Random geometric graphs for modelling the pore space of fibre-based materials

Ralf Thiedmann · Ingo Manke ·
Werner Lehnert · Volker Schmidt

Abstract A stochastic network model is developed which describes the 3D morphology of the pore space in fibre-based materials. It has the form of a random geometric graph, where the vertex set is modelled by random point processes and the edges are put using tools from graph theory and Markov chain Monte Carlo simulation. The model parameters are fitted to real image data gained by X-ray synchrotron tomography. In particular, they are specified in such a way that the distributions of vertex degrees and edge lengths, respectively, coincide to a large extent for real and simulated data. Furthermore, the network model is used to introduce a morphology-based notion of pores and their sizes. The model is validated by considering physical characteristics which are relevant for transport processes in the pore space, like geometric tortuosity, i.e., the distribution of shortest path lengths through the material relative to its thickness.

R. Thiedmann
Ulm University, 89069 Ulm, Germany
Tel.: +49-731-50 23617
Fax: +49-731-50 23649
E-mail: ralf.thiedmann@uni-ulm.de

I. Manke
Helmholtz-Zentrum Berlin für Energie und Materialien, 14109 Berlin, Germany

W. Lehnert
Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

V. Schmidt
Ulm University, 89069 Ulm, Germany

1 Introduction

A stochastic network model is developed which is based on methods from stochastic geometry and spatial statistics; see [12] and [21] for comprehensive surveys on recent results in these fields. It describes the 3D morphology of pore systems in fibre-based materials and can be used for scenario analyses e.g. with the objective of developing improved materials and technologies for renewable energies. In particular, porous materials are considered where the solid phase consists of a rather complex system of curved fibres. They mainly run parallel to some fixed 2-dimensional plane, say the x-y plane, forming wafers with small thicknesses (along the z-axis) which can be seen as stacks of thin layers of fibres. Such nonwoven-type materials are used e.g. in the gas-diffusion layer (GDL) of polymeric fuel cells; see Figure 1.
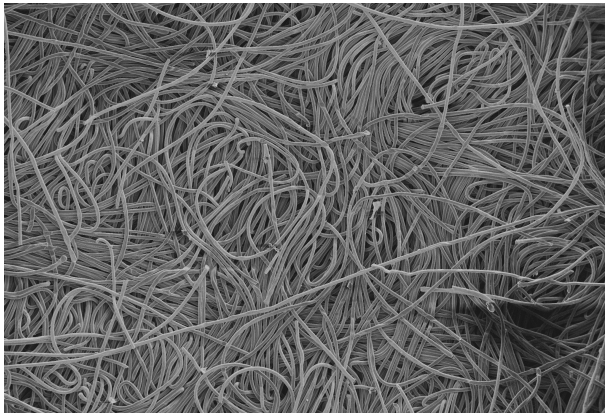

Fig. 1 2D SEM image of the considered nonwoven-type GDL

Recently, several models for the solid phase of GDL, in particular for the fibre system itself, have been proposed where the pore space is considered as complementary set ([16], [17], [36], [43], [48]). However, this indirect description of pore space often leads to very complex geometric structures, i.e., it is described by huge sets of voxels, which make numerical simulations of transport processes quite complicated and computer time consuming, especially for large domains.

In the present paper, a stochastic network model is developed which has the form of a random geometric graph, representing the pore space directly. It can be applied e.g. to investigate transport processes in GDL on a large scale. Furthermore, the model can be used to introduce a morphology-based notion of pores and their sizes. The model parameters are fitted to real image data gained by X-ray synchrotron tomography. In particular, they are specified in such a way that the distributions of vertex degrees and edge lengths, respectively, coincide to a large extent for real and simulated data.

The vertex set of the random geometric graph is constructed by a stack of 2D point processes, which can be seen as a point process in 3D, whose points have continuous x- and y-coordinates, but discrete z-coordinates. Then, by 'smearing' the z-coordinates in an appropriate way, a point process in 3D is obtained, where all three coordinates of its points are continuous random variables. These points can physically be interpreted as candidates for pore centres. Their minimum distances to the solid phase, so-called contact distances, can be seen as marks which describe pore sizes. Note that in the network extracted from synchrotron data, the contact distances of neighbouring vertices are strongly (positively) correlated. Thus, they cannot be modelled just by independent marking, but a certain moving-average procedure is proposed, which mimics this correlation structure quite well. For details concerning point processes in multidimensional spaces and their statistical inference and simulation, we refer e.g. to [4], [15], [28] and [41].

The edges are constructed combining tools from graph theory and Markov chain Monte Carlo (MCMC) simulation; see e.g. [22] and [45]. Candidates for vertex degrees, i.e. the numbers of edges outgoing from vertices, are sampled in an independent and identically distributed (iid) way, using the histogram of vertex degrees which has been computed from synchrotron data. This is followed by an acceptance-rejection procedure which ensures that the conditions of the Erdös-Gallai theorem are fulfilled, regarding the existence of graphs for a given configuration of vertex degrees. Then, for an admissible configuration of vertex degrees, edges are put using the well-known Hakimi-Havel algorithm of graph theory. However, this algorithm does not take into account the locations of vertices, which means that in general the distribution of edge lengths computed from synchrotron data is not fitted well. Thus, in order to minimize this discrepancy, the Hakimi-Havel algorithm is supplemented by an MCMC procedure to rearrange edges in such a way that the distribution of vertex degrees is kept fixed and, simultaneously, the fit of the empirical distribution of edge lengths computed from synchrotron data is improved.

Finally, the network model is validated by considering characteristics like the minimum spanning tree and geometric tortuosity, i.e., the distribution of shortest path lengths through the material relative to its thickness. It turns out that both characteristics coincide quite nicely for real and simulated data.

As already mentioned above, the network model developed in the present paper is motivated by computer-based scenario analyses with the general objective of developing improved materials and technologies for renewable energies. In particular, our model has been fitted to synchrotron data for nonwoven-type materials used e.g. in the GDL of fuel cells which is responsible for transport/diffusion of oxygen and hydrogen towards the electrode, where electricity is produced. Furthermore, as a by-product of the electrochemical processes in low temperature fuel cells, liquid water is produced which has to be drained off. Note that all these transport processes take place in the pore phase of the GDL, i.e., the phase which is not occupied by fibres or binder, see e.g. [27].

Recently, several models have been proposed in literature describing the solid phase of fibre-based GDL, i.e., the fibres themselves and the binder ([16],

[17], [36], [43], [48]). However, these models are focused on GDL materials of paper-type where the fibres can be approximated by straight lines. For the solid phase of nonwoven-type GDL with clearly curved fibres, no stochastic models are available so far which would describe their microstructure sufficiently well. However, there can also be found models for fibre-based materials and their analysis in the literature in other contexts, see e.g. [3], [10].

Moreover, an important disadvantage of the microstructure models existing so far for the solid phase of paper-type GDL is that the pore space is only described indirectly, as complement of the solid phase. This results in a description of the pore space by a huge set of voxels which complicates numerical computations especially with respect to run time and memory requirements. Furthermore, it restricts the size of domains in which the microstructure of GDL can be analysed by numerical computations. At first glance, an alternative could be to reduce the resolution of data, but this would coarsen the microstructure which causes inaccuracies. Thus, to avoid these conflicts between run time and accuracy, we propose a direct description of the pore space by 3D random geometric graphs. The advantages of this representation are manifold. First of all, the pore space is now described directly and numerical computations on the edges of a graph can be done relatively easily. In addition, the proposed graph model is off-grid, i.e., the computations on the graph can be realized in terms of Euclidean coordinates which do not depend on any given resolution.

We also mention that the idea to represent pore systems by 3D graphs is not completely new; see e.g. [5] and [44]. But no off-grid models exist so far which could be used for stochastic simulation and scenario analysis based on real 3D image data. On the other hand, some authors consider grid-based graph models for the pore space of GDL; see [13], [38] and [39]. But these graph models for the pore space do not take into account its real microstructure. The pores are located just on a grid and the models are calibrated with respect to global physical characteristics as, e.g., permeability. In contrast to this type of global model fitting, the model which we propose in the present paper is fitted to local microstructural characteristics of the pore space. They are computed from real 3D image data gained by means of X-ray synchrotron tomography; see [14] and [26].

