Tests based on simplicial depth for AR(1) models with explosion

Christoph P. Kustosz*†
Technische Universität Dortmund
Vogelpothsweg 87
D-44221 Dortmund
Germany
E-mail: kustosz@statistik.tu-dortmund.de

Anne Leucht
Technische Universität Braunschweig
Institut für Mathematische Stochastik
Pockelsstr. 14
D-38106 Braunschweig
Germany
E-Mail: a.leucht@tu-bs.de

Christine H. Müller*
Technische Universität Dortmund
Vogelpothsweg 87
D-44221 Dortmund
Germany
E-mail: cmueller@statistik.tu-dortmund.de

Abstract

We propose an outlier robust and distributions-free test for the explosive AR(1) model with intercept based on simplicial depth. In this model, simplicial depth reduces to counting the cases where three residuals have alternating signs. Using this, it is shown that the asymptotic distribution of the test statistic is given by an integrated two-dimensional Gaussian process. Conditions for the consistency of the test are given and the power of the test at finite samples is compared with five alternative tests, using errors with normal distribution, contaminated normal distribution, and Fréchet distribution in a simulation study. The comparisons show that the new test outperforms all other tests in the case of skewed errors and outliers. Finally we apply the proposed methods to crack growth data and compare the results with an ordinary least squares approach. Although we deal with the AR(1) model

---

*Research supported by DFG via the Collaborative Research Centre SFB 823 'Statistical modelling of nonlinear dynamic processes' (Project B5).

†Corresponding Author

with intercept only, the asymptotic results hold for any simplicial depth which reduces to alternating signs of three residuals.

2000 Mathematics Subject Classification. Primary 62M10, 62G10; secondary 62G35, 62F05.
Keywords. Alternating sign, autoregression, distribution-free test, robustness, simplicial depth, consistency.
Short title. Simplicial depth for AR(1) models.

version: February 19, 2015

1. Introduction

Consider the AR(1) model

$$Y_n = \theta_0 + \theta_1 Y_{n-1} + E_n, \quad n = 1, \ldots, N, \tag{1.1}$$

where $E_1, \ldots, E_N$ are i.i.d. errors with $\text{med}(E_n) = 0$, $P(E_n = 0) = 0$, and $Y_0 = y_0$ is the starting value. Moreover, it is assumed that Y_n is almost surely strictly increasing, so that Y_n is an exploding process and $\theta_1 > 1$. An example of such an exploding process is crack growth, where a stochastic version of the Paris-Erdogan equation provides $Y_n = \theta_1 Y_{n-1} + \widetilde{E}_n$, whereby $\widetilde{E}_n$ is nonnegative, see Kustosz and Müller (2014). Setting $\theta_0 = \text{med}(\widetilde{E}_n)$ and $E_n = \widetilde{E}_n - \theta_0$, we obtain model (1.1) for this case. The aim is to test a hypothesis $H_0 : \theta = (\theta_0, \theta_1)^\top \in \Theta_0$, where Θ_0 is a subset of $[0, \infty) \times (1, \infty)$. In particular, the aim is to test hypotheses on the median of the distribution of $\widetilde{E}_n$ as well.

While there is a vast literature on stationary AR(1) models with $|\theta_1| < 1$ and for the unit root case with $\theta_1 = 1$, there exist only few results for the explosive case. Anderson (1959) derived the asymptotic distribution of estimators when the errors E_n are assumed to be independent and normally distributed. Basawa et al. (1989) and Stute and Gründer (1993) used bootstrapping methods to derive the asymptotic distribution of estimators and predictors without assuming the normal distribution. Special maximum likelihood estimators for AR(1) processes with nonnormal errors were treated by Paulaauskas and Rachev (2003). Recently Hwang and Basawa (2005), Hwang et al. (2007) and Hwang (2013) investigated the asymptotic distribution of the least squares estimator of explosive AR(1) processes under some modifications of the process like dependent errors E_n. Further the limit distribution of the Ordinary Least Squares estimator in case of explosive processes was examined by Wang and Yu (2013) and differs from the stationary case depending on the underlying error distribution.

The least squares estimator is known to be sensitive to innovation outliers and additive outliers as discussed in Fox (1972). Outlier robust methods for time series were mainly proposed recently, see e.g. Grossi and Riani (2002), Agostinelli (2003), Fried and Gather (2005), Maronna (2006), Grillenzoni (2009), Gelper et al. (2009). However these methods deal only with estimation and forecasting. An asymptotic distribution was not derived in these papers so that no tests can be used. Only Huggins (1989) proposed a sign test for stochastic processes based on a M-estimator. The robustness of the sign test was studied in Boldin (2011). Moreover Bazarova et al. (2014) derived the asymptotic distribution of trimmed sums for AR(1) processes. However all these approaches base heavily on the stationarity of the process so that they cannot be used for explosive processes.

