*Journal of The Electrochemical Society*, **155** (4) B391-B399 (2008)
0013-4651/2008/155(4)/B391/9/$23.00 © The Electrochemical Society


# Stochastic 3D Modeling of the GDL Structure in PEMFCs Based on Thin Section Detection

**Ralf Thiedmann,[a] Frank Fleischer,[b] Christoph Hartnig,[c] Werner Lehnert,[c,*,d,z] and Volker Schmidt[a]**

[a]*Institute of Stochastics, Ulm University, 89069 Ulm, Germany*
[b]*Medical Data Services/Biostatistics, Boehringer Ingelheim Pharma GmbH and Company KG, 88397 Biberach, Germany*
[c]*Center for Solar Energy and Hydrogen Research Baden-Württemberg, 89069 Ulm, Germany*

We propose a mathematical model to describe the microstructure of the gas diffusion layer (GDL) in proton exchange membrane fuel cells (PEMFCs) based on tools from stochastic geometry. The GDL is considered as a stack of thin sections. This assumption is motivated by the production process and the visual appearance of relevant microscopic images. The thin sections are modeled as planar [two-dimensional (2D)] random line tessellations which are dilated with respect to three dimensions. Our 3D model for the GDL consists of several layers of these dilated line tessellations. We also describe a method to fit the proposed model to given GDL data provided by scanning electron microscopy images which can be seen as 2D projections of the 3D morphology. In connection with this, we develop an algorithm for the segmentation of such images which is necessary to obtain the required structural information from the given grayscale images.
 [DOI: 10.1149/1.2839570] 

Manuscript submitted October 15, 2007; revised manuscript received December 27, 2007.
Available electronically February 20, 2008.

In order to use hydrogen as an energy carrier, an effective way to convert hydrogen into electricity is necessary, with the fuel cell technology being an appropriate means due to its high efficiency. A key component of a fuel cell is the gas diffusion layer (GDL); see Fig. 1. The GDL is responsible for the transport of hydrogen and oxygen to the electrodes where the electrochemical reaction takes place.[1] Also, its task is to remove the water produced. In this context it is important that just the right water content is achieved, because water is essential for the conductivity of the membrane, whereas excess water in the GDL leads to flooding of the pores, which in turn limits current density. Hence, the balance between water drainage and water storage is the key to high performance. Additionally, the GDL acts as a conductor for the electricity produced.

For a better understanding of the physical processes within the GDL and for an optimization of its design, a well-fitted structural model is needed. In the present paper we propose a model of the GDL using tools from stochastic geometry, in particular, random tessellations. Note that the advantage of this type of stochastic model is that the complex morphology of the GDL can be described by few parameters. In our case, the morphology of the GDL is formed by a large number of irregularly located long fibers with negligible curvatures. The fundamental idea of our structural model arises from the production process of the GDL considered, where fibers are disposed (almost) horizontally. This leads to the assumption that a GDL can be seen as a stack of thin sections, i.e., thin (separated) sublayers of fibers, which fits the structure shown in Fig. 1 quite well. Furthermore, Fig. 1 shows that the individual fibers can be approximated by straight cylinders. Therefore, in a first step the thin sections of the GDL are represented by planar random line tessellations, which are built by intersecting lines located at random in the Euclidean plane. Then, these lines are "blown up" to three-dimensional (3D) objects, i.e., they are dilated in 3D. Each 3D-dilated line tessellation represents a thin section, i.e., a separated (sub)layer of the GDL. Finally, a certain number of such dilated line tessellations are stacked together and, in this way, a 3D model for the whole GDL is obtained.

At present, the process of obtaining 3D images of GDL, e.g., by means of synchrotron tomography,[2] is a complex and expensive task and therefore not always practicable. 2D electron microscopy images can be obtained much more easily. They can be used, for example, in the production control of the GDL. Therefore, it is a research approach to describe the morphology of the GDL by a multilayer model on the basis of 2D projection data. Another stochastic model for the GDL was recently considered in Ref. 3 and 4. This model originally stems from Ref. 5, where it was applied to the morphological structure of nonwoven porous materials. In these papers, a 3D dilated Poisson line process is used to model the fibers of the GDL. The anisotropy in the $z$ direction is modeled with an anisotropy parameter $\beta$ which has to be estimated from thin section images of the GDL. In contrast to this model, our model assumes that the orientation in the $z$ direction is negligible for this kind of GDL material. Therefore, an anisotropy parameter need not be determined, and with the segmentation algorithms proposed in the present paper, all required information for the fiber modeling can be gained from only 2D scanning electron microscopy (SEM) images. In addition to the structural model for the fibers of the GDL, we have introduced a modeling component for binder material.

An important aspect of constructing a suitable model is the possibility of fitting it to real data. In our case, data are given as 2D projection images from electron microscopy (see Fig. 1a). To get structural information from such images, it is necessary to apply suitable segmentation algorithms. Thus, the development of such an algorithm is another main part of the present paper, where we additionally consider a 3D control image of a GDL (see Fig. 1b), which is used to validate the quality of our segmentation techniques and model fit.

This paper is organized as follows. In the next section, the proposed multilayer model for the GDL diffusion layer is explained. Next, the segmentation of the image data, as illustrated in Fig. 1, is described. Then, a method for the identification of single thin sec-

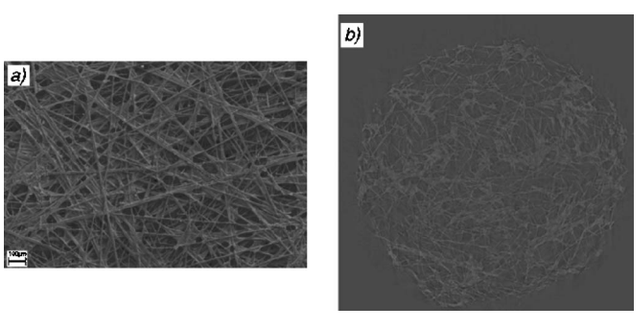


**Figure 1.** Visualization of a GDL: (a) 2D projection image and (b) 3D thin-section image.

* Electrochemical Society Active Member.
[d] Present address: Jülich Research Center, 52425 Jülich, Germany.
[z] E-mail: w.lehnert@fz-juelich.de

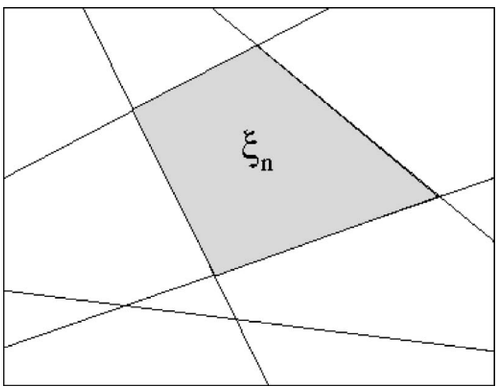


**Figure 2.** Cell $\xi_n$ of a planar line tessellation.

tions is discussed. This is a first step toward a statistical model fitting algorithm explained in the Statistical Model Fitting section. Here, some first results of our statistical analysis of the GDL and a validation of our multilayer model are also presented.

