, **13**, 801–819.

[20] Stoyan, D., Kendall, W.S. and Mecke, J. (1995). Stochastic Geometry and its Applications. 2nd ed., Chichester: J. Wiley & Sons.

[21] Thiedmann, R., Fleischer, F., Lehnert, W. and Schmidt, V. (2008). Stochastic 3D modelling of the GDL structure in PEM fuel cells, based on thin section detection. *Journal of the Electrochemical Society*, **155**, B391–B399.

[22] Thiedmann, R., Hartnig, C., Manke, I., Schmidt, V. and Lehnert, W. (2009). Local structural characteristics of pore space in GDL's of PEM fuel cells based on geometric 3D graphs. *Journal of the Electrochemical Society*, **156**, B1339–B1347.

[23] Thiedmann, R., Manke, I., Lehnert, W. and Schmidt, V. (2011). Random geometric graphs for modelling the pore space of fibre-based materials. *Journal of Materials Science*, **46**, 7745–7759.