Data depth for simple orthogonal regression with application to crack orientation

Christine H. Müller *
University Kassel

April 9, 2009

Abstract

This paper studies tangential and simplicial data depth for orthogonal regression through the origin. For both depth notions, it is proved that orthogonal lines have the same depth. As robustness measure of maximum depth estimators, exact-fit points for one line and for two orthogonal lines are defined and calculated. Since the simplicial depth has a simple asymptotic distribution, tests can be easily derived. These tests are used for checking whether data are distributed around two orthogonal lines. But since distributions which are invariant with respect to rotations with angle of $\pi/2$ have constant depth functions, the tests can only be used to reject the hypothesis of a distribution around two orthogonal lines. To verify such a hypothesis, it is proposed to transform the data appropriately and then to check the depth function for the transformed data. This approach is applied to check whether micro cracks have an orientation of approximately 45° and 135° to strain in an initial stage.

Keywords: Orthogonal regression through the origin; Tangential data depth; Simplical data depth; Statistical tests; Crack orientation

1 Introduction

The understanding of crack initiation and crack growth is very important for predicting the life time of products as wheels of trains or hip replacement. Many experiments in which material was exposed specific strains were done in the past. Thereby photos of small cracks which can be analyzed with modern methods of pattern recognition (see e.g. Fletcher et al. (2003), Iyer and Sinha 2005, Fujita et al. 2006, Gunkel et al. 2009) were also obtained by microscopes. This provides the possibility to analyze a huge amount of crack data. While the growth of large cracks (macro cracks) follows more or less the deterministic mechanical laws described e.g.

*Research supported by the SFB/TR TRR 30 Project D6

in Pook (2000) and Ludwig et al. (2003), the small cracks (micro cracks) show much more a random behavior which must be described by probabilistic laws. Some first attempts can be found in Ihara and Tanaka (2000) and Brückner-Foit et al. (2003). However, probabilistic laws should be validated by data.

In this article, a new statistical method is described for validating the following hypothesis about the orientation of micro cracks:

A) In an initial stage of the growing process, a micro crack has an orientation with an angle of approximately 45° or 135° ($\pi/4$ or $3\pi/4$) with respect to the strain.

B) When a micro crack becomes a macro crack, its orientation is tending to an orientation perpendicular to the strain.

These hypotheses were checked up to now only by a qualitative analysis regarding selected cracks (see Besel et al. 2008, Besel and Brückner-Foit 2008, Brückner-Foit and Huang 2008). Here, we in particular show how Hypothesis A) can be verified statistically.

Since orientations are circular, the simple mean of the observed angles is a misleading quantity. The only paper which deals with statistical analysis of crack orientations, the paper of Mann et al. (2003) concerning cracks in cemented femoral components, uses circular statistics given by Fisher (1995) and Zar (1999). A mean crack angle is calculated there as $\tan^{-1}\left(\frac{m_{sin}}{m_{cos}}\right)$ where m_{sin} is the mean of $\sin(\alpha_n)$ and m_{cos} is the mean of $\cos(\alpha_n)$ when α_n are the angles of the cracks. With this statistic, it is tested whether the distribution of the angles is uniform distributed and whether the angles in different regions have the same distribution.

But every mean is sensitive to outlying observations. Hence the mean should be replaced by the median. But every mean angle and median angle have the disadvantage that they do not take the length of the cracks into account. Using orthogonal regression through the origin, also the length of the cracks has an impact. In particular, longer cracks have more influence than short cracks, an important property since small cracks are often falsely detected cracks or caused by impurities of the material. While one generalization of the median for regression, the L_1 regression estimator is not outlier robust (see He et al. 1990, Mizera and Müller 1999), another generalization of the median which is based on data depth leads to outlier robust regression estimators. This was shown by Rousseeuw and Hubert (1999) for classical linear regression and by Wellmann and Müller (2008b) for orthogonal regression with intercept. Moreover, Wellmann and Müller (2008b) found in examples that often two orthogonal lines with intercepts have the same data depth. This is a welcome property in view of Hypothesis A). Hence, in this paper we study the data depth for orthogonal regression without intercept. In this special case, we can also give a proof for the observed property that orthogonal lines have the same depth.

Using simplicial data depth, also statistical tests can be derived. This was done by Müller (2005) and Wellmann et al. (2009) for polynomial regression, by Wellmann and Müller (2008a) for multiple regression, and by Wellmann and Müller (2008b) for orthogonal regression with intercept. For regression with intercept, the asymptotic distribution of the simplicial depth is given by an infinite sum of independent χ^2-distributed random variables. But here we show that the asymptotic distribution is given only by one χ^2-distributed random variable if there is no intercept.

The paper is organized as follows. Section 2 provides preliminaries about orthogonal regression and data depth. In Section 3, two depth notions, tangential depth and simplicial depth, are characterized for orthogonal regression without intercept. Some properties of these depth notions for special distributions, in particular for so-called orthogonal distributions, are shown as well. Exact-fit points for one line and for two orthogonal lines are defined as robustness measure and are derived for estimators maximizing the tangential and simplicial depth.

Section 4 deals with the problem of verifying the hypothesis that data are distributed around two orthogonal lines, e.g. the Hypothesis A) for cracks. It is shown that tests based on simplicial depth can reject such hypotheses. But they cannot distinguish between other orthogonal distributions. Therefore an approach using transformed data is proposed for the verification. This approach is applied on crack data in Section 5.

2 Preliminaries

2.1 Classical estimators

Least squares estimators and L_1 estimators for orthogonal regression through the origin can be defined as in classical linear regression by minimizing the sum of squared and absolute residuals. The difference to classical linear regression is the definition of the residuals.

For defining the residuals, it is important to note that a line through the origin in $I\!R^2$ can be expressed in different ways: as slope $\beta \in I\!R$, as angle $\alpha \in [0, \pi)$ between the line and the x-axis, and as vector $a = (a_1, a_2)^\top \in I\!R^2$ so that $\{\lambda a;\ \lambda \in I\!R\}$ contains all points of the line. Preferable, a should satisfy $\|a\| = 1$. The connections between these representations are the following:

$$a = (\cos(\alpha), \sin(\alpha))^\top,\ \beta = a_2/a_1 = \tan(\alpha).$$

In orthogonal regression, the residuals for a given line are given by the length of the difference between the data point $z_n = (x_n, y_n)^\top$ and its perpendicular projection to the line. The perpendicular projection of z_n to a line given by a is

$$a^\top z_n\, a \ \text{ if } \ \|a\| = 1.$$

Hence the absolute residuum is $\mathrm{res}(a, z_n) := \|z_n - a^\top z_n\, a\|$. Let

$$A(\alpha) = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix} \tag{1}$$

the rotation matrix which rotates the vector $a = (\cos(\alpha), \sin(\alpha))^\top$ to the point $(1, 0)^\top$, i.e. $A(\alpha)a = (1, 0)^\top$. Then the squared absolute residuum satisfies

$$\begin{aligned}
&\mathrm{res}(a, z_n)^2 = \|z_n - a^\top z_n\, a\|^2 \\
&= (z_n - a^\top z_n\, a)^\top A(\alpha)^\top A(\alpha)(z_n - a^\top z_n\, a) \\
&= \|A(\alpha)\, z_n - (A(\alpha)\, a)^\top A(\alpha)\, z_n\, A(\alpha)\, a\|^2 \\
&= \left\| \begin{pmatrix} \cos(\alpha)x_n + \sin(\alpha)y_n \\ -\sin(\alpha)x_n + \cos(\alpha)y_n \end{pmatrix} - \begin{pmatrix} \cos(\alpha)x_n + \sin(\alpha)y_n \\ 0 \end{pmatrix} \right\|^2 \\
&= (\sin(\alpha)x_n - \cos(\alpha)y_n)^2 = \|z_n^\top a^\perp\|^2,
\end{aligned}$$

where $a^{\perp} = (\sin(\alpha), -\cos(\alpha))^{\top}$ is the unit vector orthogonal to a. Hence we can set for the residuum also

$$\operatorname{res}(\alpha, z_n) := \sin(\alpha)x_n - \cos(\alpha)y_n. \tag{2}$$

This second representation of a residuum has the advantage that a derivative with respect to α can be easily calculated. This can be used to calculate the least squares estimator by Newton's method.

Definition 1
a) The least squares estimator $\hat{\alpha}$ (LS) for the angle α at $z = (z_1, \ldots, z_N)$ is defined as

$$\widehat{\alpha}_{LS}(z) \in \arg\min_{\alpha \in [0,\pi)} \sum_{n=1}^{N} \operatorname{res}(\alpha, z_n)^2.$$

b) The L_1 estimator $\hat{\alpha}$ (L1) for the angle α at $z = (z_1, \ldots, z_N)$ is defined as

$$\widehat{\alpha}_{L1}(z) \in \arg\min_{\alpha \in [0,\pi)} \sum_{n=1}^{N} |\operatorname{res}(\alpha, z_n)|.$$

2.2 Data depth

The L_1 estimator which minimizes the sum of absolute residuals is one possibility to generalize the outlier robust median to regression. However, this generalization does not lead to outlier robust regression estimators.

Another possibilty to generalize the median to regression is data depth. Tukey (1975) used the half space depth d to generalize the median to multivariate data in $I\!R^q$. The depth $d(\mu, z)$ of a location parameter $\mu \in I\!R^q$ in a sample $z = (z_1, \ldots, z_N)$ in $I\!R^q$ is the minimum number of observations $z_1, \ldots, z_N$ lying in a half space containing μ. The parameter μ which maximizes $d(\mu, z)$ is the generalization of the median. Rousseeuw and Hubert (1999) generalized the half space depth to linear regression by introducing the notion of a nonfit:

Definition 2 (Original definition of a nonfit) *A regression parameter β is a nonfit for $z_1, \ldots, z_N$ if there is another parameter $\tilde{\beta}$ so that the residuals* $\operatorname{res}(\tilde{\beta}, z_n)$ *satisfy*

$$\operatorname{res}(\tilde{\beta}, z_n)^2 < \operatorname{res}(\beta, z_n)^2 \quad \textit{for} \quad n = 1, \ldots, N,$$

which means that the regression function given by $\tilde{\beta}$ is closer to the data points $z_1, \ldots, z_N \in I\!R^2$ than the regression function given by β.

Then the regression depth $d_R(\beta, z)$ of the regression parameter β in the data set $z_1, \ldots, z_N$ is the minimum number M of observations $z_{n_1}, \ldots, z_{n_M}$ which must be removed so that β becomes a nonfit in $\{z_1, \ldots, z_N\} \setminus \{z_{n_1}, \ldots, z_{n_M}\}$.

The original Definition 2 is difficult to handle. Therefore a tangential version is usually used:

Definition 3 (Tangential version of a nonfit) *A regression parameter* $\beta \in I\!R^q$ *is a nonfit for* $z_1, \ldots, z_N$ *if there exists a vector* $u \in I\!R^q$ *with*

$$u^\top \frac{\partial}{\partial \beta} \mathrm{res}(\beta, z_n)^2 < 0 \quad for \quad n = 1, \ldots, N.$$

For classical linear regression, where the residuals are linear in β, the two definitions are identical. But Mizera (2002) pointed out that there are many situations where they are different. He called a depth based on Definition 2 global depth and a depth based on Definition 3 tangential depth. Then the tangential depth $d_T(\beta, z)$ of a parameter $\beta \in I\!R^q$ in $z_1, \ldots, z_N$ has the following simple definition

$$d_T(\beta, z) = \frac{1}{N} \min_{0 \neq u \in I\!R^q} \sharp\{n \in \{1, \ldots, N\};\ u^\top \frac{\partial}{\partial \beta} \mathrm{res}(\beta, z_n)^2 \geq 0\},$$

where $\sharp$ denotes the cardinality of a set.

However, any tangential depth has the disadvantage that it is difficult to derive its finite sample distribution and its asymptotic distribution so that tests based on it are difficult to define. Only few approaches exist for regression depth for classical regression. Bai and He (1999) only derived an implicitly given asymptotic distribution of the maximum regression depth estimator while Van Aelst et al. (2002) derived an exact test based on the regression depth only for linear regression. The development of tests becomes much easier by using the simplicial depth.