## Stochastic Multilayer Model

In this section we introduce a mathematical model for the microstructure of the GDL in proton exchange membrane fuel cells (PEMFCs) which is based on tools from stochastic geometry (see, e.g., Ref. 6-8). A closer look at the structure of the GDL shown in Fig. 1 leads to the impression that the single fibers can be approximated by dilated straight lines. Therefore, it seems appropriate to use a model which incorporates straight lines, at least for the material considered in the present paper.

The location of the fibers, visible in Fig. 1, and the production technique for this type of GDL, suggest that the fibers are mainly horizontally oriented. Additionally, we assume that the GDL is a stack of several thin sections which all contain horizontally oriented fibers. We are aware that this assumption is a simplification in comparison to the real GDL, because the fibers in one thin section are crossing each other to a small extent, which is neglected in our model. Here, the fibers within a given thin section are seen as mutually penetrating cylinders.

*Random line tessellations.*— We use planar random line tessellations as a modeling element for the fibers within the individual thin sections, in particular, so-called Poisson line tessellations (PLT). A planar line tessellation is built by the cells $\zeta_1,\zeta_2,\ldots$ which arise from intersecting lines $l_1,l_2,\ldots$ scattered in the plane. Examples of such cells can be seen in Fig. 2 and 3b. Note that the lines can be given as follows: Each line $l_n$ can be described in its normal form, i.e., its (orthogonal) signed distance $x_n$ from the origin and a direction $m_n$ (given as the angle to a predetermined direction; see Fig. 3a). Then these distances $x_n$ and directions $m_n$ are understood as realizations of certain random variables $X_n$ and $M_n$, respectively. The values of the random signed distances $X_n$ can be arbitrary real numbers, whereas

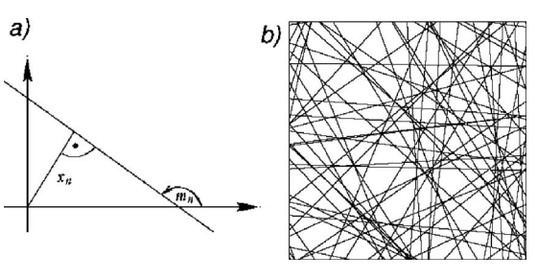


**Figure 3.** Line tessellations: (a) normal form of a line in $\mathbb{R}^2$ and (b) realization of a PLT.

**Table I. Mean value formulas for PLTs.**

| $\lambda_1(\gamma)$ | $\lambda_2(\gamma)$ | $\lambda_3(\gamma)$ | $\lambda_4(\gamma)$ | $A(\gamma)$ | $L(\gamma)$ |
|---|---|---|---|---|---|
| $\frac{1}{\pi}\gamma^2$ | $\frac{2}{\pi}\gamma^2$ | $\frac{1}{\pi}\gamma^2$ | $\gamma$ | $\frac{\pi}{\gamma^2}$ | $\frac{2\pi}{\gamma}$ |

the random directions $M_n$ take values between 0 and $\pi$ (see also Fig. 3a). The family of cells $\Xi_1,\Xi_2,\ldots$ formed by all the intersecting random lines $(X_1,M_1),(X_2,M_2),\ldots$ then forms a random line tessellation.

In the proposed model for the individual thin sections of the GDL, a special class of planar random tessellations, PLTs, is considered, where the sequence $(X_1,M_1),(X_2,M_2),\ldots$ is an independently marked Poisson process. This means that the random distances $X_1,X_2,\ldots$ form a homogeneous Poisson point process on the real line with some intensity $\gamma > 0$, where $\gamma$ is the expected number of points per unit length, i.e., $\gamma = \mathbb{E}\,\#\{n{:}X_n \in [0,1]\}$. Furthermore, the random directions $M_1,M_2,\ldots$ are independent and identically distributed random variables whose values are uniformly distributed in the interval $[0,\pi]$. Moreover, the sequences $\{X_n\}_{n\geq 1}$ and $\{M_n\}_{n\geq 1}$ are independent.

Note that the cells $\Xi_1,\Xi_2,\ldots$ of a PLT form a stationary and isotropic sequence of random polygons, which are bounded and convex with probability 1. Here, stationarity and isotropy mean that the distribution of $\{\Xi_n\}_{n\geq 1}$ does not change under translations and rotations around the origin, respectively. This model can be fully described by a single parameter, the intensity $\gamma$ of the underlying Poisson process $\{X_n\}_{n\geq 1}$. It can be shown that $\gamma$ is equal to the mean total length of the (random) set $\cup_{n=1}^{\infty}\,\partial\,\Xi_n$ of edges per unit area, i.e.

$$\gamma = \mathbb{E}\left|\bigcup_{n=1}^{\infty}\,\partial\,\Xi_n \cap [0,1]^2\right| \qquad [1]$$

This means that in an arbitrary fixed sampling window there are on average more lines (and cells, respectively) for larger values of $\gamma$, as can be seen in Fig. 3b, than for smaller values of $\gamma$, as in the realization shown in Fig. 2.

We also note that there are analytical formulas for different characteristics of PLT. In particular, denoting the mean number of vertices by $\lambda_1$, the mean number of edges by $\lambda_2$, and the mean number of cells by $\lambda_3$, all referring to the unit area, it can be shown that the model characteristics $\lambda_1 = \lambda_1(\gamma)$, $\lambda_2 = \lambda_2(\gamma)$, and $\lambda_3 = \lambda_3(\gamma)$ depend via simple formulas on the model parameter $\gamma$. Besides this, denoting the mean total edge length per unit area by $\lambda_4$, in Eq. 1 it is stated that $\lambda_4(\gamma) = \gamma$. Furthermore, simple formulas are known for the mean area $A(\gamma)$ and the mean boundary length $L(\gamma)$ of the "typical cell" of a PLT with intensity $\gamma$ (see Table I). Here, mathematically, the notion of typical cell means a cell chosen at random among all cells available in an unboundedly large sampling window. These mean values are used in the Statistical Model Fitting section.

Above, the PLT is only described in 2D. But the fibers are 3D objects. To get real 3D objects we "blow up" the lines of the tessellation. More formally, we dilate each line, which is explained in detail in the next sections.

*Morphological operators.*— For the 3D multilayer model of the GDL, which is introduced in the next section, and for the segmentation algorithm developed in a later section, two morphological operators are used. The first one is the so-called dilation in the $d$-dimensional Euclidean space $\mathbb{R}^d$, where $d \geq 1$ is an arbitrary integer. Simplified, one can imagine a dilation of a set as blowing it up. Note, however, that for purposes of statistical image analysis the cases $d = 2$ and $d = 3$ are most relevant.

