THIELE CENTRE
FOR APPLIED MATHEMATICS IN NATURAL SCIENCE

Algorithms to estimate the rose of directions of a spatial fiber system

Markus Kiderlen and Andreas Pfrang

Research Report No. 01 | January 2005

Algorithms to estimate the rose of directions of a spatial fiber system

This Thiele Research Report is also Research Report number 451 in the Stochastics Series at Department of Mathematical Sciences, University of Aarhus, Denmark.

ALGORITHMS TO ESTIMATE THE ROSE OF DIRECTIONS OF A SPATIAL FIBER SYSTEM

MARKUS KIDERLEN AND ANDREAS PFRANG

Markus Kiderlen, Department of Mathematical Sciences, University of Aarhus, Ny Munkegade, DK–8000 Aarhus C, Denmark, kiderlen@imf.au.dk
Andreas Pfrang, Institute of Applied Physics, University of Karlsruhe, D-76128 Karlsruhe, Germany, andreas.pfrang@physik.uni-karlsruhe.de

Abstract. Anisotropy of a stationary fibre process in three dimensional space can equivalently be quantified by the directional measure (which is up to normalization the rose of directions), the rose of intersections and the associated zonoid. The best accessible quantity is the rose of intersections, as its values can be estimated by counting intersections of the fibres with lower dimensional test sets. Three non-parametric algorithms to estimate the directional measure, based on this information, are investigated. We also present estimators for the associated zonoid, which turns out to be an intuitive tool for visualization. The methods are applied to two different carbon fibre architectures and to simulated data.

1. Introduction

The analysis of spatial systems of fibres is a frequent problem in biology, metallography and other applied sciences. Mathematically, these structures are often modeled as stationary fibre processes. The most basic characteristic of a stationary fibre process is the *length density* L_V, which is the mean total length of all fibres in an observation window of unit volume. Of course, this characteristic does not give any information about preferred directions of the process. To describe anisotropy of the process, there are three quantities in common use. Although, theoretically, all three describe the same first order characteristics of the fibre process, they differ considerably from a practical point of view. The first is the so called *rose of (tangent-) directions* $\mathcal{R}$, which is a measure on the unit sphere in $\mathbb{R}^3$. It is the distribution of the fibre tangent at a randomly chosen "typical" point of the fibre system. The preferred directions of a fibre system are, loosely speaking, just the directions where $\mathcal{R}$ carries most of its mass. $\mathcal{R}$ is often not directly accessible, as an estimation of $\mathcal{R}$ would require the measurement of spatial directions inside the sample. The second quantity is the *rose of intersections* γ, which is a function on the unit sphere. For a unit vector u, $\gamma(u)$ is the mean number of intersection points of the fibre process with a test window of unit area placed in a plane with normal u. (Due to the stationarity this mean number only depends on the orientation,

1991 *Mathematics Subject Classification.* Primary: 62H11; 60D05 secondary: 62M30, 62G05 60G10.
Supported in part by the Carlsberg foundation and by the Deutsche Forschungsgemeinschaft (DFG) within the Sonderforschungsbereich 551.

but not on the location of plane and probe.) The estimation of the rose of intersections (at least at a finite number of normal directions $u_1, \dots, u_k$) is a straightforward counting procedure of intersection points with lower dimensional test sets. The third way to describe anisotropy is a three dimensional set, the so called *associated zonoid* Z (or *Steiner compact*) of the fibre process. Z is convex, symmetric with respect to the origin and has the property that its support function is γ; for details, see below. Although this definition of Z is rather indirect, we will see later that it can be used to visualize the directional information in an intuitive way.

If the fibre process is isotropic, then $\mathcal{R}$ is the rotation invariant probability measure on the sphere, γ is a constant and Z is a ball. It is not necessary to reconstruct $\mathcal{R}$ or Z to detect anisotropy, as the easily accessible function γ also indicates anisotropy. A reconstruction of $\mathcal{R}$ is needed, if preferred directions are to be specified, e.g. for the purpose of quality control or if the fibre process is to be simulated to derive statistical properties of the given fibre system. The simplest model of a stationary fibre process is a Poisson process of line segments of unit length, which is determined by the two parameters $\mathcal{R}$ and L_V. Therefore, a simulation approach requires the estimation of these two quantities. In this note we compare three methods to reconstruct $\mathcal{R}$ and L_V from blurred measurements of the rose of intersections in finitely many directions and illustrate this approach with an application to carbon fibre data. In addition, we visualize the directional properties using the associated zonoid.

Several methods have been introduced in the two-dimensional case, where the rose of directions is a distribution on the unit circle (or, using the natural parametrization, on the interval $[0, 2\pi)$), the rose of intersections is derived from intersection counts with test lines and the associated zonoid is a planar convex set. For surveys of these approaches see e.g. [16] or [12].

For fibre processes in three dimensional space, parametric methods have been suggested. There are also approaches which assume that $\mathcal{R}$ has a smooth density (for example, a finite sum of spherical harmonics). For an overview of those methods we refer to [7, p. 7]. More recently, non-parametric methods have been developed which yield maximum likelihood estimators for $\mathcal{R}$, see [7], [9], [5, Section 9]. We will adopt this latter approach here and compare a least squares estimator with estimators based on the use of the EM algorithm and a linear program approach. After a more formal description of the notions in Section 2, we will describe the estimators of $L_V, \mathcal{R}$ and Z in Section 3. Section 4 is devoted to the application of the theory to a process of spatial carbon fibres. In Section 5 reconstruction from simulated data is used to illustrate the isotropic case. Section 6 concludes with some practical recommendations.

2. Notation and Preliminaries

We will now give a brief and more formal introduction to the notions and results that are needed later. For details the reader is referred to [16]. We assume that the system of fibres to be analyzed is a realization of a fibre process X, i.e. a collection of random fibres (smooth curves of finite length in three dimensional space $\mathbb{R}^3$) with the property that any bounded set only hits finitely many fibres. We also assume that X is homogeneous or *stationary*, meaning that the statistical properties of X do not

change if X is translated. A *direction* is a unit vector, that is an element of the unit sphere S^2 in $\mathbb{R}^3$. We introduce the *directional measure* η of X, which combines the rose of direction and the length density in one measure: Let $W \subset \mathbb{R}^3$ be an observation window (a measurable set) of unit volume. For an origin-symmetric measurable set A of directions let $\eta(A)$ be the mean total length of the union of all fibre points in W with normalized tangent vector in A. This definition is independent of W and gives rise to a measure η on S^2, which is even (i.e. η carries the same mass on a set and on its reflection at the origin). The total mass of η is, by definition, $L_V = \eta(S^2)$. If $0 < L_V < \infty$, we can normalize η and obtain the rose of directions:

$$\mathcal{R}(\cdot) = \frac{\eta(\cdot)}{L_V}.$$

Identifying a unit vector with the line, which is the linear span of this vector, we could consider $\mathcal{R}$ as a distribution on the space of all straight lines through the origin. This is sometimes done in the literature (e.g. in [16, p. 297]), but we prefer here the interpretation as even distribution on the unit sphere S^2. Note that the algorithms, which will be described later, naturally yield an estimator for η, and we therefore describe the estimation of η. Clearly, if $\hat{\eta}$ is such an estimator, then $\hat{\eta}(S^2)$ is an estimator for L_V and $\hat{\eta}/\hat{\eta}(S^2)$ is an estimator for $\mathcal{R}$.