Simplicial depth of a multivariate location parameter $\mu \in I\!R^q$ was introduced by Liu (1988, 1990) using the half space depth d of Tukey (1975). She defined it as

$$d_S(\mu, (z_1, ..., z_N)) = \binom{N}{q+1}^{-1} \sum_{1 \leq n_1 < n_2 < ... < n_{q+1} \leq N} I\!I\{d(\mu, (z_{n_1}, ..., z_{n_{q+1}})) > 0\}, \quad (3)$$

where $I\!I$ denotes the indicator function. This depth counts the simplexes spanned by $q+1$ data points which are containing the parameter μ. Replacing the half space depth d by any other depth notion leads to a very general concept of simplicial depth. Any notion of simplicial depth has the advantage that it is an U-statistics and for U-statistics the asymptotic distribution is in principal known from Hoeffding's theorem (see e.g. Lee 1990, p. 79, 80, 90). This advantage was used in Müller (2005), Wellmann (2007), Wellmann et al. (2009), Wellmann and Müller (2008a) to derive distribution free tests for polynomial and multiple regression. See also Wellmann et al. (2007) for the calculation of maximum simplicial depth.

3 Depth estimators for orthogonal regression through the origin

3.1 Depth notions for data

Global and tangential depth are for example different for orthogonal regression as Mizera (2002) already noticed and which was worked out by Wellmann (2007) and

Wellmann and Müller (2008b). Wellmann (2007) and Wellmann and Müller (2008b) considered only orthogonal regression for lines with intercept. In this case, also the use of the tangential depth is rather complicated.

For orthogonal regression through the origin, everything becomes much more simple. At first note that the derivative of the residuals are given by

$$\begin{aligned}\frac{\partial}{\partial \alpha}\mathrm{res}(\alpha, z_n)^2 &= \frac{\partial}{\partial \alpha}(\sin(\alpha)x_n - \cos(\alpha)y_n)^2 \\ &= 2\,(\sin(\alpha)x_n - \cos(\alpha)y_n)\,(\cos(\alpha)x_n + \sin(\alpha)y_n) = -2\,A_2(\alpha)^\top z_n\; A_1(\alpha)^\top z_n,\end{aligned}$$

where $A_1(\alpha)^\top$ and $A_1(\alpha)^\top$ are the rows of the rotation matrix $A(\alpha)$ given in (1), i.e. $A(\alpha) = \begin{pmatrix} A_1(\alpha)^\top \\ A_2(\alpha)^\top \end{pmatrix}$. Hence tangential depth for orthogonal regression through the origin can be defined as follows.

Definition 4 (Tangential depth for orthogonal regression through the origin)
The tangential depth $d_T(\alpha, z)$ of an angle $\alpha \in I\!R$ in $z_1, \ldots, z_N \in I\!R^2$ is defined as

$$d_T(\alpha, z) = \frac{1}{N}\min\{\sharp\{n;\ A_2(\alpha)^\top z_n\; A_1(\alpha)^\top z_n \geq 0\}, \sharp\{n;\ A_2(\alpha)^\top z_n\; A_1(\alpha)^\top z_n \leq 0\}\}.$$

For $\alpha = 0$, i.e. for a horizontal line, we obtain

$$d_T(\alpha, z) = \frac{1}{N}\min\{\sharp\{n;\ x_n\, y_n \geq 0\}, \sharp\{n;\ x_n\, y_n \leq 0\}\}.$$

This is the same definition of the depth of a horizontal line as for classical regression through the origin. However, for other lines the definitions are different since for classical regression through the origin the derivative of the residuals is

$$\frac{\partial}{\partial \beta}\mathrm{res}(\beta, z_n)^2 = \frac{\partial}{\partial \beta}(y_n - \beta\, x_n)^2 = 2\,(y_n - \beta\, x_n)\, x_n.$$

Definition 4 for orthogonal regression can be interpreted as follows: The data are rotated with the rotation matrix $A(\alpha)$ so that the line given by α is the horizontal line. Then the tangential depth for classical regression through the origin is used for the horizontal line and the rotated data. This interpretation was also used by Wellmann and Müller (2008b) for orthogonal regression for a line with intercept.

Lemma 1 *Orthogonal lines have the same depth, i.e.*

$$d_T(\alpha, z) = d_T(\alpha + \pi/2, z).$$

Proof. The assertion follows from

$$\begin{aligned}&\frac{\partial}{\partial \alpha}\mathrm{res}(\alpha + \pi/2, z_n)^2 \\ &= 2\,(\sin(\alpha + \pi/2)x_n - \cos(\alpha + \pi/2)y_n)\,(\cos(\alpha + \pi/2)x_n + \sin(\alpha + \pi/2)y_n) \\ &= 2\,(\cos(\alpha)x_n + \sin(\alpha)y_n)\,(-\sin(\alpha)x_n + \cos(\alpha)y_n) \\ &= -\frac{\partial}{\partial \alpha}\mathrm{res}(\alpha, z_n)^2.\square\end{aligned}$$

That orthogonal lines have the same tangential depth was also observed by examples in Wellmann and Müller (2008b) for orthogonal regression for a line with intercept. A proof was not given there. The examples in Wellmann and Müller (2008b) also showed that this property is not satisfied for the global depth so that global depth for orthogonal regression through the origin should have the same property.

To derive tests, the simplicial depth based on the tangential depth given by Definition 4 is introduced here as well.

Definition 5 (Simplicial depth for orthogonal regression through the origin)
The simplicial depth $d_S(\alpha, z)$ of an angle $\alpha \in I\!R$ in $z_1, \ldots, z_N \in I\!R^2$ is defined as

$$d_S(\alpha,(z_1,...,z_N)) = \binom{N}{2}^{-1} \sum_{1\le n_1<n_2\le N} I\!I\{d_T(\alpha,(z_{n_1},z_{n_2}))>0\}.$$

Lemma 2 *The simplicial depth for orthogonal regression through the origin satisfies*

$$\begin{aligned} & d_S(\alpha,(z_1,...,z_N)) \\ &= \binom{N}{2}^{-1}\left(\mathrm{neg}_z(\alpha)\mathrm{pos}_z(\alpha)+\mathrm{neg}_z(\alpha)\mathrm{zero}_z(\alpha)+\mathrm{pos}_z(\alpha)\mathrm{zero}_z(\alpha)+\binom{\mathrm{zero}_z(\alpha)}{2}\right)\end{aligned}$$

where

$$\begin{aligned} \mathrm{neg}_z(\alpha) &= \sharp\{n;\ A_1(\alpha)^\top z_n\, A_2(\alpha)^\top z_n < 0\}, \\ \mathrm{pos}_z(\alpha) &= \sharp\{n;\ A_1(\alpha)^\top z_n\, A_2(\alpha)^\top z_n > 0\}, \\ \mathrm{zero}_z(\alpha) &= \sharp\{n;\ A_1(\alpha)^\top z_n\, A_2(\alpha)^\top z_n = 0\}. \end{aligned}$$

Proof. Since

$$d_T(\alpha,(z_{n_1},z_{n_2}))=0$$

if and only if $A_1(\alpha)^\top z_{n_1}\, A_2(\alpha)^\top z_{n_1}$ and $A_1(\alpha)^\top z_{n_2}\, A_2(\alpha)^\top z_{n_2}$ are both positive or both negative we have

$$d_T(\alpha,(z_{n_1},z_{n_2}))>0$$

if and only if $A_1(\alpha)^\top z_{n_1}\, A_2(\alpha)^\top z_{n_1}$ and $A_1(\alpha)^\top z_{n_2}\, A_2(\alpha)^\top z_{n_2}$ have different signs or at least one of them is zero. Hence, the assertion follows. □

Since orthogonal lines have the same tangential depth for orthogonal regression through the origin, they have also the same simplicial depth, i.e. we have $d_S(\alpha, z) = d_S(\alpha + \pi/2, z)$.

3.2 Depth notions for distributions

It is straightforward to generalize the tangential depth given in Definition 4 to arbitrary distributions P^Z where Z is an arbitrary random variable on $I\!R^2$.

Definition 6 (Tangential depth for distributions)
The tangential depth $d_T(\alpha, P^Z)$ of an angle $\alpha \in I\!R$ at distribution P^Z is defined as

$$\begin{aligned} d_T(\alpha, P^Z) &= \min\big\{P^Z(\{z \in I\!R^2;\ A_2(\alpha)^\top z\, A_1(\alpha)^\top z \geq 0\}), \\ &\qquad P^Z(\{z \in I\!R^2;\ A_2(\alpha)^\top z_n\, A_1(\alpha)^\top z_n \leq 0\})\big\}. \end{aligned}$$

To generalize the simplicial depth given in Definition 5 to distributions, note that

$$d_S(\alpha, (z_1, ..., z_N)) = \frac{1}{N\,(N-1)} \sum_{n_1 \neq n_2} I\!I\{d_T(\alpha, (z_{n_1}, z_{n_2})) > 0\}.$$

Definition 7 (Simplicial depth for distributions)
The simplicial depth $d_S(\alpha, P^Z)$ of an angle $\alpha \in I\!R$ at distribution P^Z is defined as

$$d_S(\alpha, P^Z) = P^{Z_1, Z_2}(\{(z_1, z_2);\ d_T(\alpha, (z_1, z_2)) > 0\}),$$

where Z_1 and Z_2 are independent random variables with $P^{Z_1} = P^{Z_2} = P^Z$.

The following lemma is analogous to Lemma 2.

Lemma 3 *The simplicial depth at distribution P^Z satisfies*

$$\begin{aligned} &d_S(\alpha, (z_1, ..., z_N)) = 2\,P^Z(\mathrm{Neg}(\alpha))\,P^Z(\mathrm{Pos}(\alpha)) \\ &\quad + 2\,P^Z(\mathrm{Neg}(\alpha))\,P^Z(\mathrm{Zero}(\alpha)) + 2\,P^Z(\mathrm{Pos}(\alpha))\,P^Z(\mathrm{Zero}(\alpha)) + P^Z(\mathrm{Zero}(\alpha))^2, \end{aligned}$$

where

$$\begin{aligned} \mathrm{Neg}(\alpha) &= \{z \in I\!R^2;\ A_1(\alpha)^\top z\, A_2(\alpha)^\top z < 0\}, \\ \mathrm{Pos}(\alpha) &= \{z \in I\!R^2;\ A_1(\alpha)^\top z\, A_2(\alpha)^\top z > 0\}, \\ \mathrm{Zero}(\alpha) &= \{z \in I\!R^2;\ A_1(\alpha)^\top z\, A_2(\alpha)^\top z = 0\}. \end{aligned}$$

Proof. Since (see the proof of Lemma 2)

$$\begin{aligned} &P^{Z_1, Z_2}(\{(z_1, z_2);\ d_T(\alpha, (z_1, z_2)) > 0\}) \\ &= P^{Z_1, Z_2}(\{(z_1, z_2);\ z_1 \in \mathrm{Neg}(\alpha),\ z_2 \in \mathrm{Pos}(\alpha) \text{ or } z_2 \in \mathrm{Neg}(\alpha),\ z_1 \in \mathrm{Pos}(\alpha) \text{ or} \\ &\quad z_1 \in \mathrm{Neg}(\alpha),\ z_2 \in \mathrm{Zero}(\alpha) \text{ or } z_2 \in \mathrm{Neg}(\alpha),\ z_1 \in \mathrm{Zero}(\alpha) \text{ or} \\ &\quad z_1 \in \mathrm{Pos}(\alpha),\ z_2 \in \mathrm{Zero}(\alpha) \text{ or } z_2 \in \mathrm{Pos}(\alpha),\ z_1 \in \mathrm{Zero}(\alpha) \text{ or} \\ &\quad z_2 \in \mathrm{Zero}(\alpha),\ z_1 \in \mathrm{Zero}(\alpha)\}), \end{aligned}$$

the assertion follows from the independence of Z_1 and Z_2. □

As for data, orthogonal lines have the same depth, i.e. $d_T(\alpha + \pi/2, P^Z) = d_T(\alpha, P^Z)$ and $d_S(\alpha + \pi/2, P^Z) = d_S(\alpha, P^Z)$. But for special distributions, the depth can be the same for all lines and angles, respectively. We call these special distributions orthogonal distributions:

Definition 8 *P^Z is an orthogonal distribution on $I\!R^2$ if*

$$P^{Z/\|Z\|} = P^{\rho(Z)/\|\rho(Z)\|}$$

is satisfied for any rotation ρ about $\pm 90°$, i.e. for

$$\rho_1(z) := \rho_1\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \text{ and } \rho_2(z) := \rho_2\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Rotation invariant distributions like the uniform distribution on the two dimensional circle or disk are orthogonal distributions. But an orthogonal distribution can also be concentrated on or around two orthogonal lines. Namely, if

$$L_1 = \{z \in I\!R^2;\ z = \lambda(\cos(\alpha), \sin(\alpha))^\top \text{ for } \alpha \in [\alpha_0 - \alpha_1, \alpha_0 + \alpha_1],\ \lambda \in I\!R\}$$

and

$$L_2 = \{z \in I\!R^2;\ \text{ there exists } z_* \in L_1 \text{ with } z^\top z_* = 0\}$$

are areas around two orthogonal lines so that

$$P^{\rho_1(Z)}(L_1) = P^{\rho_2(Z)}(L_1) = P^Z(L_1) = P^Z(L_2) = P^{\rho_1(Z)}(L_2) = P^{\rho_2(Z)}(L_2)$$

and

$$P^Z(L_1 \cup L_2) = 1,$$

then P^Z is an orthogonal distribution. Such kind of distribution is concentrated on two orthogonal lines if $\alpha_1 = 0$. In this case, only $P^Z(L_1) = P^Z(L_2)$ and $P^Z(L_1 \cup L_2) = 1$ must be checked.

