

# CHARACTERIZATION OF BIVARIATE SIZE-ORIENTATION DISTRIBUTION OF CIRCULAR PLATE PARTICLES

KAREL BODLÁK<sup>1</sup>, ARUN BALASUNDARAM<sup>2</sup>, ARUN M GOKHALE<sup>2</sup> AND VIKTOR BENEŠ<sup>1</sup>

<sup>1</sup>Department of Probability and Mathematical Statistics, Charles University, Sokolovská 83, 18675 Praha 8, Czech Republic, <sup>2</sup>School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0245, USA

e-mail: benesv@karlin.mff.cuni.cz

(Accepted 26 August, 2002)

## ABSTRACT

The paper is devoted to the stereological unfolding problem of bivariate size-orientation distribution of platelike particles in metallography. Gokhale (1996) derived an integral equation which relates this bivariate distribution in three-dimensional (3D) space to the corresponding size-orientation distribution of planar sections of the specimen. The present paper yields a numerical algorithm which enables to transform a bivariate histogram of observed quantities to the histogram of 3D characteristics. The use of the method is demonstrated in examples with simulated data, where an easy analytical solution is available and can be compared with the results of estimation. The spectrum of unfolding problems solved numerically (cf. Ohser and Muecklich, 2000) is thus extended.

Keywords: circular plate particles, EM algorithm, unfolding VUR sampling.

## INTRODUCTION

The stereological estimation of the joint distribution of particle parameters from planar sections enables to quantify the 3D geometrical properties of the materials microstructure which is related to mechanical properties. The bivariate distribution of the size and shape factor of spheroidal particles was studied by Cruz-Orive (1976) and later for particles of some other shapes in Ohser and Muecklich (2000). The trivariate size-shape-orientation unfolding problem for spheroidal particles was solved in Beneš and Krejčíř, (1997). Two main sampling schemes were used, either the IUR (isotropic uniform random) or VUR (vertical uniform random) sections. Particles or cracks which can be modelled by circular plates have zero shape factor. In this case the investigation of size-orientation distribution is sufficient, which was investigated first by Gokhale (1996). In that paper the theoretical solution is presented for the VUR sampling design, i.e., the integral equation connecting the probability densities is derived as well as its inversion. This inversion theoretically enables to evaluate the bivariate distribution of the spatial characteristics (the size and orientation in 3D) from the planar bivariate distribution. However, the formula is based on two-fold integration followed by a two-fold differentiation and an analytical form of planar bivariate joint probability density is required. Given the bivariate histogram of

section data measured by means of an image analyser this would need first a parametrical bivariate density model to be fit. Altogether it is not recommendable to proceed using inversion formula directly, therefore a numerical histogram-based method has to be developed which is known, e.g., for size-shape problems (Ohser and Muecklich, 2000).

To do this the first step is to transform the integral equation in terms of probability densities to an equivalent integral equation in terms of distribution functions. This corresponds to the discrete data of the transformation of relative frequencies to cumulative frequencies. The integral equation for distribution function is then discretized which leads to a system of linear equations. Since a direct solution of this system may lead to negative frequencies, special numerical approaches have been adapted. Among them the EM-algorithm (Ohser and Muecklich, 2000) is the most popular one recently, since it always leads to a non-negative solution.

First attempt to the solution of the formulated problem was published by Beneš *et al.* (1996). The integral Eq. 1 was split in two subsequent univariate unfolding problems which were discretized in its original form. Thus the use of elliptic integrals was unavoidable. In the present paper an improvement consists in a direct discretization of bivariate equation which decreases estimation errors of EM-algorithm.

Further an ingenious transformation (Eq. 4) of data leads to a simpler evaluation of kernel coefficients in a closed form.

In the next Section the derivation of the integral equation between distribution functions is presented which arises to be difficult for the size-orientation problem. A special solution which needs a transformation of functions was derived, so that finally one does not deal with a distribution function exactly, but with an alternative form which is convenient for the numerical solution. A simple transformation of data then enables a standard numerical solution using EM-algorithm (Ohser and Muecklich, 2000) in Section 3.