For an arbitrary measure μ on S^2 let $\mathcal{T}(\mu, \cdot)$ be the function given by

$$\mathcal{T}(\mu, u) = \int_{S^2} |u \cdot v| \, d\mu(v), \qquad u \in S^2,$$

where

$$u \cdot v = u_1 v_1 + u_2 v_2 + u_3 v_3$$

is the usual inner product of $u = (u_1, u_2, u_3)$ and $v = (v_1, v_2, v_3)$. As u and v are both unit vectors, $u \cdot v$ is the cosine of the angle between u and v. The integral transform $\mathcal{T}(\mu, \cdot)$ is therefore called *cosine transform* of μ. The rose of intersections γ, which was defined in the introduction, obeys

$$\gamma(u) = \mathcal{T}(\eta, u), \qquad u \in S^2. \tag{1}$$

It is well known that the cosine transform determines any even measure uniquely, so γ and η carry the same directional information. To introduce the associated zonoid, we need to describe a compact convex set analytically: For a compact convex set $K \subset \mathbb{R}^3$ containing the origin and a unit vector u let $h(K, u)$ be the distance of the origin 0 to the supporting plane at K with outer normal u, see Figure 1. $h(K, \cdot)$ is called *support function of* K and determines K uniquely. If K and K' are two compact convex sets,

$$K + K' = \{a + a' \,|\, a \in K,\ a' \in K'\}$$

denotes *Minkowski addition* and $\alpha K = \{\alpha a \,|\, a \in K\}$ scaling of K with $\alpha \geq 0$, we have

$$h(\alpha K + \alpha' K', \cdot) = \alpha h(K, \cdot) + \alpha' h(K', \cdot), \qquad \alpha, \alpha' \geq 0. \tag{2}$$

To give some examples, we note that K is a ball centered at the origin if and only if $h(K, \cdot)$ is constant. If $K = [-v, v]$ is the line segment with endpoints $-v$ and v, then $h(K, u) = |u \cdot v|$. In view of (2) and (1), this shows that γ is the support function

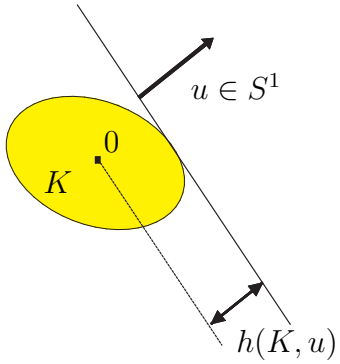

FIGURE 1. The definition of the support function $h(K, \cdot)$ in the case of a planar convex set K.

of a convex body $Z = Z(X)$, which can be considered as a "continuous Minkowski sum" of line segments. Sets with this property can be approximated arbitrarily well by finite Minkowksi sums of line segments and are called *zonoids*. A convex set is uniquely determined by its support function and vice versa, so Z, γ and η carry the same directional information.

The following connection between projections of the associated zonoid Z and projected thick sections of the underlying fibre process X follows directly from the definitions. Let $u \in S^2$ be given and let $x|L$ be the orthogonal projection of $x \in \mathbb{R}^3$ on the plane L with normal u. The support function of the projection $Z|L$ is just the rose of intersections γ, restricted to the unit circle in L. If S is a thickened plane parallel to L of thickness 1, then $(X \cap S)|L$ is a stationary fibre process in L and its (two dimensional) associated zonoid is $Z|L$. This is true if the set of points in $(X \cap S)|L$ that are projections of more than one point of $X \cap S$ has length measure zero. The latter assumption is satisfied for example, if the rose of directions has a density with respect to the natural invariant measure on the sphere, or if the fibres are mutually independent.

The next section describes, how η and Z can be estimated from finitely many measurements of γ.

3. Reconstruction algorithms

We assume that estimators of the rose of intersections are known for finitely many normal directions: Let $u_1, \dots, u_k$ (not necessarily pairwise different) unit vectors and let $\hat{\gamma}_i$ be an estimator for $\gamma(u_i)$ derived from intersection counts. To be more specific, we assume that for each i we have chosen a test window of given area $F > 0$ in a plane with normal u_i and counted the number n_i of fibre intersections with this window. Then $\hat{\gamma}_i = n_i/F$ is an unbiased estimator of $\gamma(u_i)$. The assumption that all test windows have the same area F can be omitted, but we assume it for notational convenience. We will later also need the summarized data obtained by averaging all $\hat{\gamma}_i$ for which the corresponding directions u_i coincide: To describe this formally assume that the tuple

of directions $(u_1, \dots, u_k)$ is ordered

$$(u_1, \dots, u_{m_1}, u_{m_1+1}, \dots, u_{m_2}, \dots, u_{m_{\tilde{k}-1}+1}, \dots, u_{m_{\tilde{k}}})$$

in such a way that $u_1, \dots, u_{m_1}$ all coincide (up to sign), but are different form all others, $u_{m_1+1}, \dots, u_{m_2}$ all coincide (up to sign) but are also different from all others, and so on. Then $\tilde{k}$ is the number of *different* directions and setting $m_0 = 0$, we can put

$$\tilde{\gamma}_j := \frac{\sum_{i=m_{j-1}+1}^{m_j} \hat{\gamma}_i}{m_j - m_{j-1}}, \qquad j = 1, \dots, \tilde{k}. \tag{3}$$

If all directions are pairwise linear independent, then, of course, $\tilde{k} = k$ and $\tilde{\gamma}_i = \hat{\gamma}_i$ for all $i = 1, \dots, k$. As notation for the normal direction corresponding to the measurement $\tilde{\gamma}_j$, we will use $\tilde{u}_j = u_{m_j}$, $j = 1, \dots, \tilde{k}$.

The key idea of reconstruction algorithms is the following: In view of (1), we want to find a measure μ on the unit sphere, whose cosine transform in directions $u_1, \dots, u_k$ best fits to the corresponding measurements $\hat{\gamma}_1, \dots, \hat{\gamma}_k$. The measure μ can then be considered as an estimator for η. The three algorithms we will present here specify the notion of best fit in different ways: The LSQ-method considers a best fit in the least squares sense:

$$\begin{aligned} &\text{minimize} && \sum_{i=1}^{k} \left(\hat{\gamma}_i - \mathcal{T}(\mu, u_i)\right)^2 \\ &\text{subject to} && \mu \text{ is some even measure on } S^2. \end{aligned} \tag{4}$$