Theorem 1

a) If P^Z is an orthogonal distribution and $P^{Z/\|Z\|}$ is an absolute continuous distribution, then

$$d_T(\alpha, P^Z) = d_S(\alpha, P^Z) = \frac{1}{2} \quad \textit{for all} \quad \alpha \in [0, \pi).$$

b) If P^Z is an orthogonal distribution with $P\left(Z = \binom{0}{0}\right) = 0$ which is concentrated on two lines given by α_0 and $\alpha_0 + \pi/2$, then

$$d_T(\alpha, P^Z) = d_S(\alpha, P^Z) = 1 \quad \textit{for} \quad \alpha = \alpha_0 \textit{ and } \alpha = \alpha_0 + \pi/2$$

and

$$d_T(\alpha, P^Z) = d_S(\alpha, P^Z) = \frac{1}{2} \quad \textit{for all} \quad \alpha \in [0, \pi) \setminus \{\alpha_0, \alpha_0 + \pi/2\}.$$

c) If P^Z is a distribution with $P\left(Z = \binom{0}{0}\right) = 0$ which is concentrated on one line given by α_0, then

$$d_T(\alpha, P^Z) = d_S(\alpha, P^Z) = 1 \quad \textit{for} \quad \alpha = \alpha_0 \textit{ and } \alpha = \alpha_0 + \pi/2$$

and

$$d_T(\alpha, P^Z) = d_S(\alpha, P^Z) = 0 \quad \textit{for all} \quad \alpha \in [0, \pi) \setminus \{\alpha_0, \alpha_0 + \pi/2\}.$$

Proof. At first note

$$d_T(\alpha, P^Z) = \min\{P^Z(Neg(\alpha) \cup \text{Zero}(\alpha)), P^Z(Pos(\alpha) \cup \text{Zero}(\alpha))\}$$

and $P^Z(\text{Neg}(\alpha)) = P^{Z/\|Z\|}(\text{Neg}(\alpha))$, $P^Z(\text{Pos}(\alpha)) = P^{Z/\|Z\|}(\text{Pos}(\alpha))$, $P^Z(\text{Zero}(\alpha)) = P^{Z/\|Z\|}(\text{Zero}(\alpha))$. The orthogonality of P^Z and the definitions of $A_1(\alpha)$ and $A_2(\alpha)$ provide

$$\begin{aligned}
P^Z(Pos(\alpha)) &= P^{\rho_1(Z)}(Pos(\alpha)) = P(A_1(\alpha)^\top \rho_1(Z)\ A_2(\alpha)^\top \rho_1(Z) > 0) \\
&= P\left(A_1(\alpha)^\top \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} Z\ A_2(\alpha)^\top \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} Z > 0\right) \\
&= P(\ (\sin(\alpha), -\cos(\alpha))Z\ (\cos(\alpha), \sin(\alpha))Z > 0) \\
&= P(-A_2(\alpha)^\top Z\ A_1(\alpha)^\top Z > 0) \\
&= P(A_1(\alpha)^\top Z\ A_2(\alpha)^\top Z < 0) = P^Z(Neg(\alpha)).
\end{aligned}$$

a) If $P^{Z/\|Z\|}$ is an absolute continuous distribution, then

$$P^{Z/\|Z\|}(\text{Zero}(\alpha)) = P\left(\frac{Z}{\|Z\|} = \begin{pmatrix}\cos(\alpha)\\ \sin(\alpha)\end{pmatrix} \text{ or } \frac{Z}{\|Z\|} = \begin{pmatrix}-\cos(\alpha)\\ -\sin(\alpha)\end{pmatrix} \text{ or } \frac{Z}{\|Z\|} = \begin{pmatrix}\sin(\alpha)\\ -\cos(\alpha)\end{pmatrix} \text{ or } \frac{Z}{\|Z\|} = \begin{pmatrix}-\sin(\alpha)\\ \cos(\alpha)\end{pmatrix}\right) = 0,$$

so that $P^Z(Neg(\alpha)) = P^Z(Pos(\alpha)) = \frac{1}{2}$ for all $\alpha \in [0, \pi)$.
b) If P^Z is concentrated on the two lines given by α_0 and $\alpha_0 + \pi/2$, then

$$P^{Z/\|Z\|}(\text{Zero}(\alpha_0)) = 1 = P^{Z/\|Z\|}(\text{Zero}(\alpha_0 + \pi/2))$$

and

$$P^{Z/\|Z\|}(\text{Zero}(\alpha)) = 0$$

for all $\alpha \in [0, \pi) \setminus \{\alpha_0, \alpha_0 + \pi/2\}$ because of $P\left(Z = \binom{0}{0}\right) = 0$.
c) If $P(Z \in L_1) = 1$ with $L_1 = \{z \in I\!R^2;\ z = \lambda(\cos(\alpha_0), \sin(\alpha_0))^\top,\ \lambda \in I\!R\}$ then $P(A_2(\alpha_0)Z = 0) = 1$ and $P(A_1(\alpha_0 + \pi/2)Z = 0) = 1$ so that $P^Z(\text{Zero}(\alpha_0)) = 1 = P^Z(\text{Zero}(\alpha_0 + \pi/2))$. For $\alpha \in [0, \pi) \setminus \{\alpha_0, \alpha_0 + \pi/2\}$, we obtain

$$A_1(\alpha)^\top z\ A_2(\alpha)^\top z = \lambda^2 \begin{pmatrix}\cos(\alpha)\\ \sin(\alpha)\end{pmatrix}^\top \begin{pmatrix}\cos(\alpha_0)\\ \sin(\alpha_0)\end{pmatrix} \begin{pmatrix}-\sin(\alpha)\\ \cos(\alpha)\end{pmatrix}^\top \begin{pmatrix}\cos(\alpha_0)\\ \sin(\alpha_0)\end{pmatrix}$$

for all $z \in L_1$ which is either positive or negative. Hence it holds either $P^Z(\text{Pos}(\alpha)) = 0 = P^Z(\text{Zero}(\alpha))$ or $P^Z(\text{Neg}(\alpha)) = 0 = P^Z(\text{Zero}(\alpha))$. $\square$

3.3 Depth estimators

As soon as a depth notion is given, an estimator can be defined as that parameter with maximum depth. Hence, the tangential and simplicial depth for orthogonal regression through the origin lead to the following definitions:

Definition 9 (Depth estimators)
a) The tangential depth estimator $\widehat{\alpha}_T(z)$ is defined as

$$\widehat{\alpha}_T(z) \in \arg\max_{\alpha \in [0,\pi)} d_T(\alpha, z).$$

b) The simplicial depth estimator $\widehat{\alpha}_S(z)$ is defined as

$$\widehat{\alpha}_S(z) \in \arg \max_{\alpha\in[0,\pi)} d_S(\alpha, z).$$

Note that the tangential and simplicial depth estimators are never unique since orthogonal lines have the same depth (see Lemma 1). Hence with $\widehat{\alpha}_T(z)$ and $\widehat{\alpha}_S(z)$, also $\widehat{\alpha}_T(z) + \pi/2$ and $\widehat{\alpha}_S(z) + \pi/2$ are depth estimators.

Figure 1 compares the tangential depth estimator with the least squares estimator, the L_1 estimator and the mean of the observed angles in the presence of 30% outliers. It shows that the least squares estimator as well as the L_1 estimator are heavily influenced by the outliers. The tangential depth estimator provides two lines: one which follows the majority of the data and one orthogonal to the other line. This orthogonal line is close to the line given by the mean of the angles. Hence the mean provides a line which is also far away from the majority of the data.

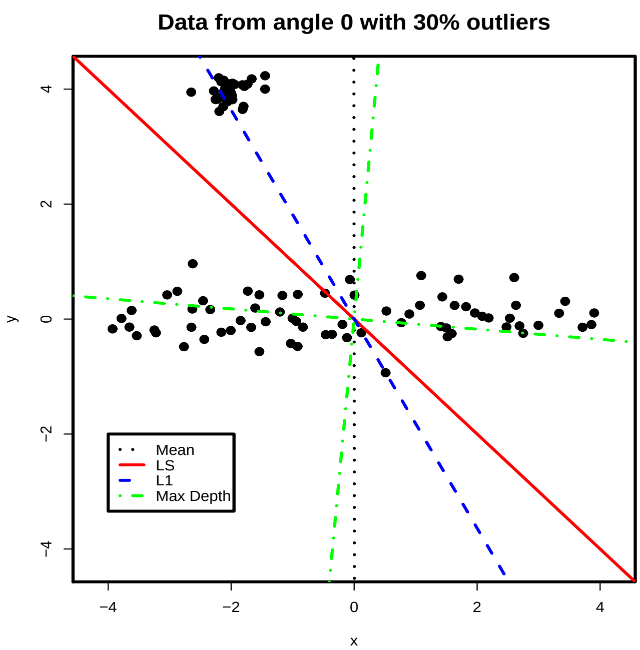

Figure 1: *Tangential depth estimator compared with mean angle, LS and L1 estimate for data with outliers*

To quantify the outlier robustness of the estimators, a robustness measure shall be used. A well known robustness measure is the breakdown point of Donoho and Huber (1983). But the breakdown point makes only sense if the parameter space is bounded by infinity or some other bounds so that convergence to such bounds means breakdown. But considering angles, the parameter space is circular so that no bound exists. However, the exact-fit point defined in Ellis and Morgenthaler (1992) can be used which has in other cases a strong relation to the breakdown point. To define the exact-fit point for orthogonal regression through the origin, let be $L(\alpha) := L_1(\alpha) \cup L_2(\alpha)$, where

$$L_1(\alpha) := \{z \in I\!R^2;\ z = \lambda(\cos(\alpha), \sin(\alpha))^\top \text{ for } \lambda \in I\!R\}$$

and

$$L_2(\alpha) := \{z \in I\!R^2; \;\; \text{there exists } z_* \in L_1(\alpha) \text{ with } z^\top z_* = 0\}.$$

We distinguish between two exact-fit properties: one where all data are lying on one line through the origin, i.e. $z_n \in L_1(\alpha)$ for $n = 1, \dots, N$, and one where all data are lying on two orthogonal lines, i.e. $z_n \in L(\alpha)$ for $n = 1, \dots, N$.