Mathematically, the dilation of a set $A \subset \mathbb{R}^d$ with a set $B \subset \mathbb{R}^d$ can be described by the Minkowski sum $A \oplus B$ of $A$ and $B$, where

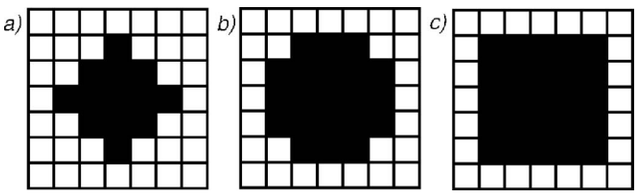


**Figure 4.** Different possibile ways to discretize the profiles of fibers: (a) lozenge, (b) square without corners, and (c) square.

$$A \oplus B = \{a + b : a \in A, b \in B\} \quad [2]$$

Furthermore, the so-called erosion is considered, where the erosion of a set $A \subset \mathbb{R}^d$ with a set $B \subset \mathbb{R}^d$ can be described by the Minkowski difference $A \ominus B$ of $A$ and $B$, which is given by

$$A \ominus B = \bigcap_{b \in B} A + b \quad [3]$$

where $A + b$ denotes the shifted set $A + b = \{a + b : a \in A\}$. Simplified, one can image erosion as shrinking of a set. Note that the two morphological operators of dilation and erosion are not inverse to each other, i.e., in general we have $(A \oplus B) \ominus B \neq A$. For more details, see, e.g., Ref. 9.

*3D dilation of PLT.*— In order to model a single thin section of fibers, the edge set $\cup_{n=1}^{\infty} \partial \Xi_n$ of the underlying PLT $\{\Xi_n\}_{n \geq 1}$ is dilated with respect to 3D, i.e., the Minkowski sum

$$\left( \bigcup_{n=1}^{\infty} \partial \Xi_n \right) \oplus B = \bigcup_{n=1}^{\infty} \partial \Xi_n \oplus B \quad [4]$$

is considered for some set $B \subset \mathbb{R}^3$, leading to an object in 3D. Thus, the stack of these 3D-dilated PLTs also represents a 3D model.

For the dilation of the edges of a PLT, set $B$ in Eq. 4 can be chosen, e.g., as a ball or a cube in 3D, which is called a structuring element by some authors. It models the profile of the fibers. Furthermore, for practical computations, a discretization of the structuring element $B$ is necessary. Figure 4 shows three possible versions of discretizations. In all three cases, five pixels are considered along both the $x$ and the $y$ axis. This choice is motivated by the fact that the diameter $2r$ of the fibers is approximately 7 μm, whereas the length of a pixel is 1.5 μm.

In the proposed multilayer model, the 3D-dilated PLTs representing the individual thin sections are assumed to be independent and identically distributed. This assumption can be (at least partly) justified by considering the production process and the visual appearance of GDLs in images like in Fig. 1. A schematic display of the multilayer model can be seen in Fig. 5.

*Modeling of the binder.*— The fibers of the GDL are adhered by means of a binder, which is introduced in the manufacturing process and can be recognized in the images as thin films. To incorporate the binder into the 3D model, the following approach is considered. Each cell of the 3D-dilated PLT in any given thin section is filled with a certain probability $p \in [0,1]$, whether or not the other cells of the PLT are filled (see Fig. 6a). This approach can be justified by looking at Fig. 1, where the binder looks like a thin, horizontally orientated set of material. For the binder, various other modeling approaches are possible. For example, 3D spheres could be considered with some (deterministic or random) radius $r_0 > 0$. These spheres could be located at some subset of vertices of the PLT which are chosen at random with some probability $b \in [0,1]$ (see Fig. 6b). This also seems to be a realistic model for the binder, because sometimes there are clusters of binder in the GDL (see, e.g., the left side of Fig. 1a). However, in the following we focus on the Bernoulli-filling approach as a model for the binder.

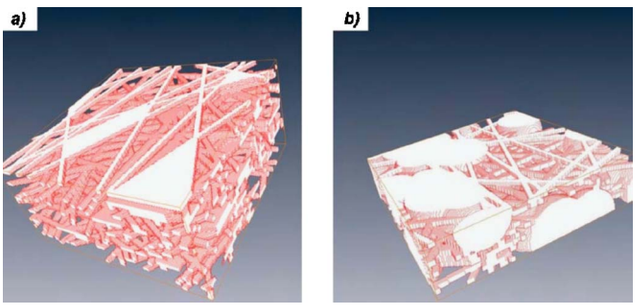


**Figure 6.** (Color online) Realizations of the 3D multilayer model with binder: (a) binder modeled with Bernoulli filling and (b) binder modeled with spheres.

## Segmentation of 2D Images

The segmentation of 2D electron microscopy images of GDL, as shown in Fig. 1a, and the identification of thin sections out of the segmented images is another important issue of the present paper. To achieve satisfactory results, different methods of image processing have to be combined.

The segmentation algorithm can be divided into two parts. First, a preprocessing step is performed which smooths the data, i.e., removes noise, while emphasizing the edges in the image. In the next step, vector data are created, i.e., line segments which follow the fibers are determined. Afterward, a method for the identification of individual thin sections is presented. The detection of these sections is necessary to fit the 3D model introduced above to real data as described in the Statistical Model Fitting section.

The goal is to gain a fully noninteractive segmentation and thin section detection algorithm for 2D SEM images. This leads finally to a fully noninteractive model-fitting procedure for the proposed multilayer model of the GDL. Therefore, techniques which are as general as possible are considered.

*Anisotropic diffusion.*— A closer look at the 2D SEM images (see Fig. 1a) indicates that the observed structures are not smooth, which makes segmentation more difficult. Therefore, the data have to be smoothed in advance of a further segmentation step. A global smoothing with, e.g., a Gaussian filter (see Ref. 9) has the disadvantage that the edges are blurred, too. In contrast to this, anisotropic diffusion[10] is a technique which smooths the data without losing information at the edges, i.e., the edges remain sharp.

The idea of anisotropic diffusion is quite simple. Before smoothing, edges are detected and smoothing is not performed across them. In other words, only the data within cells, bordered by the already detected edges, are smoothed. The result is a smoothed image with additionally emphasized edges, i.e., the noise is removed but the structure is preserved. The detailed algorithm can be found in Ref. 10.

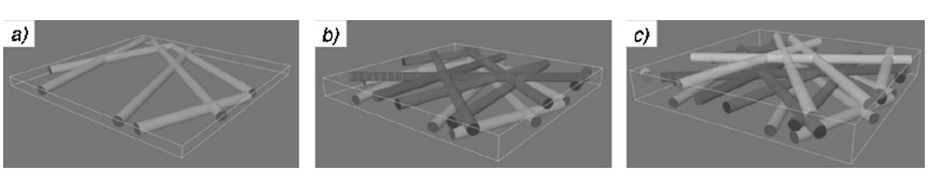


**Figure 5.** Schematic display of the multilayer model: (a) one, (b) two, and (c) three layers.

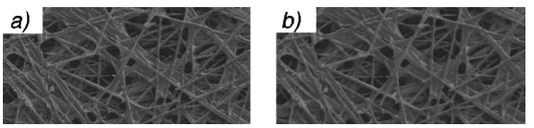


**Figure 7.** Cutout of a 2D SEM image of the GDL: (a) the original image and (b) after applying anisotropic diffusion, where noise is removed but the edges remain sharp.