The optimization problem (4) is infinite dimensional. For an implementation, it must be discretized. Amazingly, this can be done in a loss-free way. Let $V = \{\pm v_1, \dots, \pm v_m\}$ be the set of nodes associated to $\{u_1, \dots, u_k\}$. Each v_j is of the form

$$\frac{u_i \times u_{i'}}{\|u_i \times u_{i'}\|}, \qquad u_i \neq \pm u_{i'},$$

where $a \times b$ is the cross product of $a, b \in \mathbb{R}^3$ and $\|a\| = \sqrt{a \cdot a}$ is the usual Euclidean norm of a. Let $\mathcal{M}(V)$ be the set of all finite even measures with support in V. If $\mu \in \mathcal{M}(V)$, then for any set $A \subset S^2$, we have

$$\mu(A) = \sum_{j=1}^{m} \frac{\alpha_j}{2} \left(\delta_{v_j}(A) + \delta_{-v_j}(A)\right), \tag{5}$$

where δ_v is the Dirac measure on $v \in S^2$ which means that $\delta_v(A)$ is 1 if $v \in A$ and 0 otherwise. The variables $\alpha_j \geq 0$ are the total masses of μ at the points $\pm v_j$, $j = 1, \dots, m$. (As μ is even, half of this mass must be attributed to v_j and the other half to $-v_j$.) Thus, $\mu \in \mathcal{M}(V)$ is determined by the numbers $\alpha_1, \dots, \alpha_m$. It follows e.g. from [7, Corollary 1] that among all solutions of (4) there is at least one in $\mathcal{M}(V)$, so the optimal objective function value is not changed if we include the extra assumption $\mu \in \mathcal{M}(V)$ in (4). Using the special form (5) of $\mu \in \mathcal{M}(V)$, the optimization problem

reads

$$\text{(LSQ)} \qquad \begin{aligned} &\text{minimize} && \sum_{i=1}^{k}\Big(\hat\gamma_i - \sum_{j=1}^{m}\alpha_j|v_j\cdot u_i|\Big)^2 \\ &\text{subject to} && \alpha_1,\dots,\alpha_m \geq 0. \end{aligned}$$

This is a quadratic program, which can be solved using standard software. Let $\hat\alpha_1,\dots,\hat\alpha_m$ be a solution of (LSQ), and $\hat\eta_{LSQ}\in\mathcal{M}(V)$ be the measure attributing mass $\hat\alpha_j/2$ to v_j and $-v_j$, $j=1,\dots,m$. Then, $\hat\eta_{LSQ}$ is called *least squares estimator* for η and coincides with the estimator $\hat\mu_k$ of [5, Algorithm NoisyRoseLSQ].

Best fit to the given measurements must not necessarily be measured in the least squares deviation. Instead, best fit can be understood in the sense of information theory. One possibility is to minimize the I_1-divergence of the vector of observations $\hat p=(\hat\gamma_1,\dots,\hat\gamma_k)$ and the vector of expected values $p=(\mathcal{T}(\mu,u_1),\dots,\mathcal{T}(\mu,u_k))$, where μ is an estimator for η. The I_1-divergence is closely related to the Kullback-Leibler divergence, see Appendix A. This approach, together with the loss-free discretization explained above, leads to the optimization problem

$$\text{(EM)} \qquad \begin{aligned} &\text{minimize} && \sum_{i=1}^{k}\Big(\sum_{j=1}^{m}\alpha_j|v_j\cdot u_i| - \hat\gamma_i \ln\big(\sum_{j=1}^{m}\alpha_j|v_j\cdot u_i|\big)\Big) \\ &\text{subject to} && \alpha_1,\dots,\alpha_m\geq 0. \end{aligned}$$

Let $\hat\alpha_1,\dots,\hat\alpha_m$ be a solution of (EM), and $\hat\eta_{EM}\in\mathcal{M}(V)$ be the measure attributing mass $\hat\alpha_j/2$ to v_j and $-v_j$, $j=1,\dots,m$. Then, $\hat\eta_{EM}$ is called *EM estimator* for η, as (EM) can be solved numerically using the iterative EM algorithm. This algorithm is based on the idea of a two phase method (**e**stimation and **m**aximization) in each iteration step, see [7] for more details on this algorithm and properties of the estimator. (EM) can also be solved using Markov chain Monte Carlo simulations, an approach that was suggested in [11].

In [7] another estimator is introduced, which is based on the linear program

$$\text{(LP)} \qquad \begin{aligned} &\text{minimize} && \sum_{i=1}^{\tilde k}\Big(\tilde\gamma_i - \sum_{j=1}^{m}\alpha_j|v_j\cdot\tilde u_i|\Big) \\ &\text{subject to} && \sum_{j=1}^{m}\alpha_j|v_j\cdot\tilde u_i|\leq\tilde\gamma_i,\quad i=1,\dots,\tilde k, \\ & && \alpha_1,\dots,\alpha_m\geq 0. \end{aligned}$$

In the same way as above, a measure $\hat\eta_{LP}\in\mathcal{M}(V)$ can be constructed from a solution of (LP). $\hat\eta_{LP}$ is called *LP estimator*. This estimator is geometrically motivated: The associated zonoid of the fibre process lies (approximately) in the polytope

$$\hat Q = \bigcap_{i=1}^{\tilde k}\{x\in\mathbb{R}^3 : |\tilde u_i\cdot x|\leq\tilde\gamma_i\} \tag{6}$$

and $\mathcal{T}(\hat{\eta}_{LP}, \cdot)$ is the support function of a zonoid which lies inside P and has the largest possible values in the measurement directions. This interpretation also explains, why we work here with the summarized data ($\tilde{\gamma}_i$ instead of $\hat{\gamma}_i$): If we formulated (LP) with the original data, then, for any direction $\tilde{u}_j$, only the minimal value of $\hat{\gamma}_{m_{j-1}+1}, \ldots, \hat{\gamma}_{m_j}$ would be taken into account and the other values would be ignored. If the linear program (LP) is solved with the simplex algorithm, it yields a solution with at most k positive components. Simulations indicate that, in general, (LP) tends to find a solution with even fewer support points.

The motivation for the LSQ estimator comes from the following observation: If the measurements $\hat{\gamma}_1, \ldots, \hat{\gamma}_k$ are stochastically independent and Gaussian with joint variance σ^2, then the LSQ estimator is a maximum likelihood estimator. Note that the maximum likelihood problem is non-parametric, as the "parameter" is a measure on the sphere. It can be shown that the estimator is strongly consistent. This means that the estimator converges almost surely to η, if the number of normal directions tends to infinity (and the directions are suitably chosen on the sphere). This was shown in [9]. In [5] this result was strengthened, giving even the speed of convergence under slightly stronger assumptions.

The Gaussian assumption is justified if the numbers of intersections of fibres with the test planes are high. Otherwise a model working with integer valued variables would be more appropriate, as the estimators $\hat{\gamma}_i$ are derived from counts. If, for instance, the underlying fibre process is a Poisson process of line segments, then the counts are Poisson distributed variables. If the intersection counts are stochastically independent Poisson variables, then the EM estimator is a maximum likelihood estimator. Strong consistency for this estimator was shown in [7].

If there is a measure μ that fits to all measurements (i.e. the measurement vector $(\hat{\gamma}_1, \ldots, \hat{\gamma}_k)$ and $(\mathcal{T}(\mu, u_1), \ldots, \mathcal{T}(\mu, u_k))$ coincide), then the sets of solutions of (LSQ), (EM) and (LP) coincide, too. This means that in the case where the input data are very close to the exact values, all three solutions are equally good.