Definition 10 (Exact-fit point for orthogonal regression through the origin) *The fit point of an estimator $\widehat{\alpha}$ for α at a sample $z = (z_1, \dots, z_N)$ is defined as*

$$\epsilon(\widehat{\alpha}, z) = \frac{1}{N} \min\{M; \;\; \textit{there exists } \tilde{z} \in \mathcal{Z}_M(z) \textit{ such that } \widehat{\alpha}(\tilde{z}) \notin \{\widehat{\alpha}(z), \widehat{\alpha}(z) + \pi/2\}\},$$

where

$$\begin{aligned} \mathcal{Z}_M(z) \;\; := \;\; & \{(\tilde{z}_1, \dots, \tilde{z}_N); \;\; \textit{there exists } m_1, \dots, m_{N-M} \\ & \qquad \textit{such that } z_{m_i} = \tilde{z}_{m_i} \textit{ for } i = 1, \dots, N - M\}. \end{aligned}$$

a) The exact-fit point for one line of an estimator $\widehat{\alpha}$ for α is defined as

$$\epsilon^*(\widehat{\alpha}) = \min\{\epsilon(\widehat{\alpha}, z); \;\; \textit{there exists } \alpha \textit{ such that } z_1, \dots, z_N \in L_1(\alpha)\}.$$

b) The exact-fit point for two orthogonal lines of an estimator $\widehat{\alpha}$ for α is defined as

$$\epsilon^{**}(\widehat{\alpha}) = \min\{\epsilon(\widehat{\alpha}, z); \;\; \textit{there exists } \alpha \textit{ such that } z_1, \dots, z_N \in L(\alpha)\}.$$

Theorem 2
a) The exact-fit points for one and two lines of the least squares estimator and the L_1 estimator are given by

$$\begin{aligned} \epsilon^*(\widehat{\alpha}_{LS}) &= \frac{1}{N} = \epsilon^*(\widehat{\alpha}_{L1}), \\ \epsilon^{**}(\widehat{\alpha}_{LS}) &= \frac{1}{N} = \epsilon^{**}(\widehat{\alpha}_{L1}). \end{aligned}$$

b) The exact fit points for one and two lines of the tangential depth estimator are given by

$$\begin{aligned} \epsilon^*(\widehat{\alpha}_T) &= \frac{1}{N} \left\lceil \frac{N}{2} \right\rceil \stackrel{N \to \infty}{\longrightarrow} \frac{1}{2}, \\ \epsilon^{**}(\widehat{\alpha}_T) &\in \left[\frac{1}{N} \left\lceil \frac{N}{3} \right\rceil, \frac{1}{N} \left\lceil \frac{N+1}{3} \right\rceil \right] \stackrel{N \to \infty}{\longrightarrow} \frac{1}{3}. \end{aligned}$$

c) The exact fit points for one and two lines of the simplicial depth estimator are given by

$$\begin{aligned} \epsilon^*(\widehat{\alpha}_S) &= \frac{1}{N} \left\lceil \frac{N}{2} \right\rceil \stackrel{N \to \infty}{\longrightarrow} \frac{1}{2}, \\ \epsilon^{**}(\widehat{\alpha}_S) &\in \left[\frac{1}{N} \left\lceil -N + 2 + \sqrt{2N^2 - 6N + 4} \right\rceil, \frac{1}{N} \left\lceil -N + 2 + \sqrt{2N^2 - 6N + 5} \right\rceil \right] \\ & \stackrel{N \to \infty}{\longrightarrow} \sqrt{2} - 1 = 0.4142136. \end{aligned}$$

The proof of Theorem 2 is given in the Appendix.

4 Verification of a distribution around orthogonal lines

4.1 Testing parameters of orthogonal lines

To check whether the observations are distributed around orthogonal lines, one possibility is to test a hypothesis of the form $H_0 : \alpha \in [\alpha_0 - \eta, \alpha_0 + \eta] \cup [\alpha_0 + \pi/2 - \eta, \alpha_0 + \pi/2 + \eta]$, where $\eta < \pi/4$. Tests for such hypotheses can be based on the simplicial depth for orthogonal regression through the origin. For deriving the asymptotic distribution of this simplicial depth, we need some distributional assumptions: the observations $z_1 = (x_1, y_1)^\top, \ldots, z_N = (x_N, y_N)^\top$ are realizations of independent and identically distributed random variables $Z_1 = (X_1, Y_1)^\top, \ldots, Z_N = (X_N, Y_N)^\top$ with

$$Z_n = \begin{pmatrix} X_n \\ Y_n \end{pmatrix} = \begin{pmatrix} U_n \\ V_n \end{pmatrix} + \begin{pmatrix} D_n \\ E_n \end{pmatrix}.$$

We make the following assumptions on U_n, V_n, D_n, E_n:

(A) There exists $\epsilon \in [0, 1]$ and $a, b \in I\!R^2$ with corresponding angles $\alpha \in [0, \pi/2)$ and $\alpha + \pi/2$ such that $a^\top b = 0$, $\|a\| = 1 = \|b\|$ and

$$P_{\alpha,\epsilon}\left(a^\top \begin{pmatrix} U_n \\ V_n \end{pmatrix} = 0\right) = \epsilon, \quad P_{\alpha,\epsilon}\left(b^\top \begin{pmatrix} U_n \\ V_n \end{pmatrix} = 0\right) = 1 - \epsilon.$$

(B) The distribution of $\begin{pmatrix} D_n \\ E_n \end{pmatrix}$ is invariant with respect to all rotations, i.e. there exist random variables B_n and C_n such that

$$A(\alpha) \begin{pmatrix} D_n \\ E_n \end{pmatrix} \sim F \sim \begin{pmatrix} B_n \\ C_n \end{pmatrix}; \text{ for all } \alpha \in [0, 2\pi],$$

where $A(\alpha)$ is the rotation matrix given in (1) and F is a rotation invariant distribution.

(C) Additionally, we assume that

$$\begin{pmatrix} U_n \\ V_n \end{pmatrix},\ B_n,\ C_n \text{ are stochastically independent and have continuous distributions.}$$

Condition (C) in particular implies $P_{\alpha,\epsilon}\left(\binom{U_n}{V_n} = 0\right) = 0$ so that Condition (A) means $P_{\alpha,\epsilon}\left(a^\top \binom{U_n}{V_n} = 0 \text{ or } b^\top \binom{U_n}{V_n} = 0\right) = 1$. This implies that the random vectors $\binom{U_n}{V_n}$ are lying on the two orthogonal lines given by a and b almost surely. If $\epsilon = 0$ or $\epsilon = 1$, then the Assumptions (A), (B), and (C) are often made in models where orthogonal regression shall be used like in errors-in-variables models.

According to Condition (B), the distribution of $\binom{D_n}{E_n}$ is invariant with respect to rotations about $\pm 90°$. If we additionally assume that $\binom{U_n}{V_n}$ has this invariance property as well, then $P^{\rho(Z_n)} = P^{Z_n}$ for any rotation ρ about $\pm 90°$. In particular, P^{Z_n} is an orthogonal distribution in the sense of Definition 8. The invariance of $\binom{U_n}{V_n}$ implies for example for $\rho_1(z) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} z$, $a = (\cos(\alpha), \sin(\alpha))^\top$, and $b = (-\sin(\alpha), \cos(\alpha))^\top$

$$1-\epsilon = P_{\alpha,\epsilon}\left(b^\top \begin{pmatrix} U_n \\ V_n \end{pmatrix} = 0\right) = P_{\alpha,\epsilon}\left(b^\top \rho_1\left(\begin{pmatrix} U_n \\ V_n \end{pmatrix}\right) = 0\right) = P_{\alpha,\epsilon}\left(a^\top \begin{pmatrix} U_n \\ V_n \end{pmatrix} = 0\right) = \epsilon$$

because of $b^\top \rho_1(z) = a^\top z$. Hence ϵ must be $\frac{1}{2}$ for orthogonal distributions.

The asymptotic distribution for the simplicial depth is obtained via Hoeffding's theorem (see e.g. Lee 1990, p. 79, 80, 90). For that, we need the conditional expectation of the kernel function $I\!I\{d_T(\alpha, (Z_1, Z_2)) > 0\}$ of the U-statistics d_S given $Z_1 = z_1$.

Lemma 4

$$\mathrm{E}_{\alpha,\epsilon}(I\!I\{d_T(\alpha, (Z_1, Z_2)) > 0\} | Z_1 = z_1) = \frac{1}{2}.$$

Proof. Let α be the angle corresponding to the vector a providing the line on which the random variables $(U_n, V_n)^\top$ are lying. Then the rotation matrix $A(\alpha)$ given by (1) satisfies

$$A(\alpha) = \begin{pmatrix} A_1(\alpha)^\top \\ A_2(\alpha)^\top \end{pmatrix} = \begin{pmatrix} a^\top \\ b^\top \end{pmatrix}.$$

This implies using Condition (B)

$$B_n \sim A_1(\alpha) \begin{pmatrix} D_n \\ E_n \end{pmatrix} = a^\top \begin{pmatrix} D_n \\ E_n \end{pmatrix}, \; C_n \sim A_2(\alpha) \begin{pmatrix} D_n \\ E_n \end{pmatrix} = b^\top \begin{pmatrix} D_n \\ E_n \end{pmatrix}. \tag{4}$$

Since $d_T(\alpha, (Z_1, Z_2)) > 0$ if and only if $A_1(\alpha)^\top z_1 \; A_2(\alpha)^\top z_1 \; A_1(\alpha)^\top z_2 \; A_2(\alpha)^\top z_2 = a^\top z_1 \, b^\top z_1 \, a^\top z_2 \, b^\top z_2 \leq 0$, we have for z_1 with $a^\top z_1 \, b^\top z_1 \geq 0$

$$\begin{aligned} & \mathrm{E}_{\alpha,\epsilon}(I\!I\{d_T(\alpha, (Z_1, Z_2)) > 0\} | Z_1 = z_1) \\ = \;& P_{\alpha,\epsilon}(a^\top Z_1 \, b^\top Z_1 \, a^\top Z_2 \, b^\top Z_2 \leq 0 | a^\top Z_1 \, b^\top Z_1 \geq 0) \\ = \;& P_{\alpha,\epsilon}(a^\top Z_2 \;\; b^\top Z_2 \leq 0) \\ \overset{(A),(C)}{=} \;& P_{\alpha,\epsilon}(a^\top Z_2 \;\; b^\top Z_2 \leq 0, \;\; a^\top W_2 = 0) \\ & + P_{\alpha,\epsilon}(a^\top Z_2 \;\; b^\top Z_2 \leq 0, \;\; b^\top W_2 = 0), \end{aligned}$$

where $W_2 = \binom{U_2}{V_2}$. The first summand satisfies

$$
\begin{aligned}
& P_{\alpha,\epsilon}(a^\top Z_2 \quad b^\top Z_2 \leq 0, \;\; a^\top W_2 = 0) \\
= \quad & P_{\alpha,\epsilon}\left(a^\top \begin{pmatrix} X_2 \\ Y_2 \end{pmatrix} \leq 0, \;\; b^\top \begin{pmatrix} X_2 \\ Y_2 \end{pmatrix} \geq 0, \;\; a^\top W_2 = 0\right) \\
& + P_{\alpha,\epsilon}\left(a^\top \begin{pmatrix} X_2 \\ Y_2 \end{pmatrix} \geq 0, \;\; b^\top \begin{pmatrix} X_2 \\ Y_2 \end{pmatrix} \leq 0, \;\; a^\top W_2 = 0\right) \\
= \quad & P_{\alpha,\epsilon}\left(a^\top W_2 + a^\top \begin{pmatrix} D_2 \\ E_2 \end{pmatrix} \leq 0, \;\; b^\top W_2 + b^\top \begin{pmatrix} D_2 \\ E_2 \end{pmatrix} \geq 0, \;\; a^\top W_2 = 0\right) \\
& + P_{\alpha,\epsilon}\left(a^\top W_2 + a^\top \begin{pmatrix} D_2 \\ E_2 \end{pmatrix} \geq 0, \;\; b^\top W_2 + b^\top \begin{pmatrix} D_2 \\ E_2 \end{pmatrix} \leq 0, \;\; a^\top W_2 = 0\right) \\
\overset{(4)}{=} \quad & P_{\alpha,\epsilon}\left(0 + B_2 \leq 0, \;\; b^\top W_2 + C_2 \geq 0, \;\; a^\top W_2 = 0\right) \\
& + P_{\alpha,\epsilon}\left(0 + B_2 \geq 0, \;\; b^\top W_2 + C_2 \leq 0, \;\; a^\top W_2 = 0\right) \\
\overset{(C)}{=} \quad & P_{\alpha,\epsilon}\left(b^\top W_2 + C_2 \leq 0, \;\; a^\top W_2 = 0\right) \; P_{\alpha,\epsilon}\left(B_2 \geq 0\right) \\
& + P_{\alpha,\epsilon}\left(b^\top W_2 + C_2 \geq 0, \;\; a^\top W_2 = 0\right) \; P_{\alpha,\epsilon}\left(B_2 \leq 0\right) \\
\overset{(B),(C)}{=} \quad & P_{\alpha,\epsilon}\left(b^\top W_2 + C_2 \leq 0, \;\; a^\top W_2 = 0\right) \; \frac{1}{2} \\
& + P_{\alpha,\epsilon}\left(b^\top W_2 + C_2 \geq 0, \;\; a^\top W_2 = 0\right) \; \frac{1}{2} \\
= \quad & P_{\alpha,\epsilon}\left(a^\top W_2 = 0\right) \; \frac{1}{2} \overset{(A)}{=} \epsilon \, \frac{1}{2}.
\end{aligned}
$$

Analogously, it holds $P_{\alpha,\epsilon}(a^\top Z_2 \quad b^\top Z_2 \leq 0, \quad b^\top W_2 = 0) = (1-\epsilon)\frac{1}{2}$ so that $\mathrm{E}_{\alpha,\epsilon}(I\!I\{d_T(\alpha,(Z_1,Z_2)) > 0\} | Z_1 = z_1) = \frac{1}{2}$.□

Hence the simplicial depth for orthogonal regression through the origin is a degenerated U-statistic as this was shown by Müller (2005) for polynomial regression, by Wellmann and Müller (2008a) for multiple regression, and by Wellmann and Müller (2008b) for orthogonal regression for a line with intercept. To obtain the asymptotic distribution of a degenerated U-statistic, the spectral decomposition for the conditional expectation of the kernel function $I\!I\{d_T(\alpha,(Z_1,Z_2)) > 0\} - \frac{1}{2}$ given $Z_1 = z_1, Z_2 = z_2$ is needed. This spectral decomposition consists of infinite eigenfunctions for polynomial regression (Müller 2005, Wellmann et al. 2009), for multiple regression (Wellmann and Müller 2008a), and for orthogonal regression with intercept (Wellmann and Müller 2008b). Hence in these cases, the asymptotic distribution is an infinite sum of random variables basing on squared normal distributed random variables. However, for orthogonal regression through the origin, the spectral decomposition and the asymptotic distribution is much more simple. This is shown in the following theorem. Thereby note, that the assertion of the following theorem holds also for classical linear regression through the origin with very similar arguments.