Here we propose an outlier robust and distribution free test for hypotheses on $\theta = (\theta_0, \theta_1)^\top$ for the explosive AR(1) model given by (1.1) which is based on simplicial depth. Simplicial depth was originally introduced by Liu (1988, 1990) to provide another generalization of the outlier robust median to multivariate data. The direct generalization of the median to multivariate data is the half-space depth proposed by Tukey (1975). In this definition, the depth of a p-dimensional parameter μ within a p-dimensional data set is the minimum relative number of data points lying in a half-space containing the parameter μ. If there are $p+1$ data points then they span a p-dimensional simplex and all points inside of the simplex have the half-space depth of $1/(p+1)$ and all points outside of the simplex have the depth 0. The simplicial depth of Liu defines the depth of a parameter μ as the relative number of simplices spanned by $p+1$ data points which contain the paramter μ, i.e. where the half-space depth of μ with respect to the $p+1$ data points is greater than 0. Replacing the half-space depth by other depth notions leads to a corresponding simplicial depth. For example, Rousseeuw and Hubert (1999) generalized the half-space depth to regression by introducing the concept of nonfit. Thereby the depth of a regression function or respectively the regression parameter θ within a data set is the minimum relative number of data points which must be removed so that the regression function becomes a nonfit. The corresponding simplicial regression depth of a p-dimensional parameter θ is then the relative number of subsets with $p+1$ data points so that θ is not a nonfit with respect to these $p+1$ data points, see Müller (2005). Mizera (2002) proposed a general depth notion by introducing a quality function. Usually the quality function is given by the residuals. Here the residuals of the AR(1) process are used as well.

Simplicial depth has the advantage that it is a U-statistic so that its asymptotic distribution is known in principle. This is usually not the case for the original depth notion. However, the simplicial depth often is a degenerated U-statistic. There are only few cases where this is not the case, see Denecke and Müller (2011, 2012, 2013, 2014). For regression problems, the simplicial depth is a degenerated U-statistic. Deriving the spectral decomposition of the conditional expectation, Müller (2005), Wellmann et al. (2009) and Wellmann and Müller (2010a,b) derived the asymptotic distribution for several regression problems with independent observations. Only the most simple case, namely linear regression through the origin, can be transferred to AR(1) regression with no intercept θ_0. In this case, the asymptotic distribution is given by one χ^2-distributed random variable. This was done in Kustosz and Müller (2014). However, as soon as more than one parameter is unknown, the approaches for an asymptotic distribution developed for regression with independent observations cannot be transferred to autoregression. For example for polynomial regression with independent observations, the asymptotic distribution is given by an infinite sum of χ^2-distributed random variables. Here we show, that the asymptotic distribution for the AR(1) model is given by an integrated two-dimensional Gaussian process. Crucial for this result is that simplicial depth in this model reduces to counting the subsets of three data points where the residuals have alternating signs. Therefore this asymptotic distribution does not hold only for AR(1) models with intercept but also for other models where simplicial depth is given by the number of alternating signs of three residuals. For example it also holds for the nonlinear AR(1) model $Y_n = \theta_1 Y_{n-1}^{\theta_2} + E_n$ under similar assumptions as stated above.

In Section 2, we provide the simplicial depth for the AR(1) model with intercept given by (1.1) and the test statistics based on this simplicial depth for hypotheses about $\theta = (\theta_0, \theta_1)^\top$. To obtain the critical values of the tests, the asymptotic distribution of the simplicial depth is derived in Section 3. In particular, this asymptotic distribution does not depend on the starting value y_0 of the process as it is the case for the asymptotic distribution of the least squares estimator, see e.g. Hwang (2013). In Section 4, we derive the consistency of the tests given in Section 2. Although the test statistic is given by a very simple form, the calculation of it could be lengthy since all subsets of three observations have to be considered. In Section 5 an efficient algorithm for its calculation is shortly described as well as the efficient calculation of quantiles of the asymptotic distribution. Using this, a power simulation of the new test is given and the power is compared with the power of five other tests, where three of them are based on a simplified version of the depth notion given in Kustosz et al. (2015). In Section 6 we apply the proposed test to calculate parameter confidence regions for crack growth data by Maurer and Heeke (2010) and compare the results with regions based on the Ordinary Least Squares estimator.

2. Simplicial depth for the AR(1) model

According to Mizera (2002), we need a quality function to define a depth notion for the AR(1) model (1.1). A natural quality function is the function given by the squared residuals. Set $\theta = (\theta_0, \theta_1)^\top$ and define the residuals by

$$r_n(\theta) = y_n - \theta_0 - \theta_1 y_{n-1}, \ \ n = 1, \dots, N,$$

where y_n is the realization of Y_n for $n = 0, \dots, N$.

Tangential depth of the parameter θ in the sample $y_* = (y_0, \dots, y_N)^\top$ is then (see Mizera, 2002)

$$d_T(\theta, y_*) = \frac{1}{N} \min_{|u|=1} \sharp \left\{ n \in \{1, \dots, N\}; \ u^\top \frac{\partial}{\partial \theta} r_n(\theta)^2 \le 0 \right\}$$

so that it becomes

$$d_T(\theta, y_*) = \frac{1}{N} \min_{|u|=1} \sharp \left\{ n \in \{1, \dots, N\}; \ r_n(\theta) u^\top \begin{pmatrix} 1 \\ y_{n-1} \end{pmatrix} \le 0 \right\}$$

here. To define simplicial depth for autoregression, it is useful to write the sample in pairs, i.e. the sample is given by $z