In the following, let $\hat{\eta}$ be one of the estimators $\hat{\eta}_{LSQ}$, $\hat{\eta}_{EM}$ or $\hat{\eta}_{LP}$. Then, $\hat{\eta}$ is of the form (5) with $(\alpha_1, \ldots, \alpha_m) = (\hat{\alpha}_1, \ldots, \hat{\alpha}_m)$. Thus, $\hat{\eta}$ has the property that its cosine transforms in directions $u_1, \ldots, u_k$ approximate the measurements $\hat{\gamma}_1, \ldots, \hat{\gamma}_k$ best possible in the specified sense. This can also be expressed in the following way: Let $s_1, \ldots, s_\ell$ be all the non-degenerate line segments in $[-\hat{\alpha}_1 v_1, \hat{\alpha}_1 v_1], \ldots, [-\hat{\alpha}_m v_m, \hat{\alpha}_m v_m]$ and let Y be a stationary Poisson process of line segments with intensity 1 where only the line segments $s_1, \ldots, s_\ell$ occur and they occur with equal probability $1/\ell$. Then, among all possible roses of intersections, the rose of intersections of Y best fits to the observed rose of intersections of X in the directions $u_1, \ldots, u_k$.

It is not difficult to derive an estimator $\hat{Z}$ of the associated zonoid from $\hat{\eta}$: With the notation introduced in the last paragraph, we set

$$\hat{Z} = s_1 + \ldots + s_\ell. \tag{7}$$

The consistency results of the estimators for η carry over to $\hat{Z}$: under the same assumptions that assure consistency of $\hat{\eta}$, the set $\hat{Z}$ converges to Z almost surely (in the Hausdorff metric). Alternative to this approach one might think to reconstruct the

associated zonoid more directly by using the known approximate values of the rose of directions and the fact that they are support function values of Z. One could for example use $\hat{Q}$ in (6) as estimator for Z. But this reconstruction is not very stable. Typically, one may obtain a singleton with positive probability. A more robust estimation $\tilde{Z}$ of Z could be obtained by minimizing the deviation of $h(\tilde{Z},\cdot)$ from the given data $\hat{\gamma}(u_1),\dots,\hat{\gamma}(u_k)$ in a least squares sense. This, however, requires to solve a complicated optimization problem given by [5, Algorithm NoisySupportLSQ]. Both, $\hat{Q}$ and $\tilde{Z}$, approximate the zonoid Z by a convex set, which is symmetric at the origin but not a zonoid.They are therefore unsatisfactory. The fact that there are origin-symmetric compact convex sets which are not zonoids, discriminates the higher dimensional setting from the two dimensional case treated in [13], see [7] for details. In the following, the associated zonoid will be estimated using (7). Due to (2), the support function of $\hat{Z}$ is given by

$$h(\hat{Z},v) = h\Big(\sum_{i=1}^{\ell} s_i, v\Big) = \sum_{i=1}^{k} \alpha_i |u_i \cdot v| = \mathcal{T}(\hat{\eta},v), \qquad v \in S^2.$$

We will apply all three estimation procedures to two different carbon fibre architectures, as described in the next section.

4. Application to carbon fibre architectures

4.1. **Experimental setting.** A carbon fibre preform was infiltrated with pyrolytic carbon consisting of sp^2 hybridized carbon. The preform (Sintec, Halblech, Germany) consisted of high tenacity carbon fibres. The preform was a stack of alternating layers: layers of type A consist of nearly parallel fibres, whereas layers of type B are made up of nearly isotropically oriented carbon fibres. The preform was pressed into the reaction volume of the reactor mainly applying pressure in the stacking direction.

The infiltration was carried out in a hot wall reactor at a temperature of $1095°C$ from pure methane at a methane pressure of 20 kPa. An infiltration time of 120 h was chosen. The gas flow was adjusted to reach a maximum residence time of 0.1 s at

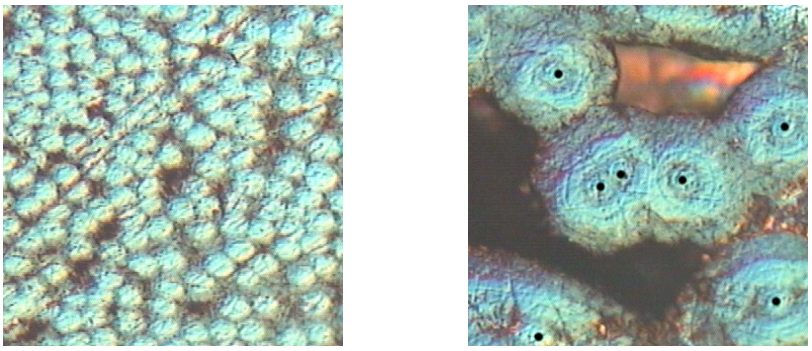

Figure 2. Polarized light micrographs of intersection plane 1 showing fibre architectures of type A (left) and type B (right). The image size is 150μm $\times 150\mu$m. The black dots in the right image mark fibers.

the end of the fibre preform. The residence time is defined as the time the gas needs from its entrance of the hot reaction zone to the point where the deposition of pyrolytic carbon takes place. After infiltration a cuboid was cut from the felt: the height of the cuboid in stacking direction (or z-direction) was 5 mm, the area perpendicular to the stacking direction was 18mm×5mm (in x-direction and y-direction, respectively). The residence time at the investigated part of the sample was 0.025 s. Differently oriented intersection planes (compare Table 1) were prepared by polishing using the cuboid as starting point. The intersection planes were polished using silicon carbide paper with an average grain size of 10μm (FEPA 2400) in the final stage. This grain size allows a clear differentiation of the carbon fibers from the surrounding pyrolytic carbon matrix in polarized light microscopy.

Light microscopy was carried out on these polished planes using a DM LM Microscope (Leica, Wetzlar, Germany) equipped with polarizer and analyzer. Imaging conditions were chosen to allow for a maximum contrast between carbon fibres and surrounding pyrolytic carbon matrix (for details of the polarized light microscopy of pyrolytic carbon see [1] and [4]). Images of a surface area of 310μm×240μm were acquired. An example of the light microscopy data is shown in Fig. 2 for intersection plane 1. A section of type A is shown on the left, one of type B on the right. The fibers were marked in the image files by dots manually (see right image in Fig. 2) and then counted manually. The summarized results of this counting procedure are given in Tables 1 and 3 ($j = 1$). The number of investigated images per direction was chosen larger for lower fibre counts area and smaller for higher fibre counts per area.

4.2. **The estimators.** In the following we present the results of the estimation for the three different estimators. The EM algorithm ran 2000 iterations. As mentioned in Section 3 we were using the summarized data (3) for the LP estimator. To calculate the representation of the associated zonoids from the estimators of η, we used the *Mathematica* notebook from *www.georgehart.com/zonohedra/zonohedrification.html*, see also [6].