Theorem 3 *The asymptotic distribution of the simplicial depth for orthogonal regression through the origin is given by*

$$
\mathcal{L}\left(N\left(d_S(\alpha,(Z_1,\ldots,Z_N)) - \frac{1}{2}\right)\right) \overset{N\to\infty}{\longrightarrow} \mathcal{L}\left(\frac{1}{2}(1-W^2)\right),
$$

where W is a random variable with standard normal distribution.

Proof. Hoeffding's theorem (see Lee 1990, p. 79, 80, 90) states for a degenerated U-statistic U with kernel function ψ: If

$$\mathrm{E}(\psi(Z_1,\ldots,Z_{q+1})|Z_1 = z_1, Z_2 = z_2) - \mathrm{E}(\psi(Z_1,\ldots,Z_{q+1})) = \sum_{l=1}^{\infty} \lambda_l \varphi_l(z_1)\varphi_l(z_2)$$

holds almost surely, where $\lambda_l \in I\!R$ and the functions φ_l are L_2-integrable, normalized, and orthogonal, then

$$\mathcal{L}\left(N\ [U(Z_1,\ldots,Z_N) - \mathrm{E}(\psi(Z_1,\ldots,Z_{q+1}))]\right) \stackrel{N\to\infty}{\longrightarrow} \mathcal{L}\left(\binom{q+1}{2}\lambda_l(W_l^2-1)\right),$$

where $W_1, W_2, \ldots$ are independent random variables, each with standard normal distribution.

The conditional expectation of $I\!I\{d_T(\alpha,(Z_1,Z_2)) > 0\}$ given $Z_1 = z_1, Z_2 = z_2$ has the form (compare the proof of Lemma 4)

$$\begin{aligned}
&\mathrm{E}_{\alpha,\epsilon}(I\!I\{d_T(\alpha,(Z_1,Z_2)) > 0\}|Z_1 = z_1, Z_2 = z_1)\\
&= I\!I\{d_T(\alpha,(z_1,z_2)) > 0\}\\
&= I\!I\{A_1(\alpha)^\top z_1\ A_2(\alpha)^\top z_1\ A_1(\alpha)^\top z_2\ A_2(\alpha)^\top z_2 \le 0\}\\
&= I\!I\{A_1(\alpha)^\top z_1\ A_2(\alpha)^\top z_1 \ge 0\}\ I\!I\{A_1(\alpha)^\top z_2\ A_2(\alpha)^\top z_2 \le 0\}\\
&\quad + I\!I\{A_1(\alpha)^\top z_1\ A_2(\alpha)^\top z_1 \le 0\}\ I\!I\{A_1(\alpha)^\top z_2\ A_2(\alpha)^\top z_2 \ge 0\}\\
&= I\!I\{r(z_1) \ge 0\}\ I\!I\{r(z_2) \le 0\} + I\!I\{r(z_1) \le 0\}\ I\!I\{r(z_2) \ge 0\},
\end{aligned}$$

where $r(z_i) = A_1(\alpha)^\top z_i\ A_2(\alpha)^\top z_i$. Hence it holds almost surely

$$\begin{aligned}
\phi(z_1,z_2) &:= \mathrm{E}_{\alpha,\epsilon}(I\!I\{d_T(\alpha,(Z_1,Z_2)) > 0\}|Z_1 = z_1, Z_2 = z_1) - \frac{1}{2}\\
&= I\!I\{r(z_1) > 0\}\ I\!I\{r(z_2) < 0\} + I\!I\{r(z_1) < 0\}\ I\!I\{r(z_2) > 0\} - \frac{1}{2}\\
&= -\frac{1}{2}\ (I\!I\{r(z_1) < 0\} - I\!I\{r(z_1) > 0\})\ (I\!I\{r(z_2) < 0\} - I\!I\{r(z_2) > 0\}).
\end{aligned}$$

Since $\varphi(z_i) = (I\!I\{r(z_i) < 0\} - I\!I\{r(z_i) > 0\})$ is a L_2-integrable, normalized function which is orthogonal to the constant function, it is the only eigenfunction of the operator $\Phi(\varphi)(z_1) = \int \phi(z_1,z_2)\,\varphi(z_2)P(dz_2)$ with nonzero eigenvalue and the corresponding eigenvalue is $\lambda = -\frac{1}{2}$. Hence the spectral decomposition of the conditional expectation is found and Hoeffding's theorem provides the result. □

Since

$$N\ \left(d_S(\alpha,(Z_1,\ldots,Z_N)) - \frac{1}{2}\right) \approx \frac{1}{2} - \frac{1}{2}W^2$$

we have

$$T_\alpha(Z_1,\ldots,Z_N) := 1 - 2\,N\ \left(d_S(\alpha,(Z_1,\ldots,Z_N)) - \frac{1}{2}\right) \approx W^2$$

so that $T_\alpha(Z_1,\ldots,Z_N)$ has approximately a χ^2 distribution with 1 degree of freedom. If a hypothesis $H_0 : \alpha \in A$ with $A \subset [0,\pi)$ is not true, then the maximum simplicial depth within A, i.e. $\max_{\alpha\in A} d_S(\alpha,(Z_1,\ldots,Z_N))$ should be low. This means that $\max_{\alpha\in A} T_\alpha(Z_1,\ldots,Z_N)$ should be high. Hence we have the following asymptotic test, if $\chi_1^2(\gamma)$ denotes the γ-quantile of the χ^2 distribution with 1 degree of freedom.

Corollary 1 *The test given by*

$$\begin{aligned}
&I\!I\{\max_{\alpha\in A} T_\alpha(Z_1,\ldots,Z_N) > \chi_1^2(0.95)\} \\
=\;& I\!I\left\{\max_{\alpha\in A} d_S(\alpha,(Z_1,\ldots,Z_N)) < \frac{1}{2} + \frac{1-\chi_1^2(0.95)}{2\,N}\right\}
\end{aligned}$$

is an asymptotic 0.05-level test for $H_0 : \alpha \in A$ *against* $H_0 : \alpha \notin A$.

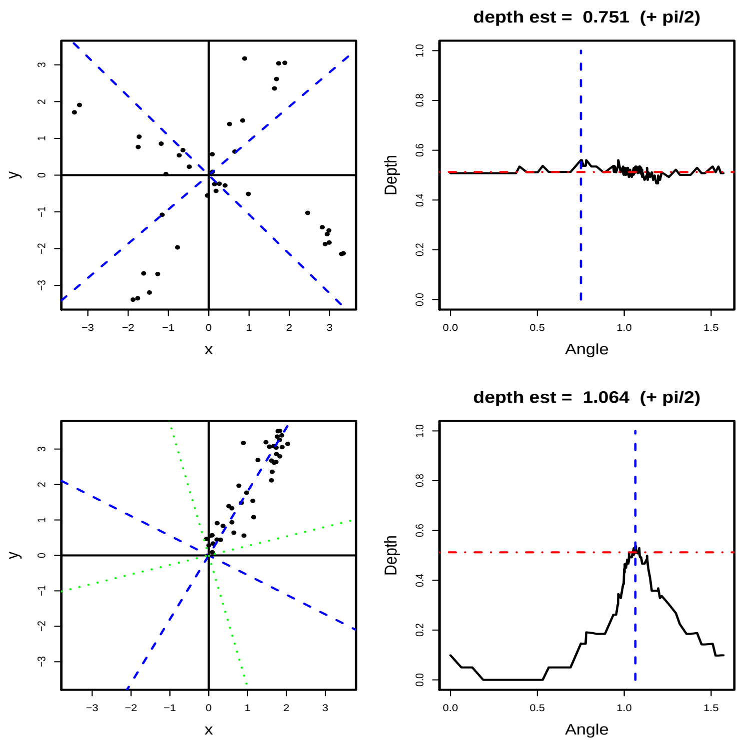

Figure 2: *20 simulated data around a line with* $\pi/3$ *and 20 simulated data around a line with* $\pi/3+\pi/2$*, upper row: original data with depth function, lower row: transformed data with depth function*

4.2 Further checks for orthogonality of a distribution

We can only reject null hypotheses with statistical tests. But we are not able to verify a null hypotheses. Moreover, for distributions which are orthogonal and continuous, the simplicial depth has the same value for all angles according to Theorem 1. Hence data coming from such a distribution will behave similarly. In particular, we cannot distinguish the two angles, which provide the two orthogonal lines around which the data are distributed, from other angles. See the upper row of Figure 2 which shows the simulated data on the left hand side and the corresponding depth function on the right hand side. Thereby, the depth function is only plotted on $[0,\pi/2)$, since the depth functions are the same for α and $\alpha+\pi/2$. The dashed lines in the left hand plot and the right hand plot display the deepest lines and deepest angle, respectively, given by the maximum simplicial depth estimate. For the data of Figure 2, the maximum simplicial depth estimate is 0.751, although the data were generated around $\alpha_0 = \pi/3 = 1.047198$ and $\pi/3+\pi/2$.

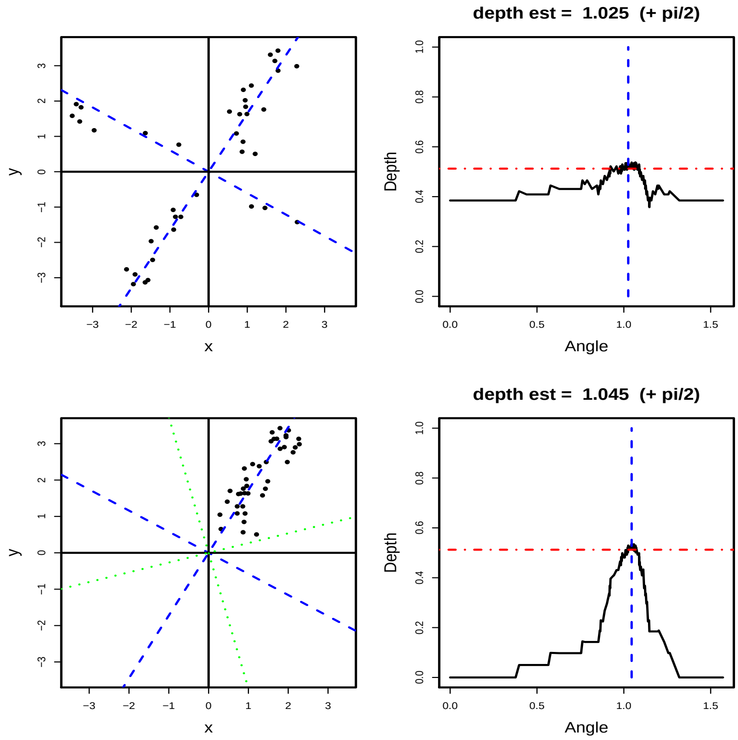

Figure 3: *30 simulated data around a line with $\pi/3$ and 10 simulated data around a line with $\pi/3 + \pi/2$, upper row: original data with depth function, lower row: transformed data with depth function*

But, as soon as the majority of the data is distributed around one line given by α_0 as in the upper row of Figure 3, then there is peak of the depth function close to α_0. This peak is the more pronounced the larger the proportion of the majority of the data is.

This leads to the following idea: If the data are distributed around two orthogonal lines, then they can be mapped to the area around one of this two lines. These mapped data will have then a pronounced peak close to α_0. However, the p-value of testing $H_0 : \alpha \in [\alpha_0 - \eta, \alpha_0 + \eta]$ for these mapped data is the same as before so that a second test makes no sense.

Therefore, we propose the following procedure for checking whether the data are distributed around two orthogonal lines given by α_0 and $\alpha_0 + \pi/2$:

Step 1 Test $H_0 : \alpha \in [\alpha_0 - \eta, \alpha_0 + \eta] \cup [\alpha_0 + \pi/2 - \eta, \alpha_0 + \pi/2 + \eta]$ for some $\eta < \pi/4$ with the test given in Corollary 1. If the hypothesis is not rejected, then continue with Step 2.