The result of anisotropic diffusion applied to the data, which are visualized in Fig. 1, can be seen in Fig. 7. The regions between edges are smoothed, whereas edges themselves are still sharp.

*Convolution and binarization.*— In the next step, edges are emphasized by a convolution of the image, which leads to an image as displayed in Fig. 8a. For applying the skeletonization method described below, a binary image, i.e., an image where every pixel has either the value 0 or 1, is necessary. To keep this binarization as universal as possible, the isodata algorithm, described, e.g., in Ref. 11, is used. This is an iterative algorithm which works as follows. The image is divided into two parts, object and background, by taking an initial threshold $t_i$. Then, the averages of pixels belonging to the object and the background are computed separately. With

$$t_c = \frac{1}{2}(\text{average background} + \text{average objects})$$

we get a value $t_c$ which is compared with $t_i$, i.e., we check if $t_c < t_i$ holds. If the inequality does not hold, the initial threshold $t_i$ is increased until the (final) threshold is found and the image is binarized by applying $t_i$. This iterative method stops when the inequality $t_c < t_i$ holds for the first time. The result can be seen in Fig. 8b. Note that in order to enhance clarity of presentation the picture is inverted, i.e., black and white pixels are interchanged in comparison to Fig. 8a.

*Skeletonization.*— For the transformation of the binary image into vector data, a skeletonization is necessary. This means that pixels belonging to the object are changed to background pixels in a way that the remaining (pixel given) lines have a thickness of one pixel, and the connectivity, or homotopy, is thereby not changed. For the skeletonization which is performed by morphological thinning,[9] the same algorithms as described in Ref. 12 are used. There the binary image is compared with a $3 \times 3$ matrix at each pixel, i.e., if the matrix fits the image, the currently considered pixel is removed from the object. This is repeated for a cycle of rotated matrices until in a whole cycle no pixel is further removed.

After skeletonization, a classification of the remaining pixels is necessary. Here we are especially interested in the nodes of the remaining structure, because we want to create a graph consisting of line segments with the nodes as start point or end point, respectively. This classification is done by counting the number of neighbors. A pixel with only one neighbor is considered an end point, a pixel with exactly two neighbors is classified as a line point, and all other pixels are considered cross points.

A problem that occurs is that by such a classification too many pixels are classified as cross points (see Fig. 9a and b). In other words, at the crossings there is typically not a unique pixel marked as a cross point but two or more pixels. Therefore, the classification

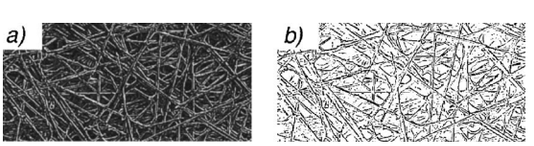


**Figure 8.** Convolution and binarization: (a) cutout after convolution and (b) cutout of the binary image.

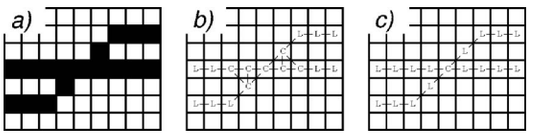


**Figure 9.** Classification after skeletonization. Crosspoints are denoted by C and line points are denoted by L: (a) original skeleton, (b) classification by number of neighbors, and (c) improved classification.

has to be improved with respect to homotopy, i.e., the number of cross points have to be reduced while the connectivity has to be preserved. But, the reduction of cross points is not unique. Therefore, an additional criterion to decide between different possible solutions has to be found. In this case the angles to the $x$ axis, i.e., the directions of the line segments, are considered. Differences of angles for the connections of a pixel to its neighbors are considered, where angles close to $\pi$ are preferred.

In the next step, nearby cross points which have a distance smaller than a given value $d_{max}$ of pixels (in our case $d_{max} = 3$ pixels) are joined together to one cross point (see Fig. 10). This is done because such closely located cross points result from crossings of fibers with quite similar orientations.

The skeletonized image is transformed into a graph structure, i.e., a data set which is no longer based on pixel values but on vector data. In our case we get a data set of line segments, i.e., we have start and end points, for which we assume that they are connected by a straight line. The start or end points, respectively, are the cross points. Two cross points are connected by a line segment if they have a connection consisting only of line points with regard to the pixel-based image. The result of skeletonization and subsequent transformation into a graph structure can be seen in Fig. 11a.

*Reconstruction of single fibers.*— The skeletonized image contains mostly short line segments. But the fibers in the original GDL can be seen as lines or at least as long line segments (see Fig. 1a). Therefore, we connect the short line segments to get longer polygonal tracks. This is done by a modified technique of simulated annealing.

The idea is to connect line segments if they have a similar orientation and if the other segment has an end point in the polytope at the end of the currently considered line segment (see Fig. 11b). Then, a new line segment which connects the two line segments is added. Afterward, the next line segment at the new end point of the current polygonal track is searched in the same way as before. The procedure is repeated until no further line segments can be found in

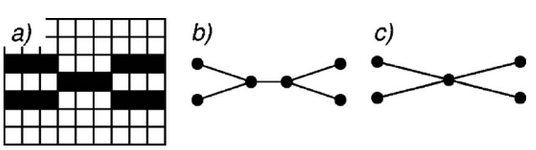


**Figure 10.** Transformation into graph structure and merging of nearby crossings: (a) nearby crossings, (b) after transformation of (a) into a graph, and (c) merging of nearby crossings.

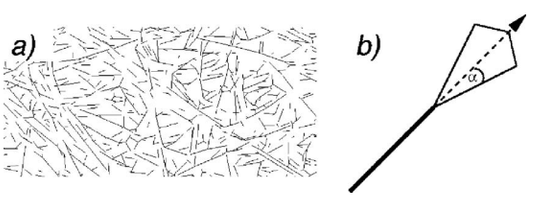


**Figure 11.** Skeletonization and connection of the line segments: (a) cutout after skeletonization and (b) polytope in which line segments are considered.

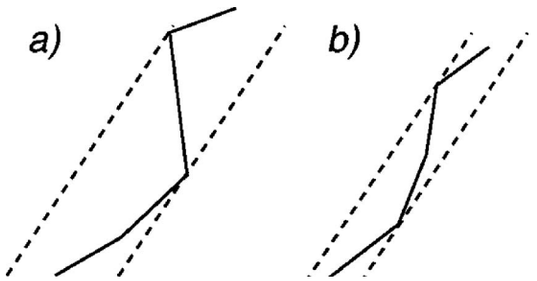


**Figure 12.** Measure for straightness: (a) wide-scattered polygonal track and (b) nearly straight polygonal track.

this way. The result is an almost straight polygonal track which is quite long, i.e., which runs (in the ideal case) across the whole sampling window.

The decision of which line segment of all possible line segments with an end point in the considered polytope is chosen as the next line segment is made as follows. All line segments to which a connection is possible are weighted. The weights are normalized such that the sum of the normalized weights equals 1. In more detail, each line segment is associated with an (sub)interval in $[0,1]$, the length of which depends on its weight. Then, the realization of a random variable which is uniformly distributed in $[0,1]$ determines the next line segment. For the weights, the following two criteria are considered.