We first consider the fibre process of type A. The measurements which served as input are summarized in the first four columns of Table 1. The second column lists the measurement directions (i.e. the normals of the intersection planes that where used). For legibility, here and in the following, we do not normalize the normal directions. The third column shows the number of measurements in the given direction and the third column shows the arithmetic mean of these measurements. For readability, we present only the summarized data although the LSQ estimator and the EM estimator both depend on all 47 individual measurements. ¿From the three algorithms we get three estimators $\hat{\eta}$ for η as output. For all three of these, 93% of the total mass is concentrated on the direction $(1, 0, 0)$ (and its antipodal). The remaining mass is then partitioned to 3 to 7 other point pairs in V. The LP estimator has the smallest support with 4 pairs of points. An illustration of the directional measure on the sphere is given in Figure 3, left. As the three estimators coincide in all relevant aspects, we only show the LSQ estimator. The spherical representations of the other two estimators appear to be identical.

No.	direction	# meas.		Cosine transform of		
j	$\tilde{u}_j$	$m_j - m_{j-1}$	$\tilde{\gamma}_j$	$\hat{\eta}_{LSQ}$	$\hat{\eta}_{EM}$	$\hat{\eta}_{LP}$
1	(-1, 1, 0)	2	5270	4241	4232	3740
2	(1, 1, 0)	2	5365	4418	4385	3824
3	(0, 1, 1)	8	173	173	173	173
4	(0, 1,-1)	8	158	189	160	158
5	(1,-1,-1)	4	3272	3587	3536	3188
6	(-1,-1, 1)	3	3580	3637	3637	3178
7	(-1,-1,-1)	3	3424	3578	3526	3066
8	(1,-1, 1)	3	2919	3340	3374	2919
9	(0, 0, 1)	0	—	233	228	233
10	(0, 1, 0)	9	127	217	131	127
11	(1, 0, 0)	3	5710	6031	6071	5281

TABLE 1. Comparison of the empirical rose of intersections with the cosine transforms of the estimators for the fibre system of type A in mm/mm^2.

To evaluate the quality of the estimators, the (summarized) empirical rose of intersections is compared with the cosine transforms of the estimators in Table 1. Recall that $\mathcal{T}(\hat{\eta}, \cdot)$ is the rose of intersections of a fibre process for which $\hat{\eta}$ is the directional measure. The deviations from the empirical measurements are most expressed in the case of the LP estimator: its cosine transform underestimates several measurements considerably. (By construction, an over-estimation is not possible.) This indicates that the geometric construction underlying the LP estimator is not well suited in the case of large measurement errors. The estimated associated zonoid, based on $\hat{\eta}_{LSQ}$ is shown in Figure 3, right. It has the shape of a very thin cylindrical object with main axes parallel to the x-axes. Like the spherical representation of $\hat{\eta}$, it reflects very clearly the high preference of the direction $\pm(1, 0, 0)$. This observation corresponds very well to the specification of the material, which states that the fibres in the type A architecture are nearly parallel.

$\hat{L}_V(LSQ)$	$\hat{L}_V(EM)$	$\hat{L}_V(LP)$	$\hat{L}_V(iso)$	$\hat{L}_V(voronoi)$
6231	6195	5448	6000	5810

TABLE 2. Estimators for the length density for type A in mm/mm^3 in comparison with two usual estimators.

The estimators for the length density L_V, which are just the total masses of the estimators for η, are given in Table 2. The last two columns of this table are usual

estimators for the length density, obtained using the formula

$$\hat{L}_V = 2 \sum_{i=1}^{\tilde{k}} \lambda_i \tilde{\gamma}_i,$$

where $\lambda_1, \ldots, \lambda_{\tilde{k}}$ are nonnegative weights summing up to 1. For $\hat{L}_V(iso)$, all weights are equal to $1/\tilde{k}$, which gives an unbiased estimator for L_V under the assumption that the fibre process is isotropic. This assumption is not satisfied here, but the estimator is reasonable, as the normal directions are quite evenly spread on S^2. For $\hat{L}_V(voronoi)$, λ_i is the relative area of the spherical voronoi cell of $\tilde{u}_i$ generated by the set $\{\tilde{u}_1, \ldots, \tilde{u}_{\tilde{k}}\}$ on S^2. This estimator does not require isotropy, see e.g. [10, p.116]. The estimator $\hat{L}_V(LP)$ is considerably smaller than all other estimators of L_V, which is again due to the above mentioned fact that $\hat{\eta}_{LP}$ is very sensitive to measurements that are considerably smaller than their mean.

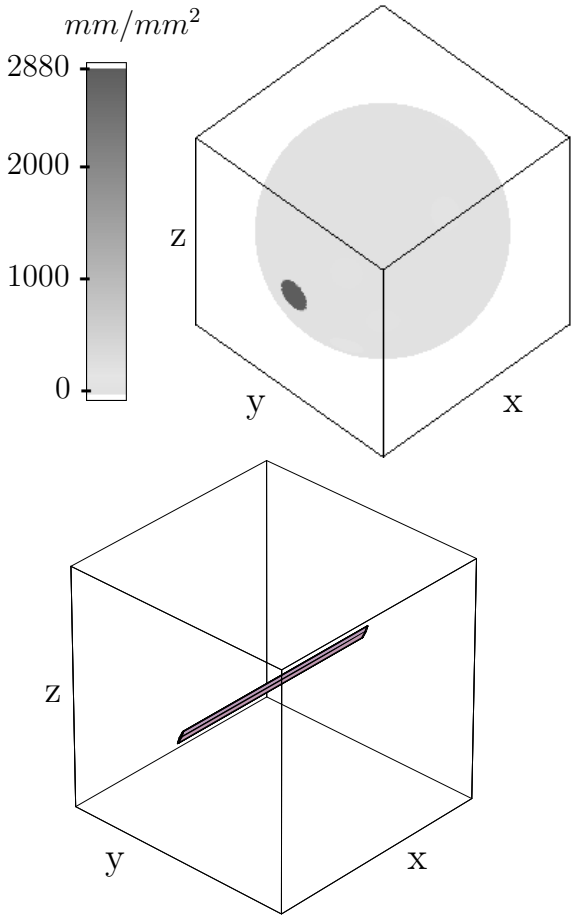

FIGURE 3. Type A fibre architectures: Spherical representation of the LSQ-estimator for η (left) and the estimator for the associated zonoid derived from it (right). The bounding box for the latter has a side length of 6070 mm.

No.	direction	# meas.		Cosine transform of		
j	$\tilde{u}_j$	$m_j - m_{j-1}$	$\tilde{\gamma}_j$	$\hat{\eta}_{LSQ}$	$\hat{\eta}_{EM}$	$\hat{\eta}_{LP}$
1	(-1, 1, 0)	12	531.9	526.5	533.6	525.8
2	(1, 1, 0)	8	538.1	584.0	603.3	538.1
3	(0, 1, 1)	8	587.0	508.1	492.3	376.0
4	(0, 1,-1)	8	497.6	495.5	491.6	409.1
5	(1,-1,-1)	8	394.7	403.3	398.1	394.7
6	(-1,-1, 1)	8	458.8	450.2	455.0	458.8
7	(-1,-1,-1)	8	602.2	554.6	535.8	474.0
8	(1,-1, 1)	8	485.8	487.1	478.9	464.0
9	(0, 0, 1)	9	60.0	134.6	74.9	60.0
10	(0, 1, 0)	7	518.5	584.0	621.2	518.6
11	(1, 0, 0)	7	543.6	543.6	543.6	543.6

TABLE 3. Comparison of the empirical rose of intersections with the cosine transforms of the estimators for the fibre system of type B in mm/mm^2.