In addition, fibre-based materials with other applications than GDL in fuel cells have been investigated with techniques from spatial statistics and stochastic geometry. For instance, a multi-layer model for the fibres of nonwoven mats with application to coalescers is discussed and analysed in [1]. Furthermore, several techniques have been developed to extract pore sizes from (2D and 3D) simulated fibre-based materials, wherefore specific properties of randomly placed fibres in the plane are used, see e.g. [9], [31], [33], [34].

The paper is organized as follows. In Section 2, the 3D image data are described which are used to fit the graph model. Then, in Section 3, the vertex model based on random point processes is explained. In Section 4, the marking of vertices is described, whereas in Section 5 the edge model is introduced.

Some issues of model validation are discussed in Section 6. Finally, Section 7 summarizes the results.

2 Preprocessing of image data

In this section, the 3D image data are described gained by means of X-ray synchrotron tomography. Then, their binarization and skeletonization is briefly explained. In particular, the construction of vertices and edges of graphs representing the pore space is described, and a morphology-based approach to the notions of pores and their sizes is given.

2.1 Data description

In order to fit a random graph model to the microstructure of real GDL materials, we use 3D image data which are gained by means of X-ray synchrotron tomography as described e.g. in [14] and [26]. These data are grey scale images which have to be preprocessed. This is done in the same way as in [43], i.e., first using a certain filter to smooth the data, then to binarize them by (global) thresholding, and, subsequently, using an opening to remove small objects which are not connected to the fibre system. The binarization threshold is chosen such that the estimated porosity in the resulting binary image, i.e., the volume fraction of the pore space, is about 75%. Further details on this type of morphological image processing can be found in e.g. [18], [30] and [40].

In order to keep the computational effort for extracting graphs from real data at a reasonable level, we consider cutouts of the original 3D data set. These cutouts are cuboids with $512 \times 512 \times 100$ voxels, which corresponds to images of size 768μm $\times$ 768μm $\times$ 150μm. To randomize the locations where the cutouts are taken we apply a bootstrap, i.e., the locations of cutouts are chosen at random; see e.g. [23]. We consider 50 such data sets as our sample drawn from the original (synchrotron) data. Notice that the corresponding cuboids do not have to be completely disjoint.

2.2 Extraction of graphs

Note that a 3D image of (segmented) synchrotron data is given as a stack of 2D binary images, i.e., the 3D information is given as a 3D matrix with entries being equal to 0 and 1 representing occupied voxels (solid phase) and pore space, respectively. To extract a graph from the voxelized pore space, a skeletonization of pore space is applied. The principal idea is to change voxels belonging to the pore space into background voxels in such a way that just a thin line is left over with thickness of one voxel, where the connectivity of the skeleton should be the same as the connectivity of the original pore space. An example in 2D is shown in Figure 2, where Figure 2(a) displays three objects (white). We are interested in skeletonization of the black phase between these

objects as our focus is directed to pore space. Figure 2(b) shows the skeleton of
the pore space. For this skeletonization in 3D, we use an algorithm described
in [11]. Subsequently, the skeleton is transformed into vector data by classi-
fying skeleton voxels into 'end voxels', 'line segment voxels', and 'junctions',
respectively, where all voxels of the skeleton with exactly one neighbour are
said to be end voxels, all voxels with exactly two neighbours are line segment
voxels, and all voxels with more than two neighbours are junctions. If a junc-
tion consists of more than one voxel, the centre of gravity is assumed to be
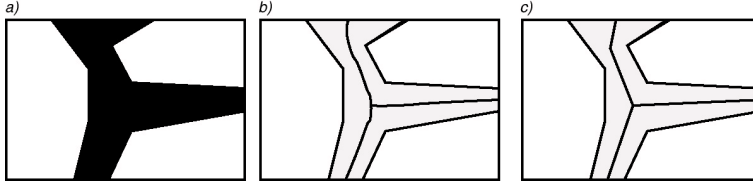the location of the junction.

Fig. 2 Skeletonization in 2D: a) pore space (black), b) skeletonized pore space (black line
in pore space), c) transformation into vector data

The end voxels and junctions form the vertices of the graph to be con-
structed. Connecting some pairs of them by line segments leads to a 3D graph.
An example in 2D is shown in Figure 2(c). Note that such a connection of ver-
tices will be represented by a polygonal track instead of just by one single
segment if the connection is not straight but curved.

To reduce boundary effects in fitting our model to the extracted graphs,
we apply a minus sampling, i.e., we neglect data which are too close to the
boundary of the sampling window.

2.3 Detection of pores

A problem in computing pore size distributions for materials with high poros-
ity (of about 75%) is the (unique) definition of pores as geometrical objects.
However, using the 3D graph introduced in Section 2.2, we can consider all
vertices of the graph as potential pore centres; see also [44]. The pore size is
then the spherical distance of such a pore centre to the solid phase. But, if
we took all vertices as pore centres, some pores would be contained partially
or completely in other pores. Therefore only those vertices are considered as
pore centres, which are not contained in larger pores, otherwise the number
of small pores would be overestimated.

The following algorithm is used to determine the pore size distribution.
For each potential pore, i.e. vertex of the graph, the spherical distance to the
solid phase is computed. This can be done very efficiently using a distance
transformation as described e.g. in [32]. The potential pores are then ordered

according to their sizes. Beginning with the largest pore, all other potential pores with pore centres belonging to that pore are deleted from the list of potential pores. Then, for the largest remaining (i.e. not yet deleted) pore the same procedure is realized, and so on. The result is a set of pores which can mutually overlap, but no pore contains a centre of another pore. An illustrating example is given in Figure 3. In Figure 3(a) all potential pores are shown, i.e., each vertex is seen as a potential pore centre and the balls around are the corresponding pores. In Figure 3(b) only those balls are shown which are classified as pores.
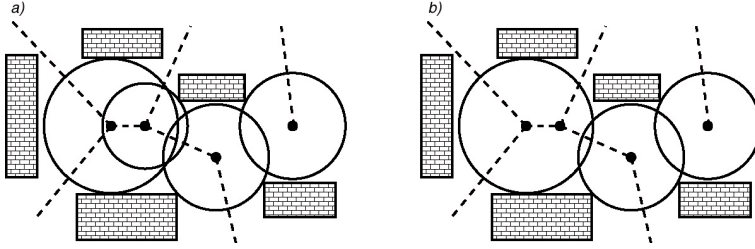

Fig. 3 Definition of pores: a) all vertices of the graph with spherical distances to the solid phase, b) deletion to those balls which are not classified as pores

The notion of pore size distribution is of special interest in electrochemistry, because characteristics of this type can be accessed directly from real GDL by porosimetric methods such as mercury or water porosimetry; see e.g. [2] and [25]. However, note that the results of physical porosimetric measurements do not coincide with the pore size distribution of graphs extracted from 3D images, because the analysis of e.g. mercury porosimetry results uses lots of assumptions about the structure of pores which are not fulfilled for real GDL materials; see e.g. [24]. A systematic comparison of our results for pore size distributions based on graphs extracted from 3D images with those obtained by porosimetric methods will be the subject of a forthcoming paper. Note that a similar attempt has been considered in [29].

2.4 Modified graph describing the pore system

According to the morphology-based definition of pores given in Section 2.3, we slightly modify the graph considered in Section 2.2, where we delete all those vertices which have not been classified as pore centres. This implies that those edges such that at least one of their endpoints is deleted, have to be changed as well. Note that these endpoints are then shifted towards the vertices classified as pore centres in whose pores they are located in. This is done in a way that all pores which were connected before are still connected in the modified

graph, see Figure 4(a). Furthermore, if there are some overlapping pores which have no common edge, we add such an edge to the graph, see Figure 4(b).
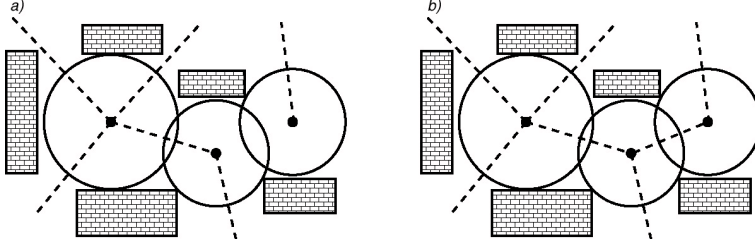

Fig. 4 Modification of the graph: a) elimination of vertices that are no pore centres, b) adding of edges if pores overlap

3 Stochastic modelling of vertices

The stochastic network model is constructed in two steps. First, the vertices of the random geometric graph are modelled which is described in this section. Then, for a given set of vertices, the edge set is constructed which is explained in Section 5.