1. The angle between the direction of the currently last line segment and the direction of the line segment which has to be included to connect the new potential line segment is considered. The weight is increased if the directions get more similar.
2. The shape of the already found polygonal track is considered. If the currently considered polygonal track is right-curved, the probability for line segments which keep the polygonal track right-curved is higher than for other line segments. The same is done for left-curved polygonal tracks. This criteria is neglected if the current polygonal track is nearly straight.

Because the polygonal tracks are constructed according to random selection rules, it is obvious that the result of a first trial cannot be assured to be optimal. To find nearly optimal solutions, the simulation is repeated several times until results of a similar good quality are found for a certain number of subsequent runs. Therefore, we use a stopping rule which takes into account the following two cases.

1. The end of the polygonal track is close to the boundary of the sampling window. In this case a better polygonal track is found if the polygonal track is straighter, where straighter means that the minimum distance between two bounding parallel lines is smaller (see Fig. 12).
2. The polygonal track ends are not close to the boundary of the sampling window. In this case the better polygonal track is the longer one. But to ensure that a polygonal track which follows a zig-zag is not preferred to a straighter one, which is in fact better, the length is measured as the (Euclidean) distance of the two end points.

Note that not all line segments of the skeletonized image are used as initial line segments for a polygonal track, but only segments which are longer than a certain threshold. This is done because for short line segments, the direction can be quite different from the direction of the original fiber. Note that for initial line segments which have no point close to the boundary of the sampling window, the search for polygonal tracks is started in both directions.

### Identification of Thin Sections

In reality (see Fig. 1a) the fibers are not directly separated into (sub)layers, but they overlap each other to a certain extent. However, in our model, we assume that the fibers interpenetrate, i.e., we suppose that the fibers of the GDL form a stack of thin horizontally oriented sections. This assumption leads to the need to detect these sections from the segmented graph, which is obtained by the image processing discussed in the previous section. This identification of thin sections delivers the necessary information for our 3D multilayer model presented previously. In particular, it suffices to identify only the top section of the GDL because, due to the production process and the visual impression of the 2D SEM images, all sections can be assumed to be identically distributed. Additionally, the top section can be determined most reliably. As a justification of these assumptions, we compare a structural characteristic of the fitted model and a 3D data set in the Spherical contact distribution function section.

A certain problem in detecting the individual thin sections is that they cannot be determined uniquely. Nevertheless, the approach presented in this section leads to plausible results for all 2D electron microscopy images considered in this paper (see also Fig. 14).

*Overlay factor*.— The general idea for separating the fibers into different sections is to regard their overlay factor. An attempt to detect different sections by visual inspection leads to the assumption that the fibers of the first section have an overlay factor of about 5%, i.e., at most 5% of their length is covered by other fibers.

The horizontal length of our 2D images is about 1750 $\mu$m. Furthermore, a fiber of the material considered in the present paper has a diameter of about 7 $\mu$m and, therefore, a horizontally oriented fiber in our images can be overlaid by about 13 other fibers and still be seen as a fiber of the first (i.e., top) section. The second section is then defined by an overlay factor between 5 and 10%, and so on. Note that for a fixed overlay factor the number of overlaying fibers is proportional to the possible maximal length. This criterion is now used for automatic detection of the different thin sections.

*Ranking of polygonal tracks*.— In order to determine the overlay factor for any given polygonal track, a ranking of all polygonal tracks is introduced, which has been extracted from the original (grayscale) image. The motivation for such a ranking is to describe how close a polygonal track is to the surface of the considered GDL. Therefore, two characterizing values of polygonal tracks are considered.

1. The first value is the percentage of that part of the track length which coincides with line segments from the skeletonized image (see Fig. 11a). This is based on the fact that fibers which are close to the surface lead to more and longer line segments in the skeletonization process than fibers which are overlain by more other fibers.
2. The second value is computed with the help of the smoothed image considered in the Anistropic diffusion section. For this purpose, the polygonal track is discretized and compared with the smoothed image. Then, the variance of the grayscale values at the pixels of the discretized polygonal track and its neighbor pixels is estimated. The smaller this variance, the larger the chance that the polygonal track belongs to the top section of the GDL. The idea for this criterion is that a fiber at the surface consists of pixels with almost the same grayscale values, whereas the fibers in the background have darker pixels in general but brighter pixels where they are crossed by fibers which are more near the surface. Therefore, the variance of the grayscale values is smaller for fibers at the surface than for fibers in the background. Then, a multiplicative linkage of these two characterizing values leads to a ranking of all polygonal tracks.

*Replacement of polygonal tracks by line segments*.— The polygonal tracks are considered to belong to the first section using the criterion based on the overlay factor introduced in the Overlay factor

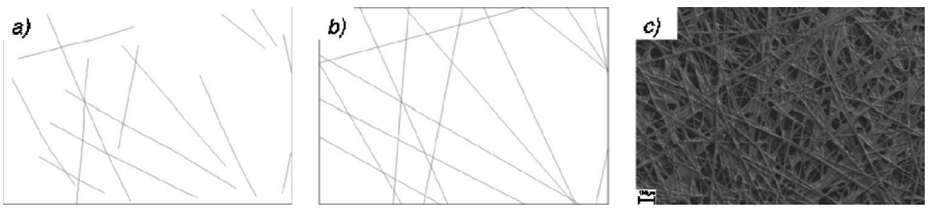


**Figure 13.** Transformation of polygonal tracks into line segments: (a) polygonal tracks, (b) line segments prolongated to the boundaries of the sampling window, and (c) original image of the GDL.

section. This overlay factor is determined by means of the ranking of all polygonal tracks as described in the Ranking polygonal tracks section.

In order to determine the overlay factor, the polygonal tracks are transformed into line segments which connect the end points of the polygon tracks (see Fig. 13a). In a further step, these line segments are prolongated until they run from boundary to boundary of the sampling window (see Fig. 13b). The ranking of the polygonal tracks is identified with the ranking of the corresponding line segments.

Now all those polygonal tracks which are not crossed by too many polygon tracks with a better ranking belong to the first section. To ensure that short polygonal tracks, which are obviously not crossed by many other polygonal tracks, are not added to the first section, we only take into account those polygonal tracks with a relative length larger than a certain threshold. The relative length is computed as the ratio of the length of the line segment which replaces the polygonal track and the maximal possible length of this line segment, i.e., the length of the line segment prolongated to the boundaries of the sampling window. Note that the prolongated line segments are used in the Statistical Model Fitting section below in order to estimate the intensity $\gamma$ of the underlying PLT.

Finally, the top section of our multilayer model is defined as follows. At first the line segment with the highest ranking is considered to belong to the top section. Then the next line segment in the ranking is regarded and so on. The currently considered line segment is added to the top section if it is not covered by too many line segments already belonging to the top section. Note that the absolute number of admissible overlappings depends on the length of the extended line segment, e.g., a line segment with extended length of 2000 $\mu$m can be overlaid by more fibers than a segment with extended length of 150 $\mu$m to have the same overlay factor.