		α_j for		
j	v_j	$\hat{\eta}_{LSQ}$	$\hat{\eta}_{EM}$	$\hat{\eta}_{LP}$
1	±(0, 0, 1)	26.5	4.9	0
2	±(1, 1, 0)	270.8	289.5	225.3
3	±(1,-1, 1)	0	0	57.4
4	±(1,-1, 0)	200.8	220.5	179.9
5	±(-1,1, 1)	21.8	0	0
6	±(1, 0, 0)	110.9	112.6	220.6
7	±(1, 1, 1)	0	1.1	0
8	±(1, 0, 1)	127.8	98.4	18.6
9	±(0, 1, 0)	241.6	260.3	208.7

TABLE 4. Type B architecture: The support points and corresponding masses of the estimated directional measures $\hat{\eta}_{LSQ}$, $\hat{\eta}_{EM}$ and $\hat{\eta}_{LP}$ corresponding to eqn. (5) in mm/mm^2.

We turn to the fibre process of type B. The input data and the cosine transforms of the estimators are summarized in Table 3. As the estimators of η now show a more different behavior, we also give a list of the estimated masses and their support directions in Table 4. All other directions in V did not carry mass for any of the estimators. In Figure 4 the spherical representation of $\hat{\eta}_{LSQ}$ and and the derived estimator $\hat{Z}$ are shown. Like $\hat{\eta}_{LSQ}$, this convex set shows that the anisotropy is by far less expressed than in the case

of the fibre architecture of type A. In Figure 5, projections of Z on axis-parallel planes are shown. As mentioned at the end of Section 1, these sets can be interpreted as associated zonoids of planar fibre processes obtained as projected thick sections of X. The projection on the x-y-plane (Figure 5, left), has the approximate shape of a disk, which indicates that this fibre architecture is almost isotropic with respect to rotations around the z-axis. According to the specification, the fibres are nearly isotropic, which implies that the associated zonoid should be close to a ball. This does not fit to our estimator $\hat{Z}$. A possible explanation is that isotropy is destroyed when the preform is pressed (in z-direction) into the reactor. This would also explain, why isotropy holds (approximately) in the x-y plane.

For comparison we show the associated zonoids derived from $\hat{\eta}_{EM}$ and $\hat{\eta}_{LP}$ in Figure 6. All three estimators for Z have approximately the same shape, but the one derived from the LP solution is slightly smaller than the others.

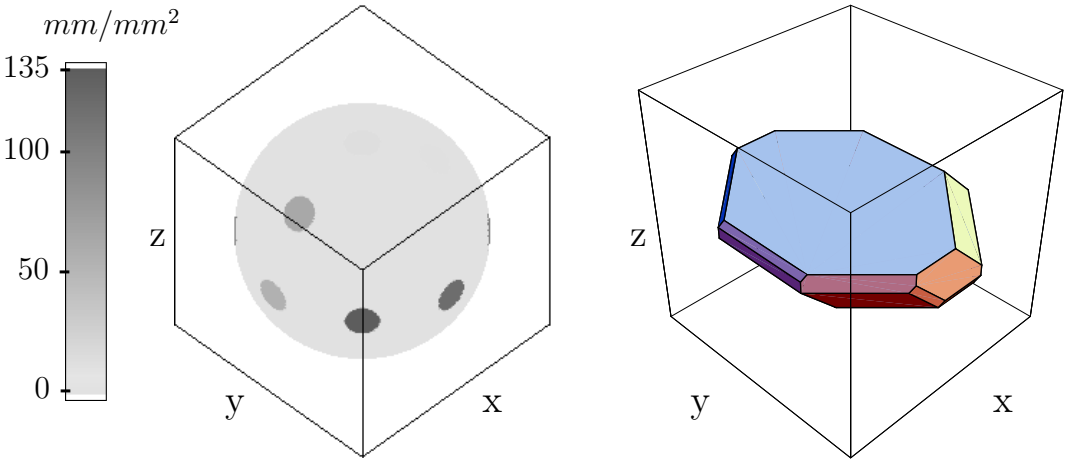

FIGURE 4. Type B fibre architectures: Spherical representation of the LSQ-estimator for η (left) and the estimator for the associated zonoid derived from it (right). The bounding box for the latter has a side length of 620 mm.

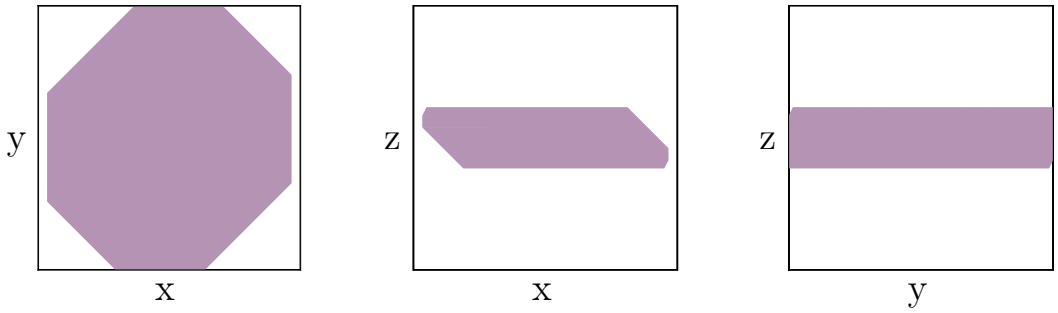

FIGURE 5. Projections of the associated zonoid in Figure 4, right, on axes-parallel planes.

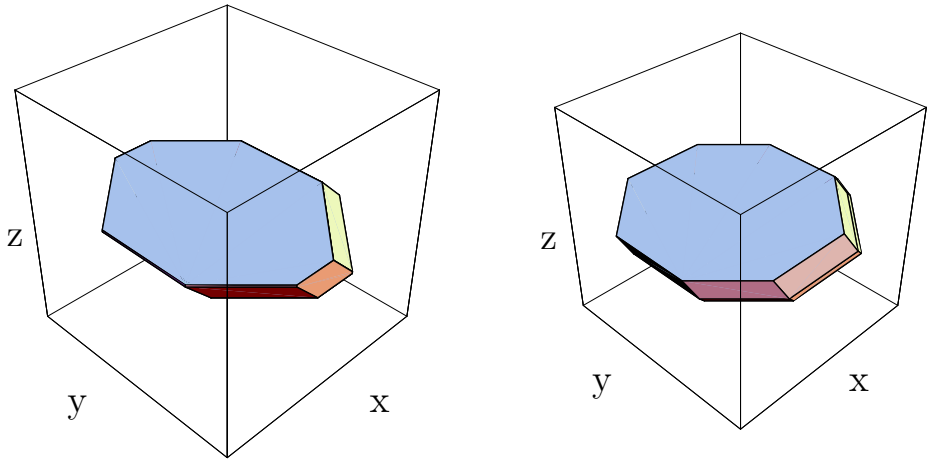

FIGURE 6. Type B fibre architectures: Estimators for the associated zonoid derived from the EM estimator (left) and the LP estimator (right). The bounding boxes have side lengths 621 mm and 544 mm, respectively.