Step 2 Rotate the data and the line given by α_0 with respect to $\pi/4 - \alpha_0$ so that the line given by α_0 becomes the line with angles $\pi/4$.

Step 3 Mirror the rotated data at the x- and y-axis so that all data end up in the positive quadrant, i.e. each data point $(x_n, y_n)^\top$ becomes a data point $(\tilde{x}_n, \tilde{y}_n)^\top \in [0, \infty)^2$.

Step 4 Rotate the new data $(\tilde{x}_1, \tilde{y}_1)^\top, \ldots, (\tilde{x}_N, \tilde{y}_N)^\top$ with respect to $\alpha_0 - \pi/4$, i.e. do the reverse rotation of Step 2.

Step 5 Plot the simplicial depth for the rotated new data. If this depth is high for angles inside $[\alpha_0 - \eta, \alpha_0 + \eta] \cup [\alpha_0 + \pi/2 - \eta, \alpha_0 + \pi/2 + \eta]$ and small outside $[\alpha_0 - \eta, \alpha_0 + \eta] \cup [\alpha_0 + \pi/2 - \eta, \alpha_0 + \pi/2 + \eta]$, then the data are distributed around the two lines given by α_0 and $\alpha_0 + \pi/4$.

The lower rows of Figure 2 and Figure 3 show the transformed data with the corresponding depth functions. The dashed lines display again the deepest lines and angles, while the dotted lines indicate the lines given by $\alpha_0 - \pi/4$ and $\alpha_0 + \pi/4$. Figure 2 and Figure 3 show that the transformed data have a much more pronounced peak. The horizontal lines on the right hand sides of these figures indicate the critical values for the 0.05-level tests, i.e. their heights are given by $\frac{1}{2} + \frac{1-\chi_1^2(0.95)}{2\,N}$. In both examples, the p-value for testing $H_0 : \alpha = \frac{\pi}{3}$ is 1.

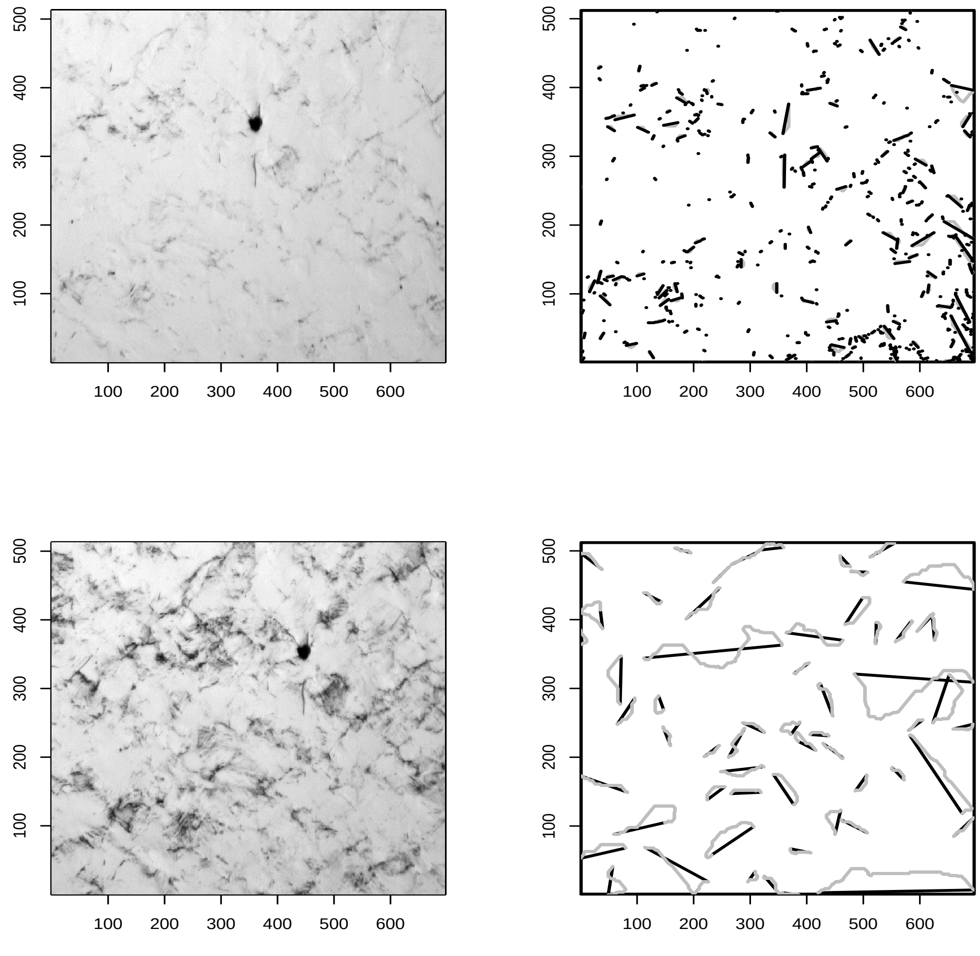

Figure 4: *Image at Time 3 with all detected cracks (upper row) and Image at Time 10 with cracks longer than 30 pixels (lower row)*

5 Application to the analysis of crack orientation

The procedure proposed in Section 4.2 is used for analyzing the orientation of micro cracks. The left hand side of Figure 4 shows two images of a small probe under strain where the strain is given in vertical direction. One image was taken at the early time point 3 and the other image was taken at a later time point 10. Thereby, a time point t means $t \cdot 1000$ load cycles. With the time, more and more micro cracks are visible and the cracks become longer. The right hand side of Figure 4 shows the cracks which were detected by the R package described in Gunkel et al. (2009) using a threshold value of 180. For time point 10, only cracks longer than 30 pixels were plotted. The crack orientations of the plotted cracks were obtained by using the difference of the start and end points of the crack paths. They are shown in the left upper corners of Figure 5 and 6.

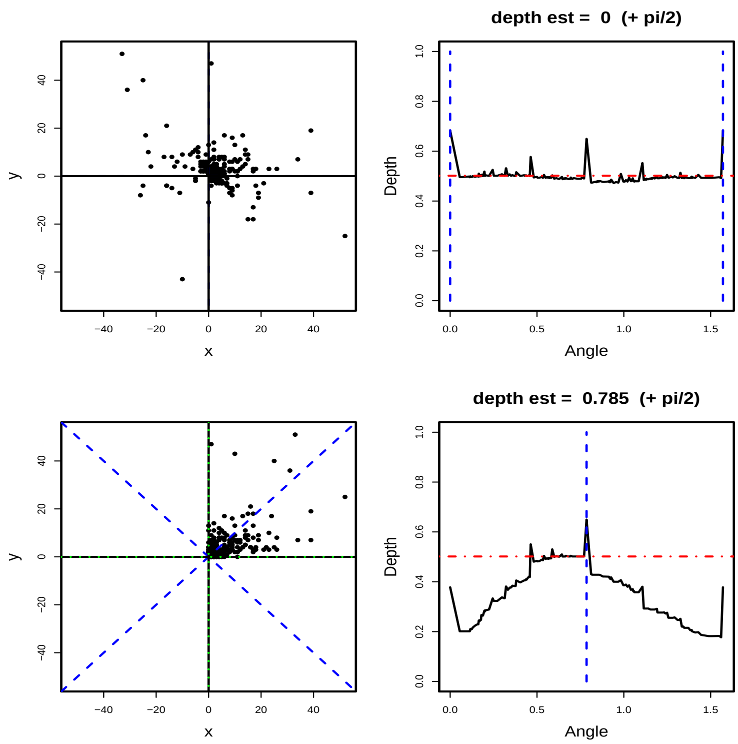

Figure 5: *Crack orientations at Time 3, upper row: original data with depth function, lower row: transformed data with depth function*

To test the hypothesis that small cracks have an orientation which has an angle of 45° and 135° to the strain, the procedure proposed in Section 4.2 is used. We also tested the hypothesis that longer cracks have an orientation perpendicular to the strain. I.e. for small cracks, the hypothesis is $H_0^A : \alpha \in [\pi/4 - \eta, \pi/4 + \eta] \cup [3\pi/4 - \eta, 3\pi/4 + \eta]$, and for longer cracks, we have $H_0^B : \alpha \in [0 - \eta, 0 + \eta]$. We used $\eta = 0.05$.

The p-value for H_0^A is 1 for the cracks at time point 3 and 0.03442 for the longer cracks at time point 10. Hence the longer cracks are not oriented in 45° direction to the strain.

However, for the cracks at time point 3, we can proceed with the Steps 2 to 5. The result is shown in Figure 5: The depth function in the upper row is more or less constant which speaks for an orthogonal distribution. The lower row shows the depth function for the data transformed according to H_0^A. The maximum depth estimate is 0.785 which is exactly $\pi/4$. This supports the hypothesis that small cracks are oriented in 45° direction to the strain. However, the peak at $\pi/4$ is not very pronounced. This indicates that several small cracks have also other orientations.

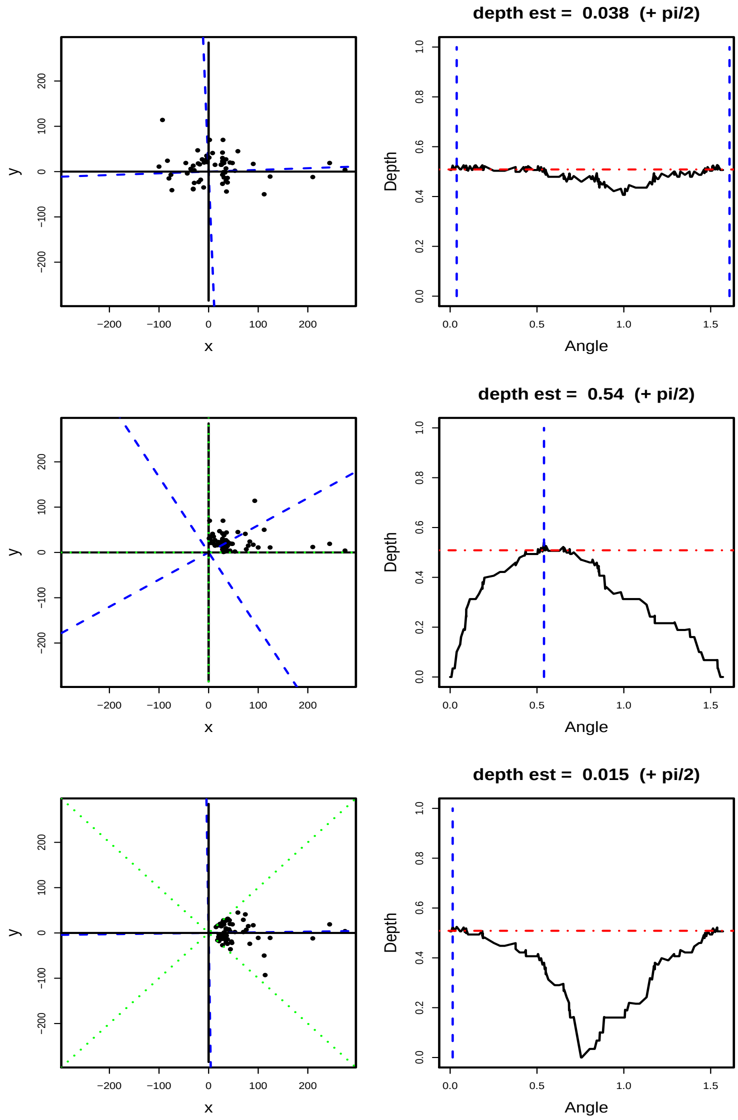

Figure 6: *Crack orientations at Time 10, upper row: original data with depth function, middle row: data transformed according to* H_0^A *with depth function, lower row: data transformed according to* H_0^B *with depth function*

For comparison, Figure 6 shows the depth function for data from time point 10. The middle row of this Figure shows the data transformed according to hypothesis H_0^A. There is a peak, but the peak is not close to $\pi/4$.

The p-value for testing H_0^B is 1. The maximum depth estimate for the original data shown in the upper row of Figure 6 is 0.038 indicating an orientation perpendicular to the strain. However, the peak at the estimate is not very pronounced. The lower row of Figure 6 shows the results when the data are transformed according to Hypothesis H_0^B. The peak close to 0 is now more pronounced supporting H_0^B for longer cracks. But the fact that the peak for the transformed data is more pronounced than for the original data indicates that there exist also many cracks with other orientations.