3.1 Multi-layer representation

The basic idea for modelling the vertices of the 3D graph described in Section 2.2 is to use a multi-layer representation of vertices. This is motivated by the microstructure of real GDL; see Figure 5 which shows the profile of the fibre-based porous material. In Figure 5 it is clearly visible that the fibres are orientated (more or less) horizontally. Thus, they can be seen as a stack of thin layers formed by planar fibre systems. Therefore it is plausible to assume that also the complement of the fibres, i.e. the pore space, has such a multi-layer structure. Note that a similar approach has been used in [43] to model the solid phase of paper-type GDL.

The fibres of the nonwoven-type GDL considered in the present paper have a thickness of about 9 to 10μm. So we assume that the fibre system forms a stack of thin layers (parallel to the surface of the GDL), each with a thickness of 9μm. Furthermore, we decompose the 3D point pattern of vertices into the same type of thin layers, with the same thickness of 9μm. In order to model these layers of vertices we project all points of a given layer onto its base, being parallel to the x-y plane, say. These 2D point patterns are then the data basis for fitting our vertex model.

Fig. 5 Profile of a nonwoven-type GDL

3.2 Point-process model

For establishing an adequate point process model for the vertices we start with analysing the given set of vertices of the 3D graph introduced in Section 2.2. Therefore, we use the pair-correlation function (or g-function) $g : [0, \infty] \to [0, \infty)$ which can be interpreted as follows. For values r, where $g(r)$ is larger than 1, there are more point pairs with distance r to each other compared to a Poisson point process with the same intensity as the currently considered point process. If $g(r)$ is smaller than 1 for a value r, there are less point pairs with this distance and, in particular, if $g(r)$ vanishes for a value of r, there are no point pairs with such a distance. Thus, the pair-correlation function in Figure 6 estimated from the point pattern of vertices of the 3D graph introduced in Section 2.2 indicates strong clustering of vertices with an unusually high peak at small distances of about 4 to 5μm; see e.g. [15] for estimators of the g-function. This suggests the idea to fit a clustered point-process model with narrow and, simultaneously, elongated clusters.
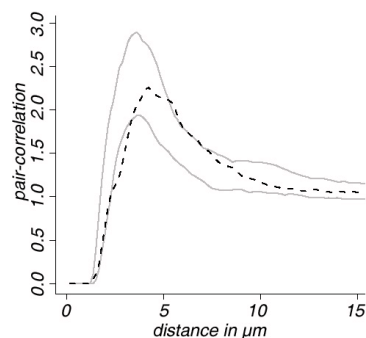

Fig. 6 Pair–correlation functions for real (dashed line) and simulated data (grey solid lines)

3.2.1 Generalized Thomas process and its pair-correlation function

As a model for the (projected) 2D point patterns described in Section 3.1, we thus use a generalized Thomas process with elliptically shaped clusters; see e.g. [7]. This cluster model has the following structure. The parent points form a stationary Poisson point process with intensity λ_p. The random number of child points per cluster is Poisson distributed with expectation c, and the random deviations of child points from their parent points are given via a 2D normal distribution $\mathrm{N}(o, C)$, with expectation vector o and covariance matrix

$$C = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \, .$$

Additionally, according to the uniform distribution on the interval $[0, 2\pi)$, the child points of each cluster are jointly rotated around their parent point. In this way, it is ensured that the generalized Thomas process is isotropic, although it possesses elliptically shaped clusters. Note that as points of the generalized Thomas process, only the child points are considered.

To fit the generalized Thomas process, its pair-correlation function $g_\theta : (0, \infty) \rightarrow [0, \infty)$ is considered, where $\theta = (\lambda_p, \sigma_1^2, \sigma_2^2)$. Note that the value $g_\theta(r)$ is proportional to the frequency of point pairs with distance $r > 0$ from each other. The following formula holds (see e.g. [7], [47]):

$$g_\theta(r) = 1 + \frac{1}{4\pi\lambda_p\sigma_1\sigma_2} \exp\left(-r^2 \frac{\sigma_1^2 + \sigma_2^2}{8\sigma_1^2\sigma_2^2}\right) I_0\left(r^2 \frac{\sigma_1^2 - \sigma_2^2}{8\sigma_1^2\sigma_2^2}\right) \, , \quad r \geq 0 \, ,$$

where I_0 denotes the modified Bessel function which can be evaluated by

$$I_0(z) = \sum_{k=0}^{\infty} \frac{(1/4z^2)^k}{(k!)^2} \, , \qquad z \in \mathbb{R} \, .$$

3.2.2 Model fitting

The pair-correlation function is estimated for overall 50 cutouts of synchrotron data, where a standard (boundary-corrected) estimator is used; see e.g. [15]. As already mentioned above, each cutout is divided into thin layers with a thickness of 9μm and the vertices are projected onto their bases. The pair-correlation function is then estimated for all these 2D data sets separately and the pointwise average of the estimated pair-correlation functions is computed which will be denoted by $\widehat{g}(r)$ in the following.

In order to fit the Thomas process to data, four parameters have to be determined: λ_p, c, σ_1^2 and σ_2^2, where a minimum-contrast method can be used with respect to the pair-correlation function. This means that the following minimization problem has to be solved:

$$f(\theta) = \int\limits_{r_1}^{r_2} \left(\widehat{g}(r) - g_\theta(r)\right)^2 \longrightarrow min$$

for an appropriately chosen pair $r_1, r_2 > 0$. Then a minimum–contrast estimator $\widehat{\theta} = (\widehat{\lambda}_p, \widehat{\sigma}_1^2, \widehat{\sigma}_2^2)$ for θ is given by $\widehat{\widehat{\theta}} = \arg\min_\theta f(\theta)$. The mean number c of child points per cluster is estimated using the formula

$$\widehat{c} = \frac{\widehat{\lambda}}{\widehat{\lambda}_p},$$

where $\widehat{\lambda}$ denotes the natural estimator of the over all intensity λ which can be estimated quite easily, just by counting the number of all points in the sampling window divided by its volume. On a scale where one distance unit corresponds to $1.0\mu m$, the result of this fitting is $\widehat{\lambda}_p = 0.000533$, $\widehat{c} = 2.28$, $\widehat{\sigma}_1^2 = 4.5$, $\widehat{\sigma}_2^2 = 78.75$. Thus, the estimated variances $\widehat{\sigma}_1^2$ and $\widehat{\sigma}_2^2$ are rather different, which means that the fitted Thomas process has clusters with clearly elongated shapes.

3.2.3 'Smearing' of points along the z-axis

Finally, the projection of vertices in z-direction onto the bases of thin layers mentioned in Section 3.1 has to be reversed. To incorporate this reversal step into the vertex model, we proceed in the following way.

Note that besides clustering, a certain hard-core effect is observed in the point pattern of vertices of the 3D graph introduced in Section 2.2. This is a result of the skeletonization and transformation into vector data, respectively, because possible vertices which are too close together are identified as one single vertex. Therefore, also in the vertex model, a (small) hard–core distance has to be included. Furthermore, analysing the z-coordinates observed in the point pattern of vertices of the 3D graph, it can be seen that they are almost uniformly distributed, see Figure 7. Besides, looking at the pair-correlation function given in Figure 6, we see that there are many point pairs with a distance of about 4 to $5\mu m$. In order to incorporate all these structural properties into the 3D vertex model, we do not shift the points of the 2D Thomas processes independently from each other along the z-axis. But we apply a dependent shifting which is based on the following property of exponential distribution: For any fixed $k \geq 1$, let $Z_1, \ldots, Z_k \sim \text{Exp}(\frac{1}{k})$ be independent and exponentially distributed random variables. Then $\min\{Z_1, \ldots, Z_k\} \sim \text{Exp}(1)$ and, therefore, $\exp(-\min\{Z_1, \ldots, Z_k\}) \sim \text{U}(0, 1)$.