$\hat{L}_V(LSQ)$	$\hat{L}_V(EM)$	$\hat{L}_V(LP)$	$\hat{L}_V(iso)$	$\hat{L}_V(voronoi)$
1000	987	910	1044	1028

TABLE 5. Estimators for the length density for type B in mm/mm^3.

The estimators of the length density are listed in Table 5. The following considerations show that the estimators for the length density are in correspondence with the nominal mean fibre length per unit volume: As the the overall fibre volume fraction of the preform is 20 % and the fibres have a diameter of approximately $9\mu m$, a very rough estimate for the length density of the preform, containing both, type A and type B layers, is 3144 mm/mm^3. Taking into account that the layers of type B cover about double the volume than type A layers, the LSQ solutions yield the estimation 2744 mm/mm^3, and the EM estimations yield $2723 mm/mm^3$. Both numbers are acceptable approximations. The corresponding value for LP is 2423 mm/mm^3, which is still reasonable, but probably an underestimation.

5. Application to simulated isotropic data

Finally, we used unblurred exact data to estimate the directional measure and the associated zonoid of an isotropic fibre process of length density 1000 mm/mm^3. As input for the algorithms, we used values of the exact rose of intersections for the same directions, which were used for the carbon fibre architecture of type A, see Table 1. To make the procedures comparable, we did not use direction number 9, as there were no measurements available for the carbon fibre process. In all other directions, the exact value 500 mm^{-2} was used as input for the algorithms. With this input, the sets of solutions of (LSQ), (EM) and (LP) are the same, but the different algorithms

find diverse estimators. The cosine transforms of all three estimators fit exactly to the corresponding given data. The estimators for the length density (total masses of $\hat{\eta}$) are listed in Table 6.

$\hat{L}_V(LSQ)$	$\hat{L}_V(EM)$	$\hat{L}_V(LP)$	$\hat{L}_V(iso)$	$\hat{L}_V(voronoi)$
1051	1065	1099	1000	1000

TABLE 6. Estimators for the length density for the simulated data in mm/mm^3.

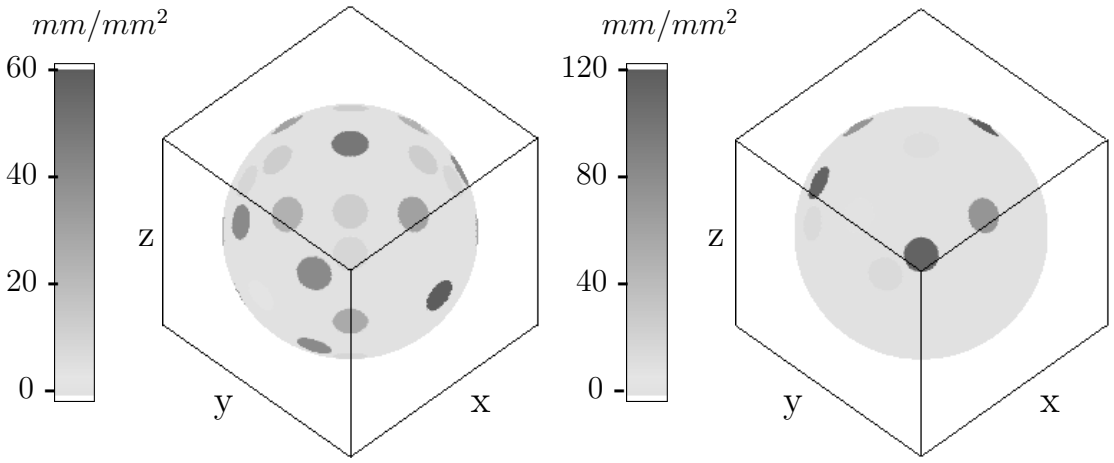

FIGURE 7. Exact isotropic data: Spherical representation of the LSQ estimator $\hat{\eta}_{LSQ}$ (left) and the LP estimator $\hat{\eta}_{LP}$ (right).

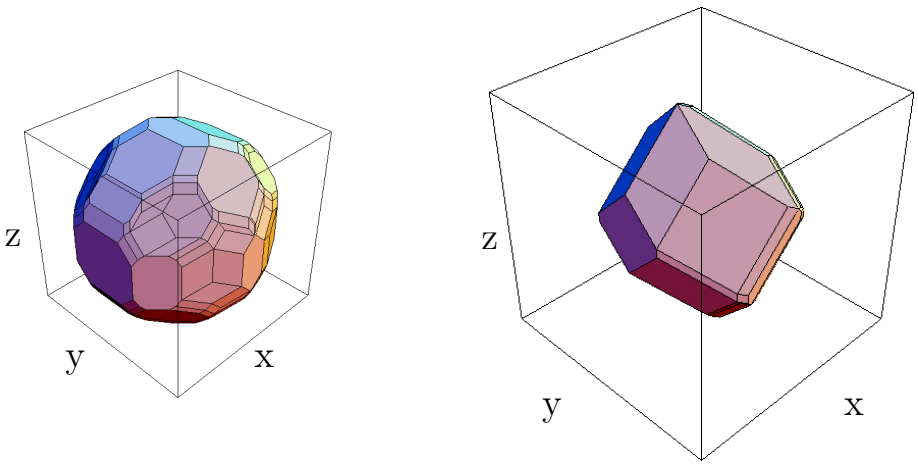

FIGURE 8. Exact isotropic data: The estimated associated zonoids derived from $\hat{\eta}_{LSQ}$ (left) and $\hat{\eta}_{LP}$ (right). The side lengths of the bounding boxes are $1022\,mm$ and $1414\,mm^2$, respectively.

In Figure 7, the LSQ and the LP estimator are shown; the corresponding estimators of Z derived from $\hat{\eta}_{LSQ}$ and $\hat{\eta}_{LP}$ are illustrated in Figure 8. As the EM estimator does not show any features qualitatively different from the LSQ estimator, $\hat{\eta}_{EM}$ is not shown here. This example indicates that the estimators for η do not clearly indicate isotropy, as they are concentrated on finitely many points with masses that are not equal. In particular, $\hat{\eta}_{LP}$ is misleading, as it indicates anisotropy which is not present. The reason for this behavior lies in the fact that the number of normal directions is very small and that the simplex algorithm, which was used to solve (LP), prefers solutions with few support points. Although $\mathcal{T}(\hat{\eta}_{LP}, \cdot)$ coincides with the given value 500 in all measurement directions, it takes the value 707 in direction $(0, 0, 1)$. This can also be seen from the estimated associated zonoid derived from $\hat{\eta}_{LP}$ (Figure 8, right): This zonotope looks too extended in z-direction. In contrast to this, the estimator in Figure 8, left, is a good approximation of the ball of radius 500, thus reflecting isotropy rather clearly. This example indicates that the information contained in just ten values of the rose of intersections is too small to give a hint on isotropy. Of course, the situation is even more involved if the measurements are blurred.