The correctness of this technique of thin section detection was tested by a visual comparison of the detected top layer with the original 2D SEM image. Examples can be seen in the following section.

*Examples*.— In Fig. 14, three further examples of 2D electron microscopy images are shown, where it is clearly visible that in all three cases the segmentation and identification algorithm developed in previous sections leads to plausible results.

## Statistical Model Fitting

For simulation purposes, it is necessary to develop a method by means of which the proposed multilayer model can be fitted to given GDL data provided by SEM images. In other words, a method has to be developed to estimate the parameters of the model.

As mentioned above, our multilayer model has four parameters: $\gamma$, $2r$, $p$, and $n$, where $\gamma > 0$ is the intensity of the underlying PLT modeling the fibers of the first sublayer, $2r > 0$ is the diameter of the fibers, $p \in [0,1]$ is the filling probability which describes the amount of binder, and $n$ is the number of sublayers in the 3D model for the whole GDL.

The parameter $2r$ can be easily obtained from data. The thickness of a GDL is also a well-known or at least easy to measure characteristic. Then, the number $n$ of sublayers is just the ratio of the thickness of the GDL and the diameter $2r$ of the fibers. Furthermore, a natural estimator for $\gamma$ is based on counting the lines which are detected in the first section, according to the procedure described in the Identification of Thin Sections section. In order to determine the filling probability $p$, the following formulas for the volume fraction of fibers and binder, respectively, are useful.

*Formulas for the volume fractions of fibers and binder*.— It is not difficult to show that the volume fraction $V_{\text{fiber}}(\gamma)$ of the set of all (interpenetrating) dilated lines of the first section is a function of the intensity $\gamma$ of the underlying PLT, which is given by

$$V_{\text{fiber}}(\gamma) = 1 - \exp(-2r\gamma) \quad [5]$$

if the profile of the fibers is given as in Fig. 4c (see, e.g., Ref. 13 and 14). The volume fraction $V_{\text{binder}}(\gamma,p)$ of binder in the GDL is a function of $\gamma$ and $p$, which can be approximated by

$$V_{\text{binder}}(\gamma,p) \approx p\left(\frac{\pi}{\gamma^2} - r\frac{2\pi}{\gamma} + 2\pi r^2\right)\frac{\gamma^2}{\pi} \quad [6]$$

Note that

$$V_{\text{binder}}(\gamma,p) = pA(\gamma,r)\lambda_3(\gamma) \quad [7]$$

Here $A(\gamma,r)$ is the expected area of the typical cell of a PLT with intensity $\gamma$ and with dilated lines, where the structuring element is a ball with radius $r$ (see Fig. 15), and $\lambda_3(\gamma)$ denotes the mean number of cells per unit area. Furthermore, the expected area $A(\gamma,r)$ in Eq. 7 can be computed as the mean area $A(\gamma)$ of the typical cell of the (undilated) PLT minus the mean area which belongs to the dilated lines (see Fig. 15). The latter area can be approximated by $rL(\gamma) - \text{corr}(r)$, where $L(\gamma)$ is the mean boundary length of the typical cell and $\text{corr}(r)$ is a correction term because, in the expression $rL(\gamma)$, the area at the vertices is considered twice. Using a well-known formula for the distribution of the angles between two (arbitrarily chosen) intersecting lines of a PLT (see Ref. 15), the correction term $\text{corr}(r)$ can be approximatively determined as

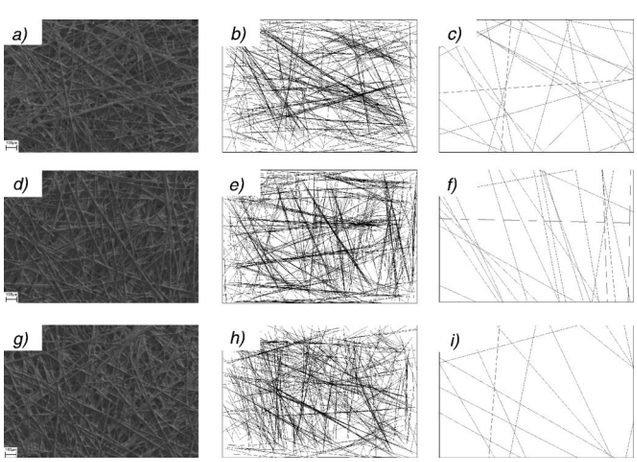


**Figure 14.** Sample images for segmentation and identification of the first section: (a) original images, (b) segmented images, and (c) detected first thin section.

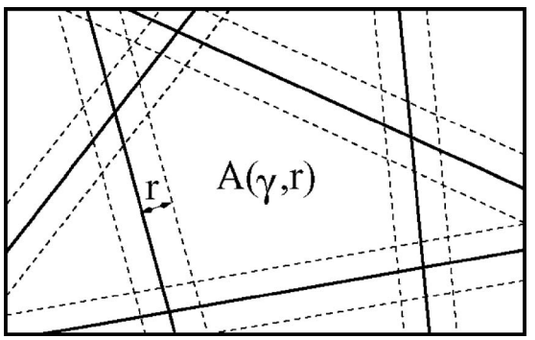


**Figure 15.** Cell of a dilated PLT.

$$\mathrm{corr}(r) \approx 2\pi r^2 \qquad [8]$$

Note, however, that this formula is not fully exact because it produces a (small) inaccuracy if the distance between two neighboring vertices is smaller than $2r$, but this is neglected in Eq. 8. Thus, altogether, Eq. 7 leads to the approximation formula

$$V_{\mathrm{binder}}(\gamma,p) \approx p[A(\gamma) - rL(\gamma) + 2\pi r^2]\lambda_3(\gamma)$$

and, inserting the mean-value formulas for $A(\gamma)$, $L(\gamma)$, and $\lambda_3(\gamma)$ given in Table I, we get that Eq. 6 holds.

Finally, the porosity $\varepsilon$ of the GDL, i.e., the volume fraction of the empty space of the GDL (not filled by fibers nor binder), is given by

$$\varepsilon = 1 - [V_{\mathrm{fiber}}(\gamma) + V_{\mathrm{binder}}(\gamma,p)] \qquad [9]$$

*Estimation of model parameters*.— In order to determine the intensity $\gamma$ from the SEM images, we consider the following estimator

$$\hat{\gamma} = \frac{\pi}{|\partial W|} \#\{n:(X_n,M_n) \cap W \neq \emptyset\} \qquad [10]$$

where $\#\{n:(X_n,M_n) \cap W \neq \emptyset\}$ is the number of lines detected in the first section of the GDL and $|\partial W|$ denotes the length of the boundary of the sampling window $W$. In Ref. 16 it is shown that the estimator $\hat{\gamma}$ given in Eq. 10 has various useful properties. In particular, $\hat{\gamma}$ is unbiased for $\gamma$, i.e., $\mathbb{E}\hat{\gamma} = \gamma$.