We use this property for $k = 4$. Thus, considering a sample of a Thomas process which has $n > 0$ points in the sampling window, we associate these points with independent random variables $Z_1, \ldots, Z_n \sim \text{Exp}(\frac{1}{4})$ and, for the ith point of these n points, $i = 1, \ldots, n$, we consider its three closest neighbours with indices $i_1, i_2, i_3 \in \{1, \ldots, n\} \setminus \{i\}$, say. Then, we shift the ith point within the corresponding layer along the z-axis, according to $\exp(-\min\{Z_i, Z_{i_1}, Z_{i_2}, Z_{i_3}\}) \sim \text{U}(0, 1)$ (suitably scaled to the thickness of the layer). This dependent shifting along the z-axis ensures that the principal structure of the clustered Thomas processes does not change.
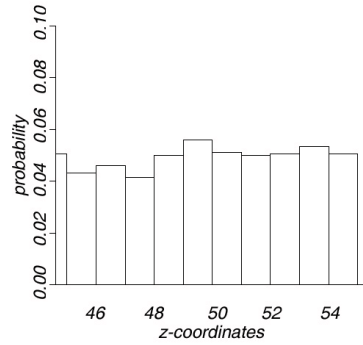
Fig. 7 Estimated distribution of z–coordinate of vertices for some layers

Finally, to incorporate a hard-core distance into the model, we apply a subsequent shift of the points along the z-axis if two points are too close to each other. Therefore we look at that pair of points of the complete 3D point pattern which are closest to each other and choose one of these two points at random. This point is then again shifted along the z-axis within the corresponding layer, according to a uniformly distributed random variable. This is repeated until the required hard-core distance of 3μm is achieved for all points, or if no further improvement is possible.

3.3 Model validation

In order to validate the point-process model proposed in Section 3.2, we consider two different characteristics of stationary point processes: the distribution function of (spherical) contact distances $H : [0, \infty) \to [0, 1]$, and the nearest-neighbour-distance distribution function $D : [0, \infty) \to [0, 1]$. Note that $H(r)$ is the probability that the distance from an arbitrary location in $\mathbb{R}^3$, chosen at random, to the closest point of the point process is not larger than r, $r > 0$. Similarly, $D(r)$ is the probability that the distance from an arbitrary point of the point process, chosen at random, to its nearest neighbour within the point process is not larger than r, $r > 0$.

Furthermore, we show that the pair-correlation functions computed from real and simulated 3D point patterns, respectively, are quite similar to each other.

To verify whether the 3D point-process model fits real data sufficiently well, we estimate $H(r)$ and $D(r)$ for all 50 cutouts of vertex sets extracted from synchrotron data, where standard (boundary-corrected) estimators are used; see e.g. [15]. The pointwise averages of these estimates are denoted by $\widehat{H}(r)$ and $\widehat{D}(r)$, respectively. Then, we compute pointwise 96% confidence bands for the two point-process characteristics mentioned above, where we generate 50

samples of the 3D point-process model with the estimated parameters as given in Section 3.2.2 in an sampling window of $768\mu m \times 768\mu m \times 150\mu m$. These bands are plotted as grey solid lines.
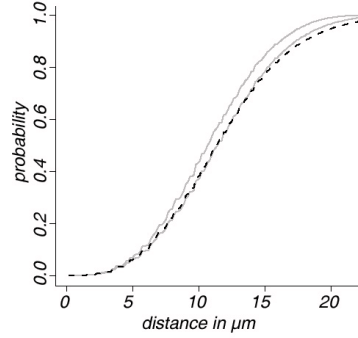

Fig. 8 Spherical–contact–distance distribution functions for real (dashed line) and simulated data (grey solid lines)

The results for $H(r)$ are visualized in Figure 8 which shows that the empirical distribution function $\widehat{H}(r)$ computed from real data (plotted as black dashed line) is more or less within the confidence band obtained from simulated data (grey solid lines). Furthermore, the results for $D(r)$ are given in Figure 9. Also for this characteristic the estimates $\widehat{D}(r)$ (black dashed line) are within the confidence band obtained from simulated data (grey solid lines).

Regarding the pair-correlation function, the estimate $\widehat{g}(r)$ which has been computed for the 3D vertex sets extracted from synchrotron data, does not match the confidence band of simulated data perfectly; see Figure 6. However, the main structural properties of $\widehat{g}(r)$ as the hard–core distance, the large peak at about $4\mu m$, and the declining rate of the tail towards the level of 1 are not too different from corresponding properties of the pair-correlation function computed from simulated data.

Considering all three characteristics together, we can conclude that the 3D vertex model introduced in Section 3.2 fits quite well to real data described in Section 2.

4 Marked point processes

In the preceding section we introduced a stochastic point-process model for the vertex set itself extracted from synchrotron data. In order to describe the pore space in more detail, we now extend this point-process approach to marked point processes, considering two different types of marks: the spherical distances of vertices to solid phase, and the numbers of edges emanating from
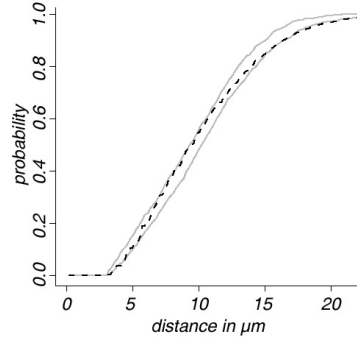
Fig. 9 Nearest–neighbour–distance distribution functions for real (dashed line) and simulated data (grey solid lines)

vertices. In the first case the marks are closely related with the notion of pore sizes introduced in Section 2.3, whereas the second kind of marks is the degree of vertices, also called the coordination number in physics and geology. Later on, in Section 5, the distribution of vertex degrees will be used in order to model the edges of the random geometric graph.

4.1 Spherical contact distances

In case of spherical contact distances of vertices, we fit a gamma distribution as their (Palm) mark distribution. However, it turns out that in the network extracted from synchrotron data, the contact distances of neighbouring vertices are strongly (positively) correlated. Thus, they cannot be modelled just by independent marking, but a certain mowing-average procedure is proposed, which mimics this correlation structure quite well.

4.1.1 Data analysis

In a first step we analyse the spherical contact distances to solid phase for the vertices extracted from synchrotron data. Their histogram is shown in Figure 10. It can be nicely fitted by a gamma distribution $\Gamma(\rho, \zeta)$ with parameters $\rho > 0$ (rate) and $\zeta > 0$ (shape) using the method of moments, see e.g. [6]. Its density function $f_{\Gamma(\rho,\zeta)} : \mathbb{R} \to [0, \infty)$ is given by $f_{\Gamma(\rho,\zeta)}(x) = \mathbb{1}_{\{x \geq 0\}} \frac{\rho^\zeta}{\Gamma(\zeta)} x^{(\zeta-1)} \exp(-\rho x)$. More precisely, let $x_1, \ldots, x_n$ be the observed spherical distances, then the estimator for ρ is given by $\widehat{m}_1/(\widehat{m}_2 - (\widehat{m}_1)^2)$ and for ζ by $(\widehat{m}_1)^2/(\widehat{m}_2 - (\widehat{m}_1)^2)$ with $\widehat{m}_k = \frac{1}{n} \sum_{i=1}^{n} x_i^k$, $k = 1, 2$. For the parameters of this gamma distribution (black curve in Figure 10), the averaged values of $\rho = 1.077$ and $\zeta = 7.331$ have been obtained, where the averages extend over all 50 cutouts from synchrotron data. Note that due to the estimation of the spherical distances from 3D image data with a resolution of

1.5μm, i.e., we can only observe discrete values, we summarized the results in a histogram with only some bins which can be seen as a smoothing of the data.
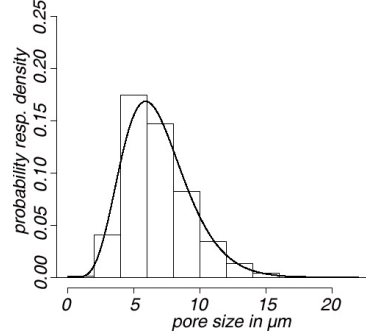

Fig. 10 Histogram of spherical distances of vertices to the solid phase and density function of the fitted gamma distribution (black solid line)

Furthermore, the correlation structure of spherical contact distances has been analysed using the so-called mark correlation function $\kappa : (0, \infty) \rightarrow [-1, 1]$ of stationary marked point processes, where $\kappa(r)$ is the correlation of the marks of an arbitrary pair of points, chosen at random, with distance $r > 0$ from each other. Note that although κ is not a correlation function in the strict sense, see the discussion in [35], we use this notion (introduced in [42]) since it can be found in various standard textbooks like [15]. Similar to the estimation of the functions $g(r)$, $H(r)$, and $D(r)$ discussed in Sections 3.2 and 3.3, the mark correlation function is estimated for all 50 cutouts extracted from synchrotron data; see e.g. [15]. The pointwise average of these estimates is denoted by $\widehat{\kappa}(r)$. It is shown in Figure 11 (black dashed line) and can be interpreted as follows. Vertices which are located close to each other have strongly (positively) correlated contact distances and, vice versa, the spherical contact distances of vertices which are far away from each other are more or less uncorrelated.