In Section 4 a synthetic example is present which demonstrates the use of the developed algorithm. Model bivariate densities in Gokhale (1996) are tested such that the inverse solution is known theoretically. Then a sample is simulated from the planar bivariate distribution and transformed to the desired bivariate histogram estimating the spatial size-orientation distribution. True and estimated marginal distributions of spatial characteristics are plotted to demonstrate the results of the method.

### ESTIMATION OF BIVARIATE SIZE-ORIENTATION DISTRIBUTION BY EM ALGORITHM

Fix an arbitrary direction in the space which is called a vertical axis. Consider a coordinate system in the 3D space, such that the vertical axis coincides with z-axis, and a system of circular plates dispersed in a reference volume. Each plate is described by its radius  $R$  and a normal orientation  $(\theta, \phi)$  in spherical coordinates, see Fig. 1.

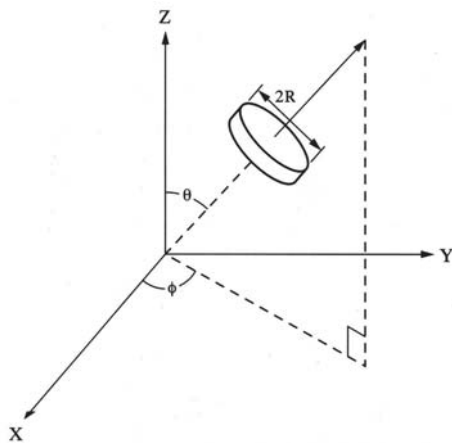


Fig. 1. *Particle parameters of a model shape in 3D.*

A vertical plane is any plane parallel to vertical axis. For a given vertical plane intersecting a plate the

intersection is a segment of length  $2r$  and orientation  $\alpha$  with respect to vertical direction, Fig. 2.

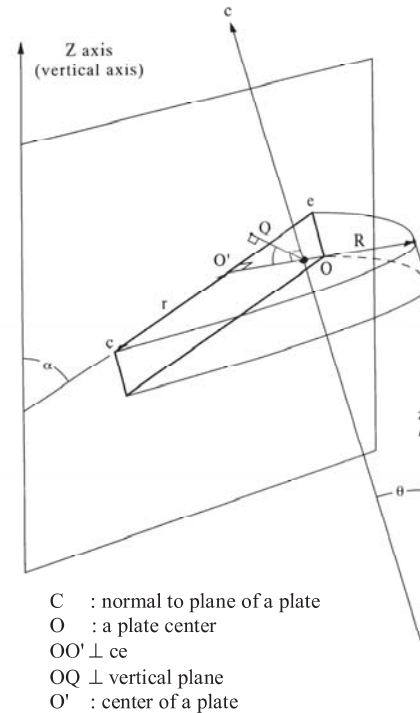


Fig. 2. *Geometry in the intersection of a platelike particle and a vertical plane.*

Let  $p(R, \theta, \phi)$  be the joint probability density of plate size and orientation, denote

$$g(R, \theta) = \frac{1}{2\pi} \int_0^{2\pi} p(R, \theta, \phi) d\phi. \quad \text{That means we will}$$

consider the distribution of size and colatitude only which arises in many applications, e.g. if a material specimen is under tension or compression in the direction of the vertical axis.

Let  $f(r, \alpha)$  be the bivariate probability density of half chord length  $r$  and chord orientation  $\alpha$  of plates if the planar section is VUR. The relationship between  $f(r, \alpha)$  and  $g(R, \theta)$  is given by a double integral equation in Gokhale (1996)

$$f(r, \alpha) = \frac{4rN_V}{\pi N_A \sin^2 \alpha} \int_r^{r_m} \int_{\pi/2-\alpha}^{\pi/2} \frac{\cos^2 \theta \sin \theta g(R, \theta) d\theta dR}{(R^2 - r^2)^{1/2} (\sin^2 \alpha - \cos^2 \theta)^{1/2}}, \quad (1)$$

where  $N_V$  is the mean number of particles per unit volume,  $N_A$  is the mean number of particle sections per unit area of a vertical plane,  $r_m$  is the largest particle's radius.