Acknowledgment. We thank Prof. Dr. Angelika Brückner-Foit, Michael Besel, and Frank Zeismann from the Institute for Materials Engineering of the University of

Kassel for providing us the series of images and for the discussions about crack behavior.

Appendix

Proof of Theorem 2 a)
Assume $\{z_1, \ldots, z_N\} \subset L_1(\alpha)$ or $\{z_1, \ldots, z_N\} \subset L(\alpha)$. Consider $\tilde{z} = (\tilde{z}_1, z_2, \ldots, z_N) \in \mathcal{Z}_1(z)$ with $\tilde{z}_1 = \lambda(\cos(\tilde{\alpha}), \sin(\tilde{\alpha}))^\top$ with $\alpha \neq \tilde{\alpha} \neq \alpha + \pi/2$. Then

$$\sum_{n=2}^{N} \operatorname{res}(\tilde{\alpha}, z_n)^2 = K > 0 \text{ and } \sin(\alpha)\cos(\tilde{\alpha}) - \cos(\alpha)\sin(\tilde{\alpha}) \neq 0.$$

Set $\lambda^2 > K/(\sin(\alpha)\cos(\tilde{\alpha}) - \cos(\alpha)\sin(\tilde{\alpha}))^2$. Then

$$\begin{aligned} \operatorname{res}(\tilde{\alpha}, \tilde{z}_1)^2 + \sum_{n=2}^{N} \operatorname{res}(\tilde{\alpha}, z_n)^2 &= K \\ &< \lambda^2(\sin(\alpha)\cos(\tilde{\alpha}) - \cos(\alpha)\sin(\tilde{\alpha}))^2 = \operatorname{res}(\alpha, \tilde{z}_1)^2 + \sum_{n=2}^{N} \operatorname{res}(\alpha, z_n)^2, \end{aligned}$$

so that $\widehat{\alpha}_{LS}(\tilde{z}) \neq \alpha$. Replacing the squares by absolute values provides also the proof for the L_1 estimator.□

Proof of Theorem 2 b)
(i) At first, assume $\{z_1, \ldots, z_N\} \subset L_1(\alpha)$.
Let be $M < \lceil \frac{N}{2} \rceil$ and $\tilde{z} \in \mathcal{Z}_M(z)$ arbitrary. Then we have $\operatorname{zero}_{\tilde{z}}(\alpha) \geq N - M > \frac{N}{2}$ so that

$$d_T(\alpha, \tilde{z}) = \frac{1}{N} \min\{\operatorname{pos}_{\tilde{z}}(\alpha) + \operatorname{zero}_{\tilde{z}}(\alpha), \operatorname{neg}_{\tilde{z}}(\alpha) + \operatorname{zero}_{\tilde{z}}(\alpha)\} > \frac{N}{2}.$$

If $\tilde{\alpha} \notin \{\alpha, \alpha + \pi/2\}$ then $\operatorname{pos}_{\tilde{z}}(\tilde{\alpha}) \geq \operatorname{zero}_{\tilde{z}}(\alpha) > \frac{N}{2}$ or $\operatorname{neg}_{\tilde{z}}(\tilde{\alpha}) \geq \operatorname{zero}_{\tilde{z}}(\alpha) > \frac{N}{2}$. In the first case, we obtain $\operatorname{neg}_{\tilde{z}}(\tilde{\alpha}) + \operatorname{zero}_{\tilde{z}}(\tilde{\alpha}) < \frac{N}{2}$ and in the second case $\operatorname{pos}_{\tilde{z}}(\tilde{\alpha}) + \operatorname{zero}_{\tilde{z}}(\tilde{\alpha}) < \frac{N}{2}$. Hence, we have $d_T(\tilde{\alpha}, \tilde{z}) < \frac{N}{2}$ in both cases so that $\widehat{\alpha}_T(\tilde{z}) \in \{\alpha, \alpha + \pi/2\}$.

Now let $M \geq \lceil \frac{N}{2} \rceil$. Chose $\tilde{z}$ such that $\tilde{z}_n = z_n$ for $n = 1, \ldots, N - M$ and $\tilde{z}_n = (\cos(\tilde{\alpha}), \sin(\tilde{\alpha}))^\top$ for $n = N - M + 1, \ldots, N$ and $\tilde{\alpha} \notin \{\alpha, \alpha + \pi/2\}$. Then $\tilde{z} \in \mathcal{Z}_M(z)$ and $\operatorname{zero}_{\tilde{z}}(\tilde{\alpha}) \geq M \geq \frac{N}{2}$. With the same arguments as above, we obtain $d_T(\alpha, \tilde{z}) \leq \frac{N}{2} \leq d_T(\tilde{\alpha}, \tilde{z})$ so that there exists an estimator $\widehat{\alpha}_T(\tilde{z}) \notin \{\alpha, \alpha + \pi/2\}$.

(ii) Assume $\{z_1, \ldots, z_N\} \subset L(\alpha)$.
Let be $M < \lceil \frac{N}{3} \rceil$ and $\tilde{z} \in \mathcal{Z}_M(z)$ arbitrary. Then we have $a := \operatorname{zero}_{\tilde{z}}(\alpha) \geq N - M > \frac{2N}{3}$ so that

$$d_T(\alpha, \tilde{z}) > \frac{1}{N} \frac{2N}{3}.$$

Set $b := \operatorname{zero}_{\tilde{z}}(\tilde{\alpha})$ for $\tilde{\alpha} \notin \{\alpha, \alpha + \pi/2\}$. Then it holds $b \leq N - a$ and $\operatorname{neg}_{\tilde{z}}(\tilde{\alpha}) + \operatorname{pos}_{\tilde{z}}(\tilde{\alpha}) = N - b$, so that $\min\{\operatorname{neg}_{\tilde{z}}(\tilde{\alpha}), \operatorname{pos}_{\tilde{z}}(\tilde{\alpha})\} \leq \frac{N-b}{2}$. This implies

$$\begin{aligned} d_T(\tilde{\alpha}, \tilde{z}) &\leq \frac{1}{N}\left(b + \frac{N-b}{2}\right) = \frac{1}{N}\left(\frac{N+b}{2}\right) \leq \frac{1}{N}\left(\frac{2N-a}{2}\right) \\ &< \frac{1}{N}\left(\frac{2N - \frac{2}{3}N}{2}\right) = \frac{1}{N}\frac{2N}{3} < d_T(\alpha, \tilde{z}) \end{aligned}$$

so that $\widehat{\alpha}_T(\tilde{z}) \in \{\alpha, \alpha+\pi/2\}$.

Now let $M \geq \lceil \frac{N+1}{3} \rceil$. Then $\{z_1, \ldots, z_N\} \subset L(\alpha)$, $\tilde{z} \in \mathcal{Z}_M(z)$, and $\tilde{\alpha} \notin \{\alpha, \alpha+\pi\}$ can be chosen so that $\text{zero}_{\tilde{z}}(\alpha) = N - M$, $\text{neg}_{\tilde{z}}(\alpha) = 0$, $\text{zero}_{\tilde{z}}(\tilde{\alpha}) = M$, $\text{neg}_{\tilde{z}}(\tilde{\alpha}) \geq \frac{N-M-1}{2} \leq \text{pos}_{\tilde{z}}(\tilde{\alpha})$. This choice provides

$$\begin{aligned}
d_T(\alpha, \tilde{z}) &= \frac{1}{N}(N-M) \leq \frac{1}{N}\left(N - \frac{N+1}{3}\right) = \frac{1}{N}\left(\frac{2}{3}N - \frac{1}{3}\right), \\
d_T(\tilde{\alpha}, \tilde{z}) &\geq \frac{1}{N}\left(M + \frac{N-M-1}{2}\right) = \frac{1}{N}\left(\frac{N+M-1}{2}\right) \\
&\geq \frac{1}{N}\left(\frac{N + \frac{N+1}{3} - 1}{2}\right) = \frac{1}{N}\left(\frac{2}{3}N - \frac{1}{3}\right)
\end{aligned}$$

so that an estimator $\widehat{\alpha}_T(\tilde{z})$ exists with $\widehat{\alpha}_T(\tilde{z}) \notin \{\alpha, \alpha+\pi/2\}$. □

Proof of Theorem 2 c)

(i) At first, assume again $\{z_1, \ldots, z_N\} \subset L_1(\alpha)$.

Let be $M < \lceil \frac{N}{2} \rceil$ and $\tilde{z} \in \mathcal{Z}_M(z)$ arbitrary. Set again $a := \text{zero}_{\tilde{z}}(\alpha)$ so that $a \geq N - M > \frac{N}{2}$ and $\text{neg}_{\tilde{z}}(\alpha) + \text{pos}_{\tilde{z}}(\alpha) = N - a$. Then we have according to Lemma 2

$$\begin{aligned}
&d_S(\alpha, \tilde{z}) \\
&= \frac{2\,\text{neg}_{\tilde{z}}(\alpha)\text{pos}_{\tilde{z}}(\alpha) + 2(\text{neg}_{\tilde{z}}(\alpha) + \text{pos}_{\tilde{z}}(\alpha))\text{zero}_{\tilde{z}}(\alpha) + \text{zero}_{\tilde{z}}(\alpha)(\text{zero}_{\tilde{z}}(\alpha) - 1)}{N(N-1)} \\
&\geq \frac{1}{N(N-1)}\left(2(N-a)a + a(a-1)\right) = \frac{1}{N(N-1)}\left(2Na - a^2 - a\right). \qquad (5)
\end{aligned}$$

Let be $\tilde{\alpha} \notin \{\alpha, \alpha+\pi/2\}$ arbitrary. Then we have without loss of generality $\text{neg}_{\tilde{z}}(\tilde{\alpha}) = a + b$ with $b \geq 0$, $\text{pos}_{\tilde{z}}(\tilde{\alpha}) = c \geq 0$, $\text{zero}_{\tilde{z}}(\tilde{\alpha}) = N - (a+b+c)$ so that

$$\begin{aligned}
&N(N-1)d_S(\tilde{\alpha}, \tilde{z}) \\
&= 2(a+b)c + 2(a+b+c)(N-(a+b+c)) \\
&\qquad + (N-(a+b+c))(N-(a+b+c)-1) \\
&= 2(a+b)c + 2(a+b+c)N - 2(a+b+c)^2 \\
&\qquad + N^2 - N(a+b+c) - N - N(a+b+c) + (a+b+c)^2 + (a+b+c) \\
&= 2(a+b)c - (a+b+c)^2 + N^2 - N + (a+b+c) \\
&= 2(a+b)c - (a+b)^2 - 2(a+b)c - c^2 + N^2 - N + (a+b+c) \\
&= -(a+b)^2 + (a+b) - c^2 + c + N^2 - N \\
&=: f(b) =: g(c).
\end{aligned}$$

Since $f'(b) = -2(a+b) + 1 \leq 0$ if and only if $b \geq \frac{1}{2} - a$ and $\frac{1}{2} - a < 0$, the function f is decreasing on $[0, \infty)$ so that it attains its maximum at $b = 0$. Since $g'(c) = -2c + 1 \leq 0$ if and only if $c \geq \frac{1}{2}$ and $c \in \mathbb{N}$, the function g is decreasing on $[0, \infty)$ as well so that it attains its maximum at $c = 0$. This implies

$$N(N-1)d_S(\tilde{\alpha}, \tilde{z}) \leq -a^2 + a + N^2 - N. \qquad (6)$$

Since

$$\begin{aligned}
&-a^2 + a + N^2 - N < 2Na - a^2 - a \\
&\iff N^2 - N < 2Na - 2a \iff N^2 - N < a(2N-2) \\
&\iff a > \frac{N(N-1)}{2(N-1)} = \frac{N}{2} \qquad (7)
\end{aligned}$$

and $a > \frac{N}{2}$, we have $d_S(\tilde{\alpha}, \tilde{z}) < d_S(\alpha, \tilde{z})$ for all $\tilde{\alpha} \notin \{\alpha, \alpha + \pi/2\}$ so that $\widehat{\alpha}_S(\tilde{z}) \in \{\alpha, \alpha + \pi/2\}$.

If $M \geq \lceil \frac{N}{2} \rceil$, then we can choose $\tilde{z} \in \mathcal{Z}_M(z)$ and $\tilde{\alpha} \notin \{\alpha, \alpha + \pi/2\}$ such that $a = N - M \leq \frac{N}{2}$ and equality holds in (5) and (6). In particular we have $\mathrm{neg}_{\tilde{z}}(\tilde{\alpha}) = a$, $\mathrm{pos}_{\tilde{z}}(\tilde{\alpha}) = 0$, $\mathrm{zero}_{\tilde{z}}(\tilde{\alpha}) = N - a$. Then $a \leq \frac{N}{2}$ implies according to (7) $d_S(\tilde{\alpha}, \tilde{z}) \geq d_S(\alpha, \tilde{z})$ so that an estimator $\widehat{\alpha}_S(\tilde{z})$ exists with $\widehat{\alpha}_S(\tilde{z}) \notin \{\alpha, \alpha + \pi/2\}$.