4.1.2 Moving-average model for dependent marking

To incorporate the correlation structure mentioned above into the model, we proceed similar as in Section 3.2.3, now using the fact that the family of gamma distributions possesses a well known stability property with respect to convolution.

If there are $n > 0$ points in the sampling window, we first associate these points with independent random variables $Z_1, \ldots, Z_n \sim \Gamma(\rho, \zeta/3)$, distributed according to the gamma distribution shown in Figure 10, where $\rho = 1.077$ and
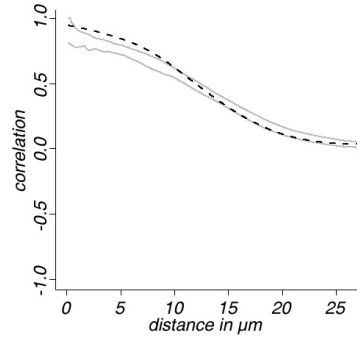
Fig. 11 Mark–correlation functions for spherical distances of real (dashed line) and simulated (grey solid lines) data

$\zeta = 7.331$. Then, for the ith point of these n points, $i = 1, \ldots, n$, we consider its three closest neighbours with indices $i_1, i_2 \in \{1, \ldots, n\} \setminus \{i\}$, say. As mark of the ith point, we finally choose the sum $Z_i + Z_{i_1} + Z_{i_2}$, where we use the fact that the sum of independent gamma distributed random variables is again gamma distributed. More precisely, it holds that

$$Z_i + Z_{i_1} + Z_{i_2} \sim \Gamma(\rho, \zeta) \,.$$

This dependent marking of points ensures that the principal structure of the empirical mark correlation function $\widehat{\kappa}(r)$ computed in Section 4.1.1 is captured quite well; see Figure 11. In this figure, pointwise 96% confidence bands are shown (grey solid lines), which were computed from 50 samples of the 3D point-process model with the moving-average marking as described above.

4.1.3 Detection of pores

For simulated vertex sets and their spherical contact distances, sampled from the model of a marked point process as described in Sections 3.2 and 4.1.2, we can proceed exactly in the same way as in Section 2.3 in order to detect pore centres within the vertex set of potential pore centres. Furthermore, for a given subset of vertices detected as pore centres, we can compute the distribution of their marks, i.e. pore sizes. Thus, still another possibility of model validation is given, comparing this distribution with the pore size distribution which has been computed in Section 2.3 for real data.

At first glance, one might think that the edge set of the random geometric graph to be constructed could be built in a similar way, directly from simulated vertex sets and their spherical contact distances. For example, just by connecting all those pairs of pore centres whose distances from each other are smaller than the sum of their pore radii. However, this happens only for a few pairs of pore centres, i.e., only a few pores overlap mutually. This means that

most edges of the random geometric graph to be constructed should be covered by pores only partially, whereas their middle parts can be interpreted e.g. as 'throats' between pores. Therefore, in order to build an edge model, we used another approach combining tools from graph theory and MCMC simulation, which will be explained in the following sections.

4.2 Degrees of vertices

An important characteristic for describing the connectivity of a graph are the degrees or coordination numbers of its vertices, i.e., the number of edges emanating from vertices. Thus, we now consider the vertex degrees as marks for the point-process model of vertices introduced in Section 3.2. However, our analysis is not directly based on the graph extracted from synchrotron data as described in Section 2.2, but on the modified graph of pore centres introduced in Section 2.4.

4.2.1 Data analysis

We first computed the empirical distribution of vertex degrees for the modified graph, which is shown in Figure 12(a). Then, we computed the mark-correlation function of vertex degrees, in the same way as this has been described in Section 4.1.1 for spherical contact distances. The result is presented in Figure 12(b), which shows that there is almost no correlation between vertex degrees. Thus, at first glance, it seems that the degrees of vertices could be modelled in an iid way, according to the distribution shown in Figure 12(a). However, this would get into conflict with the fact that not for each configuration of vertex degrees, a graph can be constructed. A possible solution of this problem is to combine iid sampling from the distribution shown in Figure 12(a) with a certain acceptance-rejection procedure which leads to admissible configurations of vertex degrees.
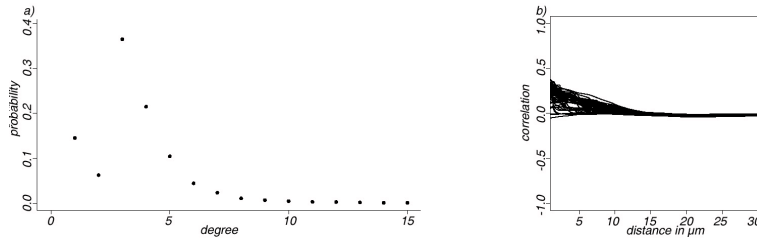

Fig. 12 Coordination number analysis: a) estimated distribution of the coordination number; b) estimated mark–correlation functions

4.2.2 Admissible configurations of vertex degrees

In order to solve the problem mentioned in the previous section, we propose a two-step algorithm for generating admissible configurations of vertex degrees. Supposing that the sample of the random graph to be constructed has n vertices in the sampling window, we generate an iid sample $d_1, \ldots, d_n > 0$ of candidates for vertex degrees according to the distribution shown in Figure 12(a). Then, using the Erdös–Gallai theorem of graph theory, see e.g. [45], [46], we check whether $d_1, \ldots, d_n$ is an admissible configuration of vertex degrees. For this purpose, we rearrange the numbers $d_1, \ldots, d_n > 0$ in descending order getting the sequence $d'_1 \geq d'_2 \geq \ldots \geq d'_n > 0$, say. In accordance with the Erdös–Gallai theorem, a simple graph, i.e., each pair of vertices has at most one direct connecting edge, can be constructed possessing the configuration $d_1, \ldots, d_n > 0$ of vertex degrees if and only if $d_1 + d_2 + \ldots + d_n$ is even, and for all $k = 1, \ldots, s$, where s is determined by

$$d'_s \geq s \qquad \text{and} \qquad d'_{s+1} < s + 1,$$

it holds that

$$\sum_{i=1}^{k} d'_i \leq k(k-1) + \sum_{i=k+1}^{n} \min\{k, d'_i\}.$$

If the sequence $d_1, \ldots, d_n > 0$ of potential vertex degrees fulfils these conditions, we can construct a graph with n vertices and vertex degrees $d_1, \ldots, d_n > 0$; see Section 5.1 below. Otherwise, we reject the sample $d_1, \ldots, d_n > 0$ and generate a new one according to the distribution shown in Figure 12(a). This procedure is repeated until a sequence of vertex degrees is generated which fulfils the conditions of the Erdös–Gallai theorem.

5 Stochastic modelling of edges

We now describe a stochastic model for putting the edges of the random geometric graph which combines tools from graph theory and MCMC simulation. In particular, the model is constructed in such a way that the distributions of vertex degrees and edge lengths, respectively, coincide to a large extent for real and simulated data.

As already mentioned in Section 4.2, candidates for vertex degrees are sampled in an iid way, according to the distribution shown in Figure 12(a). Recall that this is followed by an acceptance-rejection procedure which ensures that the conditions of the Erdös-Gallai theorem are fulfilled.

Then, assuming that an admissible configuration of vertex degrees is given, edges are put using the well-known Hakimi-Havel algorithm of graph theory. However, this algorithm does not take into account the locations of vertices, which means that in general the distribution of edge lengths computed from synchrotron data is not fitted well. In order to minimize this discrepancy, the Hakimi-Havel algorithm is supplemented by an MCMC procedure to rearrange

edges in such a way that the distribution of vertex degrees is kept fixed and, simultaneously, the fit of the empirical distribution of edge lengths computed from synchrotron data is improved.

5.1 Hakimi-Havel algorithm

Suppose that the considered sample of the random graph to be constructed has n vertices in the sampling window and that an admissible configuration $d_1, \ldots, d_n > 0$ of vertex degrees is given which has been sampled in an iid way, according to the distribution shown in Figure 12(a). Furthermore, suppose that the integers $d_1, \ldots, d_n > 0$ are numbered in descending order, i.e., $d_1 \geq d_2 \geq \ldots \geq d_n > 0$.