From practical reasons it is better to use  $\gamma = \pi/2 - \alpha$  in  $f(r, \alpha)$ . Then we get after this change of variables

$$f(r, \gamma) = \frac{4rN_V}{\pi N_A \cos^2 \gamma} \int_{\gamma}^{r_m \pi/2} \int_{\gamma}^{\pi/2} \frac{\cos^2 \theta \sin \theta g(R, \theta) d\theta dR}{(R^2 - r^2)^{1/2} (\sin^2 \gamma - \cos^2 \theta)^{1/2}} \quad (2)$$

In order to present a numerical solution we transform the Eq. 2 from the relationship between the density functions to the relationship between the distribution functions. The procedure is described as follows. After the differentiation the right hand side of Eq. 2 transforms to

$$f(r, \gamma) = \frac{4N_V}{\pi N_A \cos^3 \gamma \sin \gamma} \frac{d}{dr} \frac{d}{d\gamma} \underbrace{\left( \int_r^{r_m} \int_{\gamma}^{\pi/2} \underbrace{(R^2 - r^2)^{1/2}}_{s(R,r)} \underbrace{(\cos^2 \gamma - \cos^2 \theta)^{1/2} \cos^2 \theta \sin \theta g(R, \theta) d\theta dR}_{t(\theta, \gamma)} \right)}_{H(r, \gamma)} \quad (3)$$

In the next step we will integrate the Eq. 3 according to the variables  $r, \gamma$ . We can define

$$F(r, \gamma) = \int_r^{r_m} \int_{\gamma}^{\pi/2} f(s, \beta) \cos^3 \beta \sin \beta d\beta ds \quad (4)$$

Integrating the right hand side of Eq. 3 as in Eq. 4 we get the desired equation

$$\begin{aligned} N_A F(r, \gamma) &= \frac{4N_V}{\pi} \int_r^{r_m} \int_{\gamma}^{\pi/2} \frac{d}{ds} \frac{d}{d\beta} \underbrace{\{H(s, \beta)\}}_{I(s, \beta)} d\beta ds \\ &= \frac{4N_V}{\pi} \int_{\gamma}^{\pi/2} \left[ \int_r^{r_m} \frac{dI}{ds} ds \right] d\beta = \frac{4N_V}{\pi} \int_{\gamma}^{\pi/2} [0 - I(r, \beta)] d\beta \quad (5) \\ &= \frac{-4N_V}{\pi} \int_{\gamma}^{\pi/2} \frac{d}{d\beta} \left[ \int_r^{r_m} \int_{\beta}^{\pi/2} s(R, r) t(\theta, \beta) g(R, \theta) d\theta dR \right] d\beta \\ &= \frac{4N_V}{\pi} \int_r^{r_m} \int_{\gamma}^{\pi/2} s(R, r) t(\theta, \beta) g(R, \theta) d\theta dR \end{aligned}$$

In fact  $F$  is not a distribution function corresponding to  $f$ , but we can weight the data correspondingly to the weighting in Eq. 4 and solve the Eq. 5 with respect to  $g$  numerically.

### NUMERICAL SOLUTION

The unfolding problem belongs to a class of inverse problems which are often called ill-posed. This means that a small error in the input data may cause a large error in the estimated output. Because of this property it is suitable to use a fine discretization

relatively to the amount of data because otherwise the method can lead to large disturbances in an output.

Assume that plate sections were observed using the VUR sampling design and their parameters were measured. Suppose now that a number  $M$  of size classes and a number  $N$  of orientation classes is used for discretization. Thus we can define class intervals with limits as follows:

$$\begin{aligned} r_i &= R_i = i\delta, \quad \delta = r_m/M, \quad i = 1, \dots, M; \\ \gamma_j &= \theta_j = j\omega, \quad \omega = \pi/(2N), \quad j = 1, \dots, N. \end{aligned}$$

Typically  $r_m$  is unknown and it is estimated by means of a half of the largest observed plate trace length. Under this choice of the discretization the platelike particle's trace with size and orientation  $r, \gamma$ , respectively, belongs to an  $(i, j)$ -th class if  $r_{i-1} \leq r < r_i$  and  $\gamma_{j-1} \leq \gamma < \gamma_j$ .