(ii) Now, assume $\{z_1, \ldots, z_N\} \subset L(\alpha)$.
Let be $M < \lceil -N + 2 + \sqrt{2N^2 - 6N + 4} \rceil$ and $\tilde{z} \in \mathcal{Z}_M(z)$ arbitrary. Set again $a := \mathrm{zero}_{\tilde{z}}(\alpha)$ so that $a \geq N - M > 2N - 2 - \sqrt{2N^2 - 6N + 4}$, $\mathrm{neg}_{\tilde{z}}(\alpha) + \mathrm{pos}_{\tilde{z}}(\alpha) = N - a$, and inequality (5) holds. Let be $\tilde{\alpha} \notin \{\alpha, \alpha + \pi/2\}$ arbitrary. Then we have $\mathrm{neg}_{\tilde{z}}(\tilde{\alpha}) + \mathrm{pos}_{\tilde{z}}(\tilde{\alpha}) = a + b$ with $b \geq 0$ and $\mathrm{zero}_{\tilde{z}}(\tilde{\alpha}) = N - (a + b)$ so that $\mathrm{neg}_{\tilde{z}}(\tilde{\alpha})\mathrm{pos}_{\tilde{z}}(\tilde{\alpha}) \leq \left(\frac{a+b}{2}\right)^2$ and

$$\begin{aligned}
&N(N-1)d_S(\tilde{\alpha}, \tilde{z}) \\
&\leq 2\left(\frac{a+b}{2}\right)^2 + 2(a+b)(N-(a+b)) + (N-(a+b))(N-(a+b)-1) \\
&= \frac{1}{2}(a+b)^2 + 2(a+b)N - 2(a+b)^2 \\
&\qquad + N^2 - N(a+b) - N - N(a+b) + (a+b)^2 + (a+b) \\
&= -\frac{1}{2}(a+b)^2 + N^2 - N + (a+b) \\
&=: f(b).
\end{aligned}$$

Since $f'(b) = -(a+b) + 1 \leq 0$ if and only if $b \geq 1 - a$ and $1 - a \leq 0$, the function f is decreasing on $[0, \infty)$ so that it attains its maximum at $b = 0$. This implies

$$N(N-1)d_S(\tilde{\alpha}, \tilde{z}) \leq -\frac{1}{2}a^2 + a + N^2 - N. \qquad (8)$$

Since

$$\begin{aligned}
&-\frac{1}{2}a^2 + a + N^2 - N < 2Na - a^2 - a \\
&\iff \frac{1}{2}a^2 + 2a + N^2 - N - 2Na < 0 \\
&\iff a^2 + 4a + 2N^2 - 2N - 4Na < 0 \\
&\iff a^2 - a4(N-1) + 2N^2 - 2N < 0 \\
&\iff a > 2(N-1) - \sqrt{4(N-1)^2 - 2N^2 + 2N} \text{ and} \\
&\qquad a < 2(N-1) + \sqrt{4(N-1)^2 - 2N^2 + 2N} \\
&\iff a > 2(N-1) - \sqrt{4N^2 - 8N + 4 - 2N^2 + 2N} \text{ and} \\
&\qquad a < 2(N-1) + \sqrt{4N^2 - 8N + 4 - 2N^2 + 2N} \\
&\iff a > 2N - 2 - \sqrt{2N^2 - 6N + 4} \text{ and} \\
&\qquad a < 2N - 2 + \sqrt{2N^2 - 6N + 4} \qquad (9)
\end{aligned}$$

and $N \geq a > 2N - 2 - \sqrt{2N^2 - 6N + 4}$, we have $d_S(\tilde{\alpha}, \tilde{z}) < d_S(\alpha, \tilde{z})$ for all $\tilde{\alpha} \notin \{\alpha, \alpha + \pi/2\}$ so that $\widehat{\alpha}_S(\tilde{z}) \in \{\alpha, \alpha + \pi/2\}$.

If $M \geq \lceil -N + 2 + \sqrt{2N^2 - 6N + 5} \rceil$, then we can choose $\{z_1, \ldots, z_N\} \subset L(\alpha)$, $\tilde{z} \in \mathcal{Z}_M(z)$ and $\tilde{\alpha} \notin \{\alpha, \alpha + \pi/2\}$ such that $a = N - M \leq 2N - 2 - \sqrt{2N^2 - 6N + 5}$, equality holds in (5) and $\mathrm{neg}_{\tilde{z}}(\tilde{\alpha}) = \frac{a}{2} = \mathrm{pos}_{\tilde{z}}(\tilde{\alpha})$ if a is even and $\mathrm{neg}_{\tilde{z}}(\tilde{\alpha}) = \frac{a+1}{2}$, $\mathrm{pos}_{\tilde{z}}(\tilde{\alpha}) = \frac{a-1}{2}$, respectively, if a is odd. If a is even, then equality holds in (8) as well. Then $a \leq 2N - 2 - \sqrt{2N^2 - 6N + 4}$ implies according to (9) $d_S(\tilde{\alpha}, \tilde{z}) \geq d_S(\alpha, \tilde{z})$ so that an estimator $\widehat{\alpha}_S(\tilde{z})$ exists with $\widehat{\alpha}_S(\tilde{z}) \notin \{\alpha, \alpha + \pi/2\}$. If a is odd, then we have analogously to (9)

$$
\begin{aligned}
& d_S(\tilde{\alpha}, \tilde{z}) \geq d_S(\alpha, \tilde{z}) \qquad (10)\\
\iff\ & 2\left(\frac{a+1}{2}\right)\left(\frac{a-1}{2}\right) + 2a(N-a) + (N-a)(N-a-1) \geq 2Na - a^2 - a \\
\iff\ & \frac{1}{2}\left(a^2 - 1\right) + 2aN - 2a^2 + N^2 - Na - N - Na + a^2 + a \geq 2Na - a^2 - a \\
\iff\ & -\frac{1}{2}a^2 - \frac{1}{2} + a + N^2 - N \geq 2Na - a^2 - a \\
\iff\ & \frac{1}{2}a^2 + 2a + N^2 - N - 2Na - \frac{1}{2} \geq 0 \\
\iff\ & a^2 - a4(N-1) + 2N^2 - 2N - 1 \geq 0 \\
\iff\ & a \leq 2N - 2 - \sqrt{2N^2 - 6N + 4 + 1} \text{ or} \\
& \qquad a \geq 2N - 2 + \sqrt{2N^2 - 6N + 4 + 1}.
\end{aligned}
$$

Since $a \leq 2N - 2 - \sqrt{2N^2 - 6N + 5}$, the inequality (10) is satisfied so that also in this case an estimator $\widehat{\alpha}_S(\tilde{z})$ exists with $\widehat{\alpha}_S(\tilde{z}) \notin \{\alpha, \alpha + \pi/2\}$.$\square$

References

[1] Bai, Z.-D. and He, X. (1999). Asymptotic distributions of the maximal depth estimators for regression and multivariate location. *Ann. Statist.* **27**, 1616-1637.

[2] Besel M., Brückner-Foit, A. (2008). Surface damage evolution of engineering steel. *To appear in Fatigue and Fracture of Engineering Materials and Structures.*

[3] Besel,M., Brückner-Foit, A., Zeismann, F., Grüning, A., and Mannel, J. (2008). Damage Accumulation of Graded Steel. To appear in Engineering Failure Analysis.

[4] Brückner-Foit, A. and Huang, X. (2008). On the determination of material parameters in crack initiation laws. *To appear in Fatigue and Fracture of Engineering Materials and Structures.*

[5] Brückner-Foit, A., Meyer, S., Möslang, A. (2003). A stochastic simulation model for microcracks in a martensitic steel. *Comp. Mater. Sci.* **26**, 102-110.

[6] Donoho, D.L. and Huber, P.J. (1983). The notion of breakdown point. In *A Festschrift for Erich L. Lehmann* (P.J. Bickel, K.A. Doksum and J.L. Hodges, Jr.,eds.) 157-184. Wadsworth, Belmont, Calif.

[7] Ellis, S.P. and Morgenthaler, S. (1992). Leverage and breakdown in L_1 regression. *J. Amer. Statist. Assoc.* **87**, 143-148.

[8] Fletcher, D.I., Franklin, F.J., Kapoor, A. (2003). Image analysis to reveal crack development using a computer simulation of wear and rolling contact fatigue. *Fatigue Fract. Engng Mater. Struct.* **26**, 957–967.

[9] Fujita, Y., Mitani, Y., Hamamoto, Y. (2006). A method for crack detection on a concrete structure. *Proceedings - International Conference on Pattern Recognition* **3**, 901–904.

[10] Fisher, N.I. (1995). *Statistical Analysis of Circular Data.* Cambridge University Press, New York.

[11] Gunkel, C., Müller, A.C., Stepper, A., Müller, Ch.H. (2009). Automatic crack detection and statistical analysis. *Submitted.*

[12] He, X., Jurečková, J., Koenker, R. and Portnoy, S. (1990). Tail behavior of regression estimators and their breakdown points. *Econometrica* 58, 1195-1214.

[13] Ihara, C.; Tanaka, T. (2000). A stochastic damage accumulation model for crack initiation in high-cycle fatigue. *Fatigue Fract. Engng Mater. Struct.*, **23**, 375-280.

[14] Iyer, S., Sinha, S. K. (2005). A robust approach for automatic detection and segmentation of cracks in underground pipeline images. *Image and Vision Computing* **23**, 921–933.

[15] Lee, A.J. (1990). *U-Statistics. Theory and Practice.* Marcel Dekker, New York.

[16] Liu, R.Y. (1988). On a notion of simplicial depth. *Proc. Nat. Acad. Sci. USA* **85**, 1732-1734.

[17] Liu, R.Y. (1990). On a notion of data depth based on random simplices. *Ann. Statist.* **18**, 405-414.

[18] Ludwig, W., Buffiere, J., Savelli, S., and Cloetens, P. (2003). Study of the intera of short fatigue crack with grain boundaries in a cast Al Alloy using X-ray microtomography. *Acta Mat.*, **51**, 585-598.

[19] Mann, K.A., Gupta, S., Race, A., Miller, M.A., and Cleary, R.J. (2003). Application of circular statistics in the study of crack distribution around cemented femoral components. *Journal of Biomechanics* **36**, 1231-1234.

[20] Mizera, I. (2002). On depth and deep points: A calculus. *Ann. Statist.* **30**, 1681-1736.

[21] Mizera, I. and Müller, Ch.H. (1999). Breakdown points and variation exponents of robust M-estimators in linear models. *Ann. Statist.* **27**, 1164-1177.

[22] Müller, Ch.H. (2005). Depth estimators and tests based on the likelihood principle with application to regression. *Journal of Multivariate Analysis* **95**, 153-181.

[23] Pook, L. (2000). *Linear Elastic Fracture Mechanics for Engineers: Theory and Application.* Southhampton: WIT Press, 2000.

[24] Rousseeuw, P.J. and Hubert, M. (1999). Regression depth (with discussion). *J. Amer. Statist. Assoc.* **94**, 388-433.

[25] Tukey, J.W. (1975). Mathematics and the picturing of data. In *Proc. International Congress of Mathematicians*, Vancouver 1974, **2**, 523- 531.

[26] Van Aelst, S., Rousseeuw, P.J., Hubert, M. and Struyf, A. (2002). The deepest regression method. *J. Multivariate Anal.* **81**, 138-166.

[27] Wellmann, R. (2007). On data depth with application to regression and tests. *Ph.D. thesis.* University of Kassel, Germany.

[28] Wellmann, R., Katina, S. and Müller, Ch.H. (2007). Calculation of simplicial depth estimators for polynomial regression with applications. *Computational Statistics and Data Analysis* **51**, 5025-5040.

[29] Wellmann, R., Harmand, P. and Müller, Ch.H. (2009). Distribution free tests for polynomial regression based on simplicial depth. *J. Multivariate Anal.* **100**, 622-635.

[30] Wellmann, R., and Müller, Ch.H. (2008a). Tests for multiple regression based on simplicial depth. *Submitted*

[31] Wellmann, R., and Müller, Ch.H. (2008b). Depth notions for orthogonal regression. *Submitted*

[32] Zar, J.H. (1999). *Biostatistical Analysis.* Prentice-Hall, Upper Saddle River, NJ.