Then, a preliminary version of the edge set is constructed using an algorithm which is based on the classical Hakimi–Havel theorem, see e.g. [45]. This theorem states that there exists a simple graph with n vertices and degree sequence $d_1 \geq d_2 \geq \ldots \geq d_n > 0$ if and only if there exists one with $n - 1$ vertices and degree sequence $d_2 - 1, \ldots, d_{d_1+1} - 1, d_{d_1+2}, \ldots, d_n$.

This leads to the following algorithm: Suppose that a sequence $d_1 \geq d_2 \geq \ldots \geq d_n > 0$ of vertex degrees is given, which fulfils the conditions of the Erdös-Gallai theorem. Assign the vertex degrees $d_1, \ldots, d_n$ at random to the n vertices. Connect the vertex with degree d_1 with those vertices having degrees $d_2, \ldots, d_{d_1+1}$. Then, pick a vertex with the largest residual (i.e. free) degree, say d', and connect this vertex with d' of other vertices, with which it is not yet connected and which have the largest residual degrees. Repeat this procedure until no free degrees are left.

The result of this construction is a random graph whose distribution of vertex degrees fits to the corresponding empirical distribution of vertex degrees estimated from synchrotron data; see Section 4.2.1. However, note that the algorithm described above does not take into account the locations of vertices, which means that in general the distribution of edge lengths observed in real image data may not be fitted very well. Thus, in order to minimize this discrepancy, the Hakimi-Havel algorithm is supplemented by an MCMC procedure to rearrange edges in such a way that the distribution of vertex degrees is kept fixed and, simultaneously, the fit of the empirical distribution of edge lengths computed from synchrotron data is improved.

5.2 Distribution of edge lengths

In Figure 13(a) the empirical distribution of edge lengths is shown which has been computed for the modified graph extracted from synchtrotron data; see Section 2.4. It turned out that a shifted gamma distribution $\Gamma(\rho, \zeta, c)$ with parameters $\rho = 2.10$ (rate), $\zeta = 0.11$ (shape), and $c = 4.5\mu$m (shift) can be fitted to this empirical distribution, using maximum-likelihood estimation or the method of moments, see e.g. [6].

Thus, the goal is to construct a random geometric graph which fits sufficiently well both the distribution of vertex degrees considered in Section 4.2.1 and the shifted gamma distribution of egde lengths mentioned above.
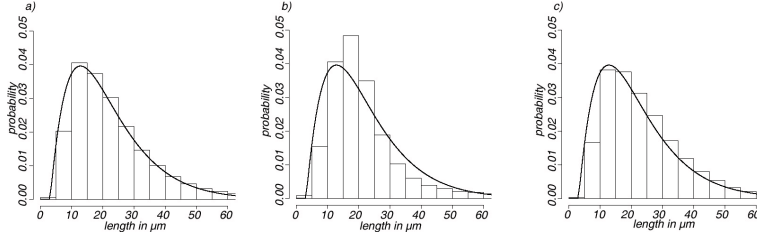

Fig. 13 Edge length analysis. Images show the length distribution for edges from a) the modified graph and a (shifted) fitted gamma distribution (black solid line); b) the graph generated by the HH–algorithm; c) the graph after MCMC–simulation

5.3 Rearrangement of edges

The graph constructed in Section 5.1, using the Hakimi-Havel algorithm, does not fit the empirical edge-length distribution sufficiently well which has been computed from real data; see Figure 13(b). Thus, in order to get a better fit, the Hakimi-Havel algorithm is supplemented by an MCMC procedure which rearranges the edges in such a way that the distribution of vertex degrees is kept fixed and, simultaneously, the fit of the empirical distribution of edge lengths computed from synchrotron data is improved. For further details on Markov chains and MCMC simulation see e.g. in [22].

Our MCMC algorithm is based on the following idea. Let $V = \{v_1, \ldots, v_n\}$ denote the vertex set considered, and let $E = \{e_1, \ldots, e_k\}$ be an edge set generated by the Hakimi-Havel algorithm, say. Furthermore, for any not directly connected pair of edges e_i, e_j, where $i \neq j$, we denote their endpoints by v_{i_1}, v_{i_2} and v_{j_1}, v_{j_2}, respectively. Then we consider all (three) possible connections of the four vertices $v_{i_1}, v_{i_2}, v_{j_1}, v_{j_2}$ by not directly connected pairs of edges and evaluate the suitability of these edge pairs. Eventually, the original pair of edges is replaced by another pair of edges which is evaluated better, where the evaluation depends on the distance between the (empirical) length distribution of the current set of edges and the shifted gamma distribution obtained in Section 5.2.

More precisely, in order to decrease the discrepance between the (empirical) edge-length distribution of the current set of edges and the shifted gamma distribution fitted in Section 5.2 to real data, we consider two different Markov chains. First, we run an auxiliary Markov chain that eliminates all those edges which are too long, replacing them by shorter ones. This Markov chain is

defined in the following way. An edge, say e_i, of the edge set $E = \{e_1, \ldots, e_k\}$ is chosen with probability

$$p_i = \frac{|e_i|}{\max_{1 \leq \ell \leq n}|e_\ell|} \,, \qquad i = 1, \ldots, k \,,$$

and another edge, say e_j, $i \neq j$, among those edges which are not directly connected to e_i is chosen with probability $p_j = |e_j|/\max_{1 \leq \ell \leq n}|e_\ell|$, where $|e|$ denotes the length of edge e. Subsequently, all three possible edge combinations which can be constructed by the endpoints $v_{i_1}, v_{i_2}, v_{j_1}, v_{j_2}$ of the chosen edges e_i, e_j are evaluated, where the 'value' η_{ij} of a pair of edges e_i, e_j is defined as the sum of their (current) selection probabilities, i.e. $\eta_{ij} = p_i + p_j$. If any of the other two possible edge pairs has a smaller value than e_i, e_j, then the edges e_i, e_j are replaced by that pair of edges. Otherwise, the pair e_i, e_j is not replaced. This procedure is continued as long as no further improvement is found.

Note that the evaluation of edges by this Markov chain just depends on their lengths, preferring edges which are short. In other words, we replace a pair of edges e_i, e_j if one of the other two, not directly connected pairs of edges has a shorter summary length. The result is a graph where edges are as short as possible. Furthermore, the MCMC algorithm described above helps to avoid numerical problems in computing the selection probabilities of the following (main) Markov chain which is defined in a similar way.

Let $f_\Gamma(r)$ denote the density of the shifted Γ–distribution derived in Section 5.2. The empirical edge-length density of the current configuration of edges is denoted by $\widehat{f}(r)$. Then an edge e_i is chosen with probability

$$\widetilde{p}_i = \max\left\{0, 1 - \frac{f_\Gamma(|e_i|)}{\widehat{f}(|e_i|)}\right\}$$

and another edge e_j among those edges which are not directly connected to e_i is chosen with probability $\widetilde{p}_j = \max\{0, 1 - f_\Gamma(|e_j|)/\widehat{f}(|e_j|)\}$. Subsequently, all three possible edge combinations which can be constructed by the endpoints $v_{i_1}, v_{i_2}, v_{j_1}, v_{j_2}$ of the chosen edges e_i, e_j are evaluated, where the 'value' $\widetilde{\eta}_{ij}$ of a pair of edges e_i, e_j is defined as the sum of their (current) selection probabilities, i.e. $\widetilde{\eta}_{ij} = \widetilde{p}_i + \widetilde{p}_j$. If any of the other two possible edge pairs has a smaller value than e_i, e_j, then the edges e_i, e_j are replaced by that pair of edges. Otherwise, the pair e_i, e_j is not replaced. Thus, the idea of this Markov chain is to eliminate edges e which have lengths occurring more often then required, i.e. $f_\Gamma(|e|) < \widehat{f}(|e|)$.