This way we can classify the input data into a bivariate histogram with class limits for the size and the orientation. Next we suppose the function  $g(R, \theta)$  being constant within each class  $(i, j)$ ,  $N_V g(R, \theta) = x_{ij}$  for  $R_{i-1} \leq R < R_i$  and  $\theta_{j-1} \leq \theta < \theta_j$ . As a discrete analogue of the stereological Eq. 5 we can obtain a system of linear equations

$$y_{kl} = \sum_{i=k}^M \sum_{j=l}^N p_{ijkl} x_{ij}, \quad i = 1, \dots, M; \quad j = 1, \dots, N. \quad (6)$$

The numerical densities  $y_{kl}$  present the (weighted, Eq. 4) mean number of the platelike particles in the class  $(k, l)$  and they are defined as follows:

$$\begin{aligned} y_{kl} &= N_A [F(r_{k-1}, \gamma_{l-1}) - F(r_k, \gamma_{l-1}) - F(r_{k-1}, \gamma_l) + F(r_k, \gamma_l)] \\ \text{and} \quad x_{ij} &= N_V g(R_i, \theta_j). \end{aligned}$$

The coefficients of the linear equation system (6) are given by

$$\begin{aligned} p_{ijkl} &= p(i, j, k-1, l-1) - p(i, j, k, l-1) \\ &\quad - p(i, j, k-1, l) + p(i, j, k, l) \end{aligned}, \quad (7)$$

where

$$\begin{aligned} p(i, j, k, l) &= \frac{4}{\pi} \int_{(i-1)\delta}^{i\delta} \int_{(j-1)\omega}^{j\omega} s(R, r_k) t(\theta, \gamma_l) d\theta dR, \quad i > k, j > l \quad (8) \end{aligned}$$

and  $p_{ijkl} = 0$  otherwise.

To calculate coefficients  $p(i, j, k, l)$  we have to integrate the right side of Eq. 8 separately according to the parameters  $R$  and  $\theta$ . For the case  $r > 0$  we obtain

$$\int_a^b s(R,r) \, dR = \frac{1}{2} \left[ b(b^2 - r^2)^{1/2} - a(a^2 - r^2)^{1/2} - r^2 \log \left( \frac{b + (b^2 - r^2)^{1/2}}{a + (a^2 - r^2)^{1/2}} \right) \right]$$

and for setting  $r = 0$  we get

$$\int_a^b s(R,r) \, dR = \frac{1}{2} [b^2 - a^2].$$

Integrating the right side of Eq. 8 in  $\theta$  we can write

$$\int_c^d t(\theta, \gamma) \, d\theta = \left[ \frac{\cos \theta}{4} (\cos^2 \gamma - \cos^2 \theta)^{3/2} - \frac{\cos^2 \gamma}{8} (\cos^2 \gamma - \cos^2 \theta)^{1/2} - \frac{\cos^4 \gamma}{8} \arctan \left( \frac{\cos \theta}{(\cos^2 \gamma - \cos^2 \theta)^{1/2}} \right) \right]_c^d.$$

In order to solve the system of linear Eq. 6 the iterative EM algorithm in Ohser and Muecklich (2000) is used. Suppose that  $y$  is said to be the matrix of incomplete data obtained by planar sampling and  $x$  represents the matrix of parameters to be estimated. Each iteration consists of two steps, an E-step and an M-step, where E stands for expectation and M for maximization. Let  $z_{ijkl}$  be the number of events occurring in the class  $(i,j)$  which contributes to the counts in the class  $(k,l)$ . The E-step yields the expected value of  $z_{ijkl}$  given input data  $y$  under the current estimate  $x^{(\lambda)}$  of the parameter matrix  $x$  as follows

$$z_{ijkl} = x_{ij}^{(\lambda)} y_{kl} p_{ijkl} / r_{kl}^{(\lambda)}. \quad (9)$$