To determine the filling probability $p$ we proceed in the following way. First, we compute $\hat{\gamma}$ according to Eq. 10. Next, we insert the expressions for $V_{\mathrm{fiber}}(\hat{\gamma})$ and $V_{\mathrm{binder}}(\hat{\gamma},p)$ given in Eq. 5 and 6, respectively, into the right side of Eq. 9. We then can solve Eq. 9 for $p$ and, in this way, get an estimator $\hat{p}$ for $p$, provided that estimators $\widehat{2r}$ and $\hat{\varepsilon}$ for $2r$ and $\varepsilon$, respectively, are at hand. As we mentioned previously, the diameter $2r$ of the fibers is 5 pixels and the length of a pixel is 1.5 $\mu$m.

For the porosity of the GDL we use the value $\hat{\varepsilon} = 0.78$, which is given in the literature and by the manufacturer (Toray) (see, e.g., Ref. 17). We are aware that this (general) porosity must not necessarily coincide with the porosity of the concrete 3D sample considered here. However, because we used this 3D synchrotron tomographic image (see Fig. 16) for the validation of our (fitted) multilayer model, we estimated the porosity $\hat{\varepsilon}'$ of the segmented 3D image with

$$\hat{\varepsilon}' = \frac{\#\{\text{black voxels in the sampling window}\}}{\#\{\text{all voxels in the sampling window}\}} \qquad [11]$$

In this case, the segmentation is done interactively. At first a threshold is chosen by hand in such a way that in the binary image the fibers are well represented. Single voxels which belong to neither fibers nor binder and can therefore be seen as noise are removed by morphological opening, i.e., dilation and subsequent erosion. From this segmented 3D sample, the value $\hat{\varepsilon}' = 0.78$ has been obtained,

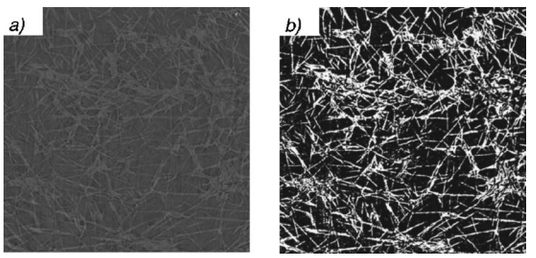


**Figure 16.** Thin slice of a GDL (from a 3D image generated by synchrotron tomography): (a) original image and (b) segmented image.

which coincides with the value given in the literature. Because of this good fit, the sample can be seen as a typical sample and is appropriate to use as a reference image for the validation. Moreover, we estimated the intensity $\gamma$ for 10 SEM images of a GDL material (see Table II) which leads to an average $\hat{\gamma} = 0.025$.

Solving Eq. 9 with $\hat{\gamma} = 0.025$ and $\hat{\gamma} = 0.78$, we get that $\hat{p} = 0.059$, i.e., the parameters $\hat{\gamma}$ and $\hat{p}$ of the multilayer model are specified. Moreover, inserting the values $\hat{\gamma} = 0.025$, $\hat{p} = 0.059$, and $2r = 7.5$ into Eq. 5 and 6, we get that

$$V_{\mathrm{fiber}}(\hat{\gamma}) = 0.17 \quad \text{and} \quad V_{\mathrm{binder}}(\hat{\gamma},\hat{p}) = 0.049 \qquad [12]$$

In other words, 17% of the considered GDL material consists of fibers and 5% of binder, which is in line with experimental findings.

We note that there are advantages in estimating parameter values from 2D SEM images instead of using measured/literature given values. With estimated parameter values a model can also be fitted to, e.g., a damaged GDL where the parameters are different from values given in the literature.

*Spherical contact distribution function*.— An important step in statistical model fitting is to check whether the fitted model sufficiently reflects the structural properties of image data obtained by real measurements. In particular, 3D synchroton-tomographic data and realizations of the model are compared. Those characteristics which are relevant in estimating the performance of the GDL should match sufficiently for both data sets.

An example of such an image characteristic is the distribution function of the spherical contact distance to the (white) foreground phase of binary images, i.e., the distribution of the nearest distance from an arbitrary point of the (black) background phase to the (white) foreground phase. This characteristic is closely related to the microstructure of the pore space of the GDL (see Fig. 16b). Mathematically, this spherical contact distribution function $H:[0,\infty) \to [0,1]$ is defined for any stationary random closed set $\Xi \subset \mathbb{R}^3$ by the formula

$$H(r) = P[\Xi \cap B(o,r) \neq \emptyset \,|o \notin \Xi] \quad r \geq 0 \qquad [13]$$

where $B(o,r) = \{x \in \mathbb{R}^3:|x| \leq r\}$ denotes the ball in $\mathbb{R}^3$ around the origin with radius $r \geq 0$. This means that $H(r)$ is the (conditional) probability that the distance to the boundary $\partial\Xi$ of the random set $\Xi$, seen from an arbitrarily chosen point outside of $\Xi$, is smaller than or equal to $r$. In our case, $\Xi$ is a model for the complement of pore space of the GDL, i.e., for the space filled by fibers or binder.

The following (edge corrected) estimator

$$\hat{H}_W(r) = \frac{\hat{p}_W(r) - \hat{p}_W(0)}{1 - \hat{p}_W(0)} \qquad [14]$$

of $H(r)$ is considered in the literature, where

**Table II. Estimated intensities of 10 2D SEM images of GDLs.**

| | | | | |
|---|---|---|---|---|
| 0.02166 | 0.02839 | 0.02839 | 0.02748 | 0.02656 |
| 0.02565 | 0.02015 | 0.02473 | 0.02107 | 0.02565 |

$$\hat{p}_W(r) = \frac{|W_r^* \cap [\Xi \oplus B(o,r)]|}{|W_r^*|} \quad \text{with} \quad W_r^* = W \ominus B(o,r) \qquad [15]$$

is an (edge corrected) estimator of the area fraction $p(r)$ of the dilated random set $\Xi \oplus B(o,r)$. In order to efficiently compute the estimator $\hat{H}_W(r)$ of $H(r)$ given in Eq. 14 for a sufficiently large number of distances $r \in [0,r_{\max}]$, one can proceed as follows. The numerator $\hat{p}_W(r) - \hat{p}_W(0)$ in Eq. 14 has to be computed for a sufficiently large number of distances $r \in [0,r_{\max}]$, where the following algorithm is used.

1. For each black pixel in the eroded window $W_r^* = W \ominus B(o,r)$, i.e., pixels which are not covered by $\Xi$, the distance to the closest pixel in $\Xi \cap W$ is computed.
2. The number of those pixels in $W_r^*/\Xi$ with a distance value smaller than or equal to $r \geq 0$ is determined.
3. The quotient of this number divided by the number of all pixels in $W_r^*$ provides a (discretized) approximation of $\hat{p}_W(r) - \hat{p}_W(0)$.

The denominator $1 - \hat{p}_W(0)$ in Eq. 14 is approximated by the quotient of the number of all pixels in $W/\Xi$ divided by the number of all pixels in $W$. Finally, the estimator $\hat{H}_W(r)$ of $H(r)$ is computed as the relative frequency of those pixels in $W_r^*/\Xi$ whose distance value is smaller than or equal to $r \geq 0$.