A realization of the random geometric graph model, where the edge lengths have been fitted to real data by the above described MCMC simulation, can be seen in Figure 14. Note that Figure 14 just shows a small cutout of a realization of the network model describing the pore space of a GDL.
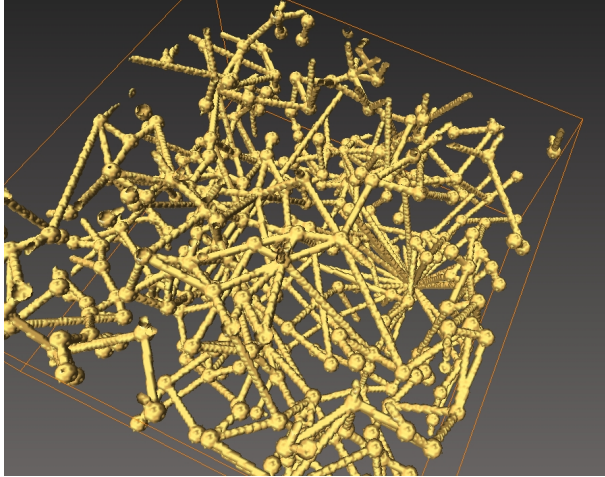
Fig. 14 Small cutout of a simulated 3D random geometric graph

5.4 Discussion of MCMC simulation

Note that the rearrangement of edges according to the MCMC procedure
described in Section 5.3 does not change the distribution of edge degrees.
Moreover, the degrees of the individual vertices are not changed at all. On the
other hand, the discrepance between the (empirical) edge-length distribution
of the finally generated set of edges and the shifted gamma distribution fitted
in Section 5.2 is essentially improved; see Figure 13(c).

However, we are aware of the contingency that it might be possible that
not all predetermined edge-length distributions can be achieved by our MCMC
procedure. This mainly depends on the spatial structure of vertex locations
which has an essential influence on possible edge-length distributions. For ex-
ample, if there is a hard-core distance of 5μm between vertices, no edge with
length smaller than 5μm can be generated. Such a 'lower bound' for the edge-
length distribution, i.e., a distribution of edge-lengths with as much as possible
probability mass next to zero, can be approximated using the first (auxiliary)
Markov chain described in Section 5.3.

For the image data considered in the present paper, such a 'lower bound'
was found by this preprocessing Markov chain. Then, the second (main) Markov
chain shifted the probability mass away from zero and, as a result of this, the
predetermined shifted Γ-distribution had a good chance to be approximated
well.

Another reason why in the present case the predetermined distribution
could be fitted well, originates from the fact that all other components of the
stochastic network model have been fitted quite well to real data. In partic-
ular, the vertex model has nicely been fitted to the vertices extracted from
synchrotron data; see Section 3.3. Furthermore, there is practically no differ-

ence between the distributions of vertex degrees for real and simulated data, respectively.

Note that in the MCMC procedure described in the previous section we considered the Kolmogorov distance of the predetermined and the current (empirical) edge-length distributions as a stopping criterion. Besides, we also fixed a maximum number of possible edge rearrangements. If this number was reached before the Kolmogorov distance fallen below a given threshold, then the simulation was stopped anyway, and the current configuration of the graph is seen as a sample from the random geometric graph to be constructed.

6 Validation of the network model

In the preceding sections of this paper, the validation of the stochastic network model has been accomplished only by means 'local' characteristics of various model components. We now consider two structural characteristics of the whole model in order to validate it. They are relevant to transport properties of pore space and have not been used for model fitting.

6.1 Minimum spanning tree

An important structural characteristic which describes the connectivity of graphs is the so-called minimum spanning tree (MST). It is based on the concept of thinning a graph with the aim to minimize its total length and, at the same time, to keep its connectivity preserved. In our context, the MST is a very useful structural characteristic, because the skeletonization algorithm considered in Section 2.2 preserves principal connectivity properties of the complete 3D pore space. So a global validation of the random geometric graph model with respect to connectivity seems to be reasonable.

Note that the total length of a graph is just the sum of the lengths of its edges. In other words, the principal idea of the MST is to look at a graph where as many edges as possible are removed without changing the connectivity, i.e., all vertices which have been connected before are still connected. Note however that the sequences of vertex degrees of the original graph and its MST are not the same.

In order to practically compute the MST, Prim's algorithm can be used; see e.g. [8] and [19]. The property which is used for model validation is the relative length ρ of the MST in comparison to the length of the original graph, i.e., we consider the ratio

$$\rho = \frac{\text{length of the MST}}{\text{length of the original graph}},$$

which provides important information about the connectivity of the graph. Note that the numerical results which we obtained for the MST of real and simulated data, respectively, are very similar to each other; see Table 1.

Table 1 Tortuosity and MST results

	tortuosity		MST
	mean value	standard deviation	rel. length
Real data	1.35	0.11	0.38
Simulated graph	1.41	0.15	0.40
Rel. failure	4%	27%	5%

6.2 Geometric tortuosity

Other basic characteristics of porous media, which are considered when investigating transport properties, are their porosity and, in order to describe the pathways through the materials, their tortuosity. Note that tortuosity is a physical characteristic which is usually defined as the ratio of the mean effective path length of a fluid through the pore space of a porous material and the material thickness, or, in other words, the mean extension of the real pathway compared to the minimum distance between two points chosen at random; see e.g. [20] and [37]. Note that this notion of tortuosity is given just by a single number, which is not really a morphology-based characteristic.

On the other hand, the notion of geometric tortuosity, which has been introduced in [44], describes the microstructure of pore space in more detail, dealing with the distribution of shortest path lengths through a porous material, i.e., it allows for the consideration of probability distributions and not just mean values.

In order to compute the lengths of shortest paths from top to bottom of GDL material, along the edges of the 3D graph representing the pore space, we first have to determine the starting points of these paths. Therefore a stationary planar Poisson point process with some intensity $\lambda > 0$ is simulated on the top of the GDL. Recall that the sampled point patterns then follow the principle of complete spatial randomness. Choosing this model for the starting points of shortest paths, we had in mind that e.g. gas molecules can start their diffusion/migration at any point of the GDL surface with the same probability and independent of each other (in ex situ experiments).

Note that the starting points of shortest paths generated by the Poisson model mentioned above were not yet included into the random geometric graph model described in the preceding sections. So we had to add these points to the graph, where each point of the Poisson point process simulated on the top of the GDL was connected to the closest vertex of the random graph representing the pore space.

Then, beginning from the starting points, the shortest paths along the edges from top to bottom were determined using Dijkstra's algorithm; see, e.g., [8] and [19]. This analysis has been done both for the graphs extracted from real 3D data and for simulated graphs, where we obtained mean values and standard deviations as given in Table 1; for the distributions of shortest path lengths see Figure 15.
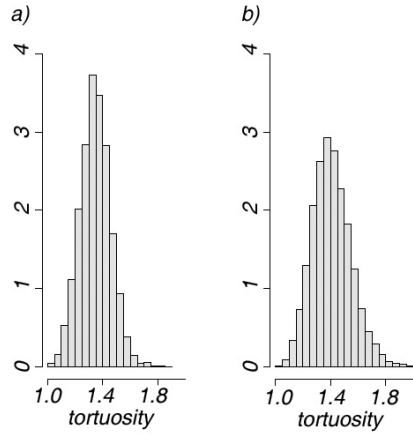
Fig. 15 Tortuosity distributions of modified graphs: a) for graphs extracted from real data; b) for simulated graphs

Although there is no perfect matching of these 'local' tortuosity characteristics for real and simulated data, respectively, the difference between the two mean values is relatively small and also the principal shapes of the two histograms representing the distributions of shortest path lengths are relatively close to each other. This shows that the stochastic network model proposed in the present paper is an appropriate tool in order to investigate the geometry of and transport processes in the pore space of this type of GDL.

7 Summary and conclusions

In the present paper, we developed a model for random geometric graphs in 3D, where we showed that our model can successfully be fitted to graph structures extracted from real image data. The model is built in several steps. We first modelled the vertices of the graph to be constructed using a multilayer approach, i.e., by a stack of (smeared) 2D point processes, where the 3D point process model obtained in this way has been validated using different morphological image characteristics from stochastic geometry. Then, for any given point pattern in 3D, we considered a depending marking of points, using spherical contact distances as marks which are closely related to the physical notion of pore size. In a third step, we proposed an edge model to reconnect a given set of vertices according to a prespecified distribution of vertex degrees. For this purpose we assigned a sequence of vertex degrees as marks to the vertices and checked whether such a configuration of vertex degrees allows the construction of a graph or not. If possible, an algorithm based on the Hakimi-Havel theorem of graph theory was used to construct a preliminary graph. However, since this algorithm does not take into account the locations

of vertices, we finally applied an MCMC simulation to rearrange edges in order to get a better fit to geometric properties of graphs observed in real image data.