The M-step yields a maximum likelihood estimate of the parameter  $x$  using the estimated data  $z_{ijkl}$  from the

$$\text{E-step. Define } t_{ij} = \sum_{k=1}^M \sum_{l=1}^N p_{ijkl} \text{ and } r_{kl}^{(\lambda)} = \sum_{i=1}^M \sum_{j=1}^N p_{ijkl} x_{ij}^{(\lambda)}.$$

Then the M-step is given by

$$x_{ij}^{(\lambda)} = \frac{1}{t_{ij}} \sum_{k=1}^M \sum_{l=1}^N z_{ijkl}. \text{ Combining these two steps we get}$$

an EM iteration given by the updating formula

$$x_{ij}^{(\lambda+1)} = \frac{x_{ij}^{(\lambda)}}{t_{ij}} \sum_{k=1}^M \sum_{l=1}^N \frac{y_{kl} p_{ijkl}}{r_{kl}^{(\lambda)}}; \quad (10)$$

$$i = 1, \dots, M, j = 1, \dots, N$$

The Eq. 10 produces a sequence  $x_{ij}^{(\lambda)}$  which

converges to a solution, i.e., the matrix of parameters  $x$ . As an initial iteration we set  $x_{ij}^{(0)} = y_{ij}$ . This choice of nonnegative initial values ensures that each of  $x$  is nonnegative. There is in fact not a unique solution. Convergence of the EM algorithm is guaranteed in theory but the solution can depend on the choice of the initial values used in the iteration scheme. See Ohser and Muecklich (2000) for more experience with the use of the EM algorithm in metallographical practice when solving unfolding problems.

## SIMULATION STUDY

The following examples are given to illustrate the use of the EM algorithm for estimation of the true bivariate size and orientation distribution from simulated planar data according to given joint density functions. The number of iterations of the EM algorithm depends on how fine is the discretization as well as on the accuracy of measured planar data. For a sample size about 10000 and a total of 10 classes the number of iterations need not exceed 32 as recommended in Ohser and Muecklich (2000).

We have chosen 5 classes of the radii and 5 classes of the orientation. The number of the simulated platelike particle traces was 2000 and the number of iterations of the EM algorithm was suggested equal to 16. Suppose first that the platelike particle's size and orientation distribution is  $f_1(r, \alpha) = 12 \cdot 10^{18} / \pi \cdot r (r_m^2 - r^2)^2$ , the radius of the largest platelike particle is given by  $r_m = 0.001$  cm. The size of a class is equal 0.0002 cm. The derived true platelike particle's size and orientation distribution (Gokhale, 1996) has a form  $g_1(R, \theta) = 10^{16} / \pi \cdot R (R_m^2 - R^2)^{3/2}$  where  $R_m = r_m$ , i.e. orientations are uniform random in this example.

If we set the number of platelike particles per  $\text{cm}^2$  of the vertical plane area at value  $N_A = 10^4$ , then the total number of platelike particles  $N_V$  is equal to  $64 / \pi \cdot 10^6$ . Using the EM algorithm to evaluate the data simulated from the density  $f$  we get the estimated bivariate size and orientation distribution  $\text{est}(g)$  in 3D, Fig. 3. The estimated total number of microcracks per a unit volume is  $\text{est}(N_V) = 2.04 \cdot 10^7$  which is very close to the true value  $N_V$ . The true and estimated marginal distribution of the radius and the orientation is plotted in Fig. 4. In order to compare the true and estimated marginal distributions we have to transfer a scale of the true distribution to classes from 1 to 5 on the horizontal axis and use a relative scale for the vertical axis.