In order to assess the structural resemblance of the multilayer model where the binder is modeled with the Bernoulli-filling approach, we simulated this model in a cubic window $W$ of $1250 \times 1250 \times 100$ voxels. As parameters we used $\hat{\gamma} = 0.025$, $\hat{p} = 0.059$, and $2r = 5$, which have been obtained from 2D SEM images as described in a previous section. Then, using the algorithm described above, we computed the values $\hat{H}_W(r)$ of the (estimated) spherical contact distribution function for several $r \in [0,40]$, both for the simulated data and for a 3D image of the GDL material shown in Fig. 16 which has been obtained by means of synchrotron tomography.

The result can be seen in Fig. 17, which shows that both spherical contact distribution functions are similar. Moreover, the porosity $\varepsilon$ has been estimated according to Eq. 11 for the simulated data and the measured (synchrotron) data. The obtained values of $\hat{\varepsilon}$ also agree very well, giving $\hat{\varepsilon} = 77.3\%$ for the simulated data and $\hat{\varepsilon} = 78.2\%$ for the synchrotron data. This indicates that the stochastic multilayer model is a suitable description of the microstructure of the GDL material considered in the present paper.

## Conclusions

In this paper we propose a stochastic multilayer model in order to describe the microstructure of porous material for the GDL in PEMFCs. The model, based on stochastic geometry, offers at least two advantages.

First, the model is easy to handle because it depends on only four parameters, the intensity $\gamma$ of a (planar) PLT, the probability of binder filling $p$, the diameter $2r$ of the fibers, and the number $n$ of thin sections. Note that $n$ and $2r$ are either known material parameters or can be easily determined. Furthermore, $n$ and $2r$ are uniquely linked because $2rn$ is the thickness of the whole GDL.

Second, the statistical fitting of our multilayer model to real data is based on 2D SEM images, which can be obtained much easier by microscopic measurements than 3D (synchrotron-tomographic) images of the GDL.

For purposes of model fitting, we developed an algorithm for the segmentation of 2D SEM images of the GDL. Moreover, a procedure is proposed to detect thin sections from the segmented SEM images. In this way, estimators for the remaining two parameters $\gamma$ and $p$ of the multilayer model can be constructed.

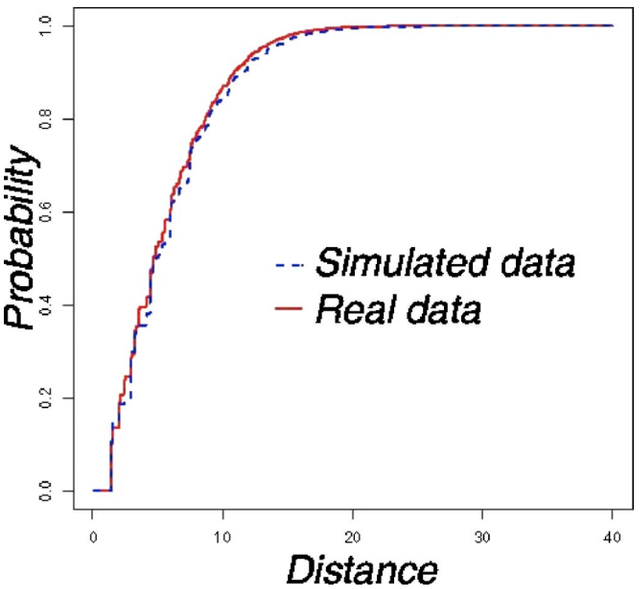


**Figure 17.** (Color online) Estimated spherical contact distribution functions for real and simulated data where the Bernoulli-filling approach is used for binder modeling.

The quality of our structural model has been assessed by simulating the fitted model in a cubic window. Then, applying the algorithm mentioned above, the distribution function of the spherical contact distance to fibers and binder has been determined, seen from an arbitrarily chosen point of the pore space. This has been done for the simulated data and for a 3D synchrotron-tomographic image of the GDL material. It turns out that both spherical contact distribution functions are similar. This indicates that our model describes the microstructure of the considered GDL material in an appropriate way.

## Acknowledgments

This research has been supported by the Forschungsallianz Brennstoffzelle Baden-Württemberg (FABZ). We thank Natalie Grünwald for her help in the implementation of the estimator for the spherical contact distribution function. Bernhard Clasen has kindly provided us with the segmentation of the 3D image in Fig. 16.

*Center for Solar Energy and Hydrogen Research Baden-Wüerttemberg assisted in meeting the publication costs of this article.*

## References

1. M. F. Mathias, J. Roth, J. Fleming, and W. Lehnert, *Handbook of Fuel Cells*, Vol. III, W. Vielstich, A. Lamm, and H. Gasteiger, Editors, Chap. 42, J. Wiley & Sons, Chichester (2003).
2. I. Manke, C. Hartnig, M. Grünerbel, W. Lehnert, N. Kardjilov, A. Haibel, A. Hilger, and J. Banhart, *Appl. Phys. Lett.*, **90**, 174105-1 (2007).
3. V. P. Schulz, J. Becker, A. Wiegmann, P. P. Mukherjee, and C.-Y. Wang, *J. Electrochem. Soc.*, **154**, B419 (2007).
4. P. K. Sinha, P. P. Mukherjee, and C.-Y. Wang, *J. Mater. Chem.*, **17**, 3089 (2007).
5. K. Schladitz, S. Peters, D. Reinel-Bitzer, A. Wiegmann, and J. Ohser, *Comput. Mater. Sci.*, **38**, 56 (2006).
6. A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu, *Spatial Tessellations*, 2nd ed., J. Wiley and Sons, Chichester (2000).
7. R. Schneider and W. Weil, *Stochastische Geometrie*, Teubner, Stuttgart (2000).
8. D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic Geometry and its Applications*, 2nd ed., J. Wiley & Sons, Chichester (1995).
9. B. Jahine, *Digital Image Processing*, 6th ed., Springer, Berlin (2005).
10. P. Perona and J. Malik, *IEEE Trans. Pattern Anal. Mach. Intell.*, **12**, 629 (1990).
11. T. W. Ridler and S. Calvard, *IEEE Trans. Syst. Man Cybern.*, **SMC-8**, 630 (1978).
12. M. Beil, H. Braxmeier, F. Fleischer, V. Schmidt, and P. Walther, *J. Microsc.*, **220**,

84 (2005).
13. P. Davy, Doctoral Dissertation, Australian National University, Canberra, Australia (1978).
14. J. Ohser and F. Mücklich, *Statistical Analysis of Microstructure in Materials Science*, J. Wiley and Sons, Chichester (2000).
15. R. Miles, *Proc. Natl. Acad. Sci. U.S.A.*, **52**, 901 (1964).
16. A. J. Baddeley and L. M. Cruz-Orive, *Adv. Appl. Probab.*, **27**, 2 (1995).
17. J. T. Gostick, M. W. Fowler, M. A. Ioannidis, M. D. Pritzker, Y. M. Volfkovich, and A. Sakars, *J. Power Sources*, **156**, 375 (2006).