These steps altogether led to a 3D random geometric graph which can be simulated relatively easily. A realization of this network model can be seen in Figure 14. Moreover, by partitioning the model construction into different steps, our stochastic network model becomes quite flexible having the potential to be successfully applied also to other porous materials than those considered in the present paper.

To validate the 3D geometric graph model, morphology-based image characteristics were considered which were not used for model fitting. They are related to physical properties of the fibre-based porous material analysed in this paper. One of these characteristics was the relative length of the minimum spanning tree which was used for analysing the connectivity of the graph. The other one was geometric tortuosity, i.e., the ratio of the shortest path length through a material and the material thickness. Note that also this morphology-based image characteristic is very important in order to investigate transport processes in porous materials.

In an ongoing research we are analysing the statistical properties of transport processes along the edges of the stochastic network model developed in the present paper. Note that there are clear computational advantages to investigate the properties of transport processes on geometric graphs, i.e. on one-dimensional structures instead of considering a complete 3D model. In particular, the graph representation of pore space enables the consideration of much larger domains and, simultaneously, keeping computer time at reasonable low levels.

References

1. Abdel-Ghani M.S. and Davies G.A. (1985) Simulation of non-woven fibre mats and the application to coalescers. Chemical Engineering Science, **40**, 117–129.
2. Armatas G.S. (2006) Determination of the effects of the pore size distribution and pore connectivity distribution on the pore tortuosity and diffusive transport in model porous networks. Chemical Engineering Science, **61**, 4662–4675.
3. Berhan L., Yi Y.B. and Sastry A.M. (2004) Effect of nanorope waviness on the effective moduli of nanotube sheets. Journal of Applied Physics, **95**, 5027–5034.
4. Benes V. and Rataj J. (2004) Stochastic Geometry: Selected Topics. Kluwer Academic Publishers, Boston.
5. Blunt M.J., Jackson M.D., Piri M. and Valvatne P.H. (2002) Detailed physics, predictive capabilities and macroscopic consequences for pore–network models of multiphase flow. Advances in Water Resources, **25**, 10691089.
6. Casella G. and Berger R.L. (2002) Statistical Inference. 2nd ed., Duxbury, Pacific Grove.
7. Daley D.J. and Vere-Jones D. (2003) An Introduction to the Theory of Point Processes Volume I: Elementary Theory and Methods. Springer, New York.
8. Diestel R. (2005) Graph Theory. Springer, Heidelberg.
9. Eichhorn S.J. and W.W. Sampson (2005) Statistical geometry of pores and statistics of porous nanofibrous assemblies. Journal of the Royal Society Interface, **2**, 309–318.

10. Eichhorn S.J. and W.W. Sampson (2010) Relationship between specific surface area and pore size in electrospun polymer fibre networks. Journal of the Royal Society Interface, **7**, 641–649.

11. 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.

12. Gelfand A.E., Diggle P.J., Fuentes M. and Guttorp P. (2010) Handbook of Spatial Statistics. Chapman & Hall/CRC, Boca Raton.

13. Gostick J.T., Ioannidis M.A., Fowler M.W. and Pritzker M.D. (2007) Pore network modeling of fibrous gas diffusion layers for polymer electrolyte membrane fuel cells. Journal of Power Sources, **173**, 277–290.

14. Hartnig C., Kuhn R., Krüger P., Manke I., Kardjilov N., Goebbels J., Müller B.R. and Riesemeier H. (2008) Water management in fuel cells – A challenge for non–destructive high resolution methods. MP Materials Testing, **50**, 609–614.

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

16. Inoue G., Matsukuma Y. and Minemoto M. (2007) Numerical analysis of two-phase transport in GDL of polymer electrolyte fuel cells. Proceedings of the 2nd European Fuel Cell Technology and Applications Conference, EFC2007-39024.

17. 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.

18. Jähne B. (2005) Digital Image Processing. 6th revised and extended ed., Springer, Berlin.

19. Jungnickel D. (1999) Graphs, Networks and Algorithms. Springer, Berlin.

20. Keil F. (1999) Diffusion und Chemische Reaktionen in der Gas/Feststoff–Katalyse. Springer, Berlin.

21. Kendall W.S. and Molchanov I. (eds.) (2010) New Perspectives in Stochastic Geometry. Springer, Berlin.

22. Kroese D.P., Taimre T. and Botev Z.I. (2011) Handbook of Monte Carlo Methods. J. Wiley & Sons, Hoboken.

23. Lahiri S.N. (2003) Resampling Methods for Dependent Data. Springer, New–York.

24. León y León C.A. (1998) New perspectives in mercury porosimetry. Advances in Colloid and Interface Science, **76-77**, 341–372.

25. Maheshwaria P.H., Mathur R.B. and Dhamia T.L. (2008) The influence of the pore size and its distribution in a carbon paper electrode on the performance of a PEM fuel cell. Electrochimica Acta, **54**, 655–659.

26. Manke I., Hartnig C., Grünerbel M., Lehnert W., Kardjilov N., Haibel A., Hilger A. and Banhart J. (2007) Investigation of water evolution and transport in fuel cells with high resolution synchrotron X–ray radiography. Applied Physics Letters, **90**, 174105-1 – 174105-3.

27. Mathias M.F., Roth J., Fleming J. and Lehnert W. (2003) Diffusion media materials and characterisation, Handbook of Fuel Cells, Volume III, Chapter 42, Ed. W. Vielstich, A. Lamm, H. Gasteiger. J.Wiley & Sons, Chichester.

28. Møller J. and Waagepetersen R.P. (2004) Statistical Inference and Simulation for Spatial Point Processes. Chapman & Hall / CRC, Boca Raton.

29. Münch B. and Holzer L. (2008) Contradicting geometrical concepts in pore size analysis attained with electron microscopy and mercury intrusion. Journal of the American Ceramic Society, **91**, 4059–4067.

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

31. Piekaar H.W. and L.A. Clarenburg (1967) Aerosol filters – pore size distribution in fibrous filters. Engineering Science, **22**, 1399–1408.

32. Saito T. and Toriwaki J.-I. (1994) New algorithms for euclidean distance transformations of an n-dimensional digitized picture with applications. Pattern Recognition, **27**, 1551–1565.

33. Sampson W.W. (2003) A multiplanar model for the pore radius distribution in isotropic near-planar stochastic fibre networks. Journal of Materials Science, **38**, 1617–1622.

34. Sampson W.W. and Urquhart S.J. (2008) The contribution of out-of-plane pore dimensions to the pore size distribution of paper and stochastic fibrous materials. Journal of Porous Materials, **15**, 411–417.
35. Schlather M. (2001) On the second-order characteristics of marked point processes. Bernoulli, **7**, 99–117.
36. 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.
37. Shen L. and Chen Z. (2007) Critical review of the impact of tortuosity on diffusion. Chemical Engineering Science, **62**, 3748–3755.
38. Sinha P.K., Mukherjee P.P. and Wang C.-Y. (2007) Impact of GDL structure and wettability on water management in polymer electrolyte fuel cells. Journal of Materials Chemistry, **17**, 3089–3103.
39. Sinha P.K. and Wang C.-Y. (2007) Pore–network modeling of liquid water transport in gas diffusion layer of a polymer electrolyte fuel cell. Electrochimica Acta, **52**, 7936–7945.
40. Soille P. (1999) Morphological Image Analysis. Springer, Berlin.
41. Stoyan D., Kendall W.S. and Mecke J. (1995) Stochastic Geometry and its Applications. 2nd ed., J. Wiley & Sons, Chichester.
42. Stoyan D. (1984) On correlations of marked point processes. Mathematische Nachrichten, **116**, 197–207.
43. Thiedmann R., Fleischer F., Hartnig C., Lehnert W. and Schmidt V. (2008) Stochastic 3D modeling of the GDL structure in PEM fuel cells based on thin section detection. Journal of the Electrochemical Society, **155**, B391 – B399.
44. 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.
45. Thulasiraman K. and Swamy M.N.S. (1992) Graphs: Theory and Algorithms. J.Wiley & Sons, Chichester.
46. Tripathi A. and Vijay S. (2009) A note on a theorem of Erdös & Gallai. Discrete Mathematics, **265**, 417–420.
47. Weil H. (1954) The distribution of radial error. Annals of Mathematical Statistics, **25**, 168–170.
48. Yoneda M., Takimoto M. and Koshizuka S. (2007) Effects of microstructure of gas diffusion layer on two-phase flow transport properties. ECS Transactions, **11**, 629–635.