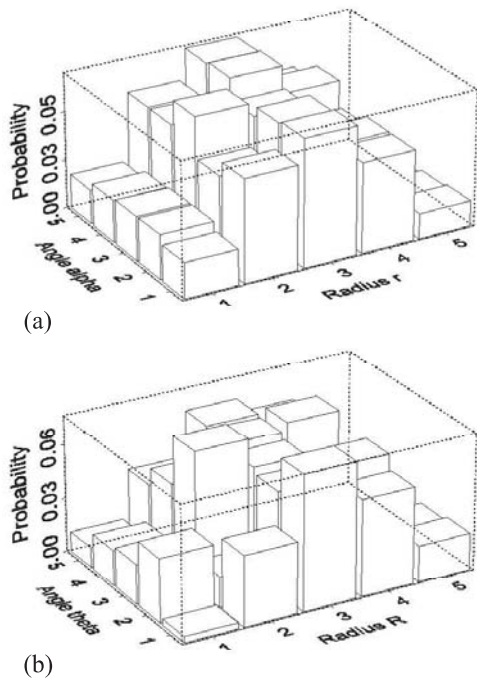


Fig. 3. *Platelike particle's distribution with the uniform distribution of the angle. (a) The bivariate histogram given by simulated planar data from  $f_1$ . (b) Estimated bivariate distribution in 3D.*

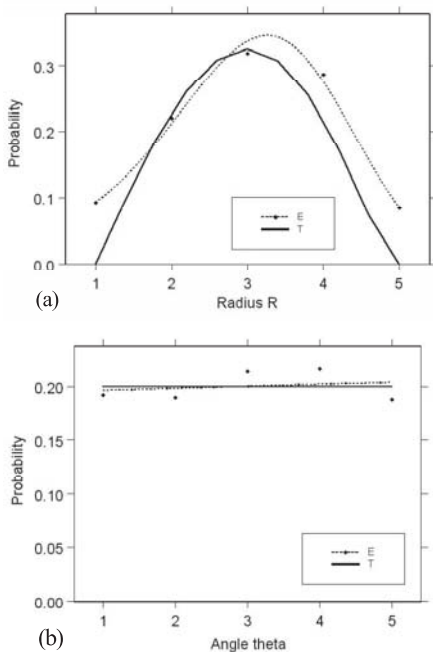


Fig. 4. *The solid lines (T) show the true marginal distributions of the spatial size and orientation distribution corresponding to Fig. 3. We indicate smoothed estimated marginal distribution corresponding to the size and orientation by dotted line (E). (a) The curve E is fitted by the spline of the third order. (b) For smoothing the curve E we use the least squares method.*

In the next example suppose that platelike particle's trace  $r$  and orientation  $\alpha$  are still uncorrelated but the orientations are not uniform random. The planar probability density has a form  $f_2(r,\alpha) = 24 \cdot 10^{18} / \pi \cdot r (r_m^2 - r^2)^2 \sin^2 \alpha$  and true platelike particle's size and orientation distribution is  $g_2(R,\theta) = 2 \cdot 10^{16} / \pi \cdot R (R_m^2 - R^2)^{3/2} \cos^2 \theta$ . For the same setting of  $N_A$  the total number of platelike particles  $N_V$  is equal to  $256 / (3\pi) \cdot 10^6$ . From the simulation we receive the estimated value  $est(N_V) = 2.8 \cdot 10^7$ . The estimated bivariate distribution in 3D is depicted in Fig. 5, analogously to Fig. 4 the true and estimated marginal distribution of 3D characteristics is plotted in Fig. 6.

The last simulation example covers the situation of correlated platelike particle's trace  $r$  and its orientation angle  $\alpha$ . The planar probability density is given by  $f_3(r,\alpha) = 4 \cdot 10^{12} / \pi \cdot r (r_m^2 - r^2) + 12 \cdot 10^{18} / \pi \cdot r (r_m^2 - r^2)^2 \sin^2 \alpha$ . The associated distribution in 3D has a form  $g_3(R,\theta) = 3 \cdot 10^9 / \pi \cdot R (R_m^2 - R^2)^{1/2} + 10^{16} / \pi \cdot R (R_m^2 - R^2)^{3/2} \cos^2 \theta$  and the associated total number of platelike particles is  $N_V = 16 \cdot 10^7 / (6\pi) + 128 \cdot 10^6 / (3\pi)$ . The estimated total number of platelike particles is  $est(N_V) = 2.1 \cdot 10^7$  which is again very close to the true value  $N_V$ . The simulated planar and estimated spatial bivariate distributions are showed in Fig. 7, the true and estimated spatial marginal distribution is in Fig. 8.