6. Conclusion

We have shown that all three estimators for the directional measure (and their derived estimators for the associated zonoid) are useful to indicate anisotropy of a stationary fibre process. This is even true, if the number of measurements is very small, provided that the anisotropy is expressed and the measurement errors are small. On the other hand, it is very difficult to distinguish isotropic and slightly anisotropic processes if the number of normal directions is small. For simulation examples with a higher number of normal directions, see [7]. Considering the comparison between the different estimators, the LP estimator has clearly the weakest performance, as it tends to produce estimators of η which have few support points and therefore might indicate preferred directions that are not present. In addition, $\hat{\eta}_{LP}$ becomes worse, if the number of measurement directions increases, which was already noted in [7]. One way out is the use of smoothing neighboring measurements before using them as input for (LP), as suggested in [13] for the planar case, but these methods of course also destroy information present in the data and introduce new parameters whose optimal choice is nontrivial. We therefore recommend to use either the LSQ or the EM estimator in applications. If the input measurements are known to be (approximately) Poisson distributed, then $\hat{\eta}_{EM}$ is to be preferred, as it is tailor-made for this situation. In all other cases, the LSQ estimator is a good choice: From a theoretical point of view, this is true as the mean number of intersection points in a lower dimensional test window is approximately normal distributed if the area of the window is large. From a practical point of view, the LSQ estimator is easier to implement than the EM estimator, as it uses only standard algorithms.

The three estimators are particularly useful, whenever there is no a priory information on preferred directions available. In contrast to the established parametric estimators (which usually work with spherical distributions that are rotationally symmetric), the

presented estimators can also handle multi modal directional measures. As their implementation is relatively simple, they can be used in practical applications, whenever enough values for the rose of intersections are available.

Acknowledgments: The authors would like to thank S. Lichtenberg, G. Schoch and K.J. Hüttinger for the infiltration of the preform and helpful remarks and I. Walter for cutting the sample.

Appendix A. The I_1–divergence

When introducing the EM estimator, we made implicit use of the I_1–divergence. The purpose of this appendix is a short introduction to this notion. Properties, generalizations and applications can be found in [8].

Let $\hat{p} = (\hat{p}_1, \dots, \hat{p}_k)$ and $p = (p_1, \dots, p_k)$ be two vectors in k-dimensional space, where the components of $\hat{p}$ are non-negative and those of p are positive. The *I_1–divergence of $\hat{p}$ and p* is defined by

$$I_1(\hat{p}, p) = \sum_{i=1}^{k} \left(\hat{p}_i \, \ln \frac{\hat{p}_i}{p_i} - \hat{p}_i + p_i \right).$$

This is a special case of [8, Definition 2.4 and (1.1)], where the underlying measure ρ is the counting measure on $\{1, \dots, k\}$. If p and $\hat{p}$ are probability vectors, then

$$I_1(\hat{p}, p) = \sum_{i=1}^{k} \hat{p}_i \ln \frac{\hat{p}_i}{p_i}$$

is the *Kullback-Leibler divergence of p and $\hat{p}$*, which is also called log-likelihood ratio statistic. The Kullback Leibler divergence is not a distance function, as it is neither symmetric nor does it satisfy the triangle inequality, but $I_1(\hat{p}, p) = 0$ holds if and only if $p = \hat{p}$. It plays an important role in information theory. It is also used for goodness-of-fit statistics based on counting the occurrence of certain events and on comparing the resulting relative frequencies with the according expectations. This indicates a connection to our procedure of estimating the directional measure from intersection counts.

For a measure μ given by (5), we have

$$\mathcal{T}(\mu, u) = \sum_{j=1}^{m} \alpha_j |v_j \cdot u|, \qquad u \in S^2.$$

If $p = (\mathcal{T}(\mu, u_1), \dots, \mathcal{T}(\mu, u_k))$ and $\hat{p} = (\hat{\gamma}_1, \dots, \hat{\gamma}_k)$, then

$$I_1(\hat{p}, p) = \sum_{i=1}^{k} \left(\sum_{j=1}^{m} \alpha_j |v_j \cdot u_i| - \hat{\gamma}_i \ln \left(\textstyle\sum_{j=1}^{m} \alpha_j |v_j \cdot u_i| \right) \right) + \sum_{i=1}^{k} \hat{\gamma}_i \, (\ln \hat{\gamma}_i - 1)$$

coincides with the objective function in (EM) up to addition of a term, which does not depend on the unknowns $\alpha_1, \dots, \alpha_m \geq 0$.

References

[1] Bortchagovsky, E.G., Reznik, B., Gerthsen, D., Pfrang, A. and Schimmel, Th. (2003) Optical properties of pyrolytic carbon deposits deduced from measurements of the extinction angle by polarized light microscopy. *Carbon* **41**, 2430–2433.

[2] Campi, S., Colesanti, A. and Gronchi, P. (1995) Convex bodies with extremal volumes having prescribed brightness in finitely many directions. *Geom. Dedicata* **57**, 121–133.

[3] Campi, S., Haas, D. and Weil, W. (1994) Approximation of zonoids by zonotopes in fixed directions. *Disc. Comput. Geom.* **11**, 419–431.

[4] Pfrang, A. and Schimmel, Th. (2004) Quantitative analysis of pyrolytic carbon films by polarized light microscopy. *Surface and Interface Analysis* **36**, 184–188.

[5] Gardner, R. J., Kiderlen, M. and Milanfar, P. (2004+) Convergence of Algorithms for Reconstructing Convex Bodies and Directional Measures. *submitted*.

[6] Hart, G. W. (1999) Zonohedrification. *Mathematica J. 7*, **3**, *www.mathematica-journal.com/issue/v7i3*.

[7] Kiderlen, M. (2001) Non-parametric estimation of the directional distribution of stationary line and processes. *Adv. Appl. Prob.* **33**, 6–24.

[8] Liese, F. and Vadja, I. (1987) *Convex Statistical distances*. Teubner, Leipzig.

[9] Männle, S. (2002) *Ein Kleinste-Quadrat-Schätzer des Richtungsmaßes für stationäre Geradenprozesse*. Masters Thesis, University of Karlsruhe, Germany.

[10] Ohser, J. and Mücklich, F. (2000) *Statistical analysis of microstructures in materials science*. Wiley Series in Statistics in Practice, Chichester.

[11] Prokešová, M. (2003) Bayesian MCMC estimation of the rose of directions. *Kybernetika* **39**, 703–717.

[12] Rataj, J. and Béneš, V. (2004) *Stochastic Geometry: Selected Topics*. Springer, New York.

[13] Rataj, J. and Saxl, V. (1989) Analysis of planar anisotropy by means of the Steiner compact. *J. Appl. Prob.* **26**, 490–502.

[14] Schneider, R. (1993) *Convex Bodies: The Brunn-Minkowski Theory*. Cambridge University Press, Cambridge.

[15] Schneider, R. and Weil, W. (2000) *Stochastische Geometrie*. Teubner, Stuttgart.

[16] Stoyan, D., Kendall W. S. and Mecke, J. (1995) *Stochastic Geometry and its Applications*. 2nd edn., John Wiley, New York.