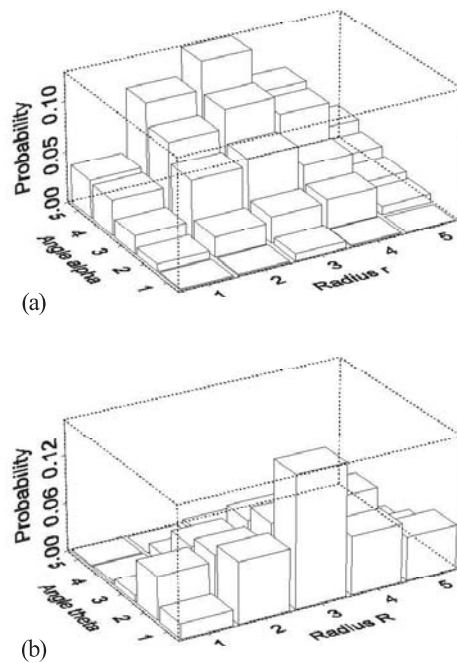


Fig. 5. *Anisotropic uncorrelated platelike particle's distribution. (a) The bivariate histogram given by simulated planar data from  $f_2$ . (b) Estimated bivariate distribution in 3D.*

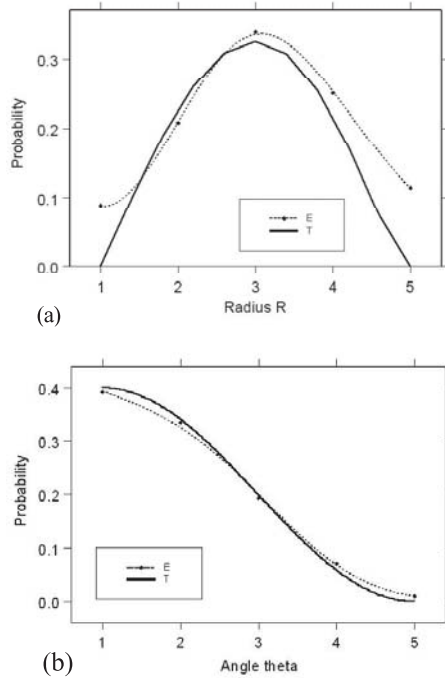


Fig. 6. The solid lines ( $T$ ) show the true marginal distributions of the spatial size and orientation distribution corresponding to Fig. 5. We indicate the smoothed estimated marginal distribution corresponding to the size and orientation by dotted line ( $E$ ). The both curves  $E$  in (a) and (b) are fitted by the spline of the third order.

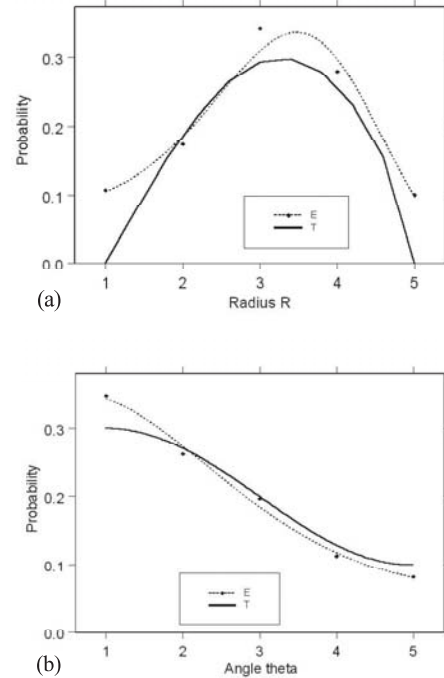


Fig. 8. The solid lines ( $T$ ) show the true marginal distributions of the spatial size and orientation distribution corresponding to Fig. 7. We indicate smoothed estimated marginal distribution corresponding to the size and orientation by dotted line ( $E$ ). (a) The curve  $E$  is fitted by the method of least squares. (b) For smoothing the curve  $E$  we use by the spline of the third order.

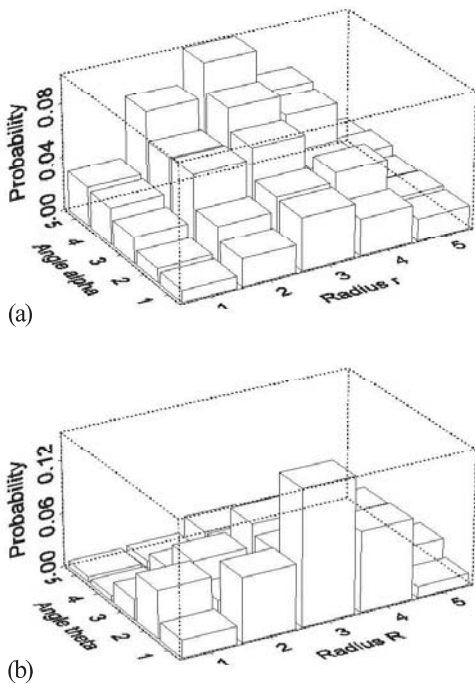


Fig. 7. Anisotropic and correlated platelike particle's distribution. (a) The bivariate histogram given by simulated data from  $f_3$ . (b) Estimated bivariate distribution in 3D.

## DISCUSSION

A method for the estimation of the spatial bivariate size-orientation distribution of circular plates using stereological methods was developed. The unfolding of true plate parameters is based on the observation in a vertical uniform random section plane. In metallographical practice the circular plates present an approximate model for cracks or platelike particles in particle reinforcement. The vertical axis then coincides with the direction of tension or compression of the material.

If a close analytical form of the true spatial bivariate platelike particle's distribution is available then (sometimes after tedious mathematical calculations) one could express the associated planar bivariate distribution in the vertical planes and vice versa. But for a real dataset it is difficult to fit its planar bivariate distribution by an analytical formula. In such a case (especially when the data are presented in a bivariate histogram form) it is easier to proceed in discretization and apply an efficient numerical procedure like EM algorithm to calculate the bivariate

distribution of 3D characteristics from the measured data. Such a method was developed in the present paper and its use was demonstrated in synthetic examples where the data were simulated from a known distribution. The results are promising for the application of the procedure in metallographical practice, they extend the supply of unfolding methods discussed in Ohser and Muecklich (2000).

### ACKNOWLEDGEMENTS

Karel Bodlák acknowledges U.S. NSF sponsored programme (DMR-9816618) and Czech Ministry of Education project KONTAKT ME335 for international collaboration between Charles University and Georgia Institute of Technology for financial support. He also thanks Georgia Institute of Technology for allowing his stay in Atlanta. A.B. and A.M.G acknowledge U.S. National Science of Foundations Research Grant No. DMR-9816618 (for which Dr

MacDonald was the Program Monitor) for financial support. V.B. was supported by project MSM11300008 of the Czech Ministry of Education.

### REFERENCES

- Beneš V, Krejčíř P (1997). Decomposition in stereological unfolding problems. *Kybernetika* 33(3):245-58.
- Beneš V, Gokhale AM, Slámová M (1996). Unfolding the bivariate size-orientation distribution. *Acta Stereol* 15(1):9-14.
- Cruz-Orive LM (1976). Particle size-shape Distributions: The General Spheroid Problem. *J Microsc* 107(3):235-53.
- Gokhale AM (1996). Estimation of bivariate size and orientation distribution of platelike particles. *Acta Metall and Mater* 44(2):475-85.
- Ohser J, Muecklich F (2000). *Statistical Analysis of Microstructures in Material Science*. New York: J Wiley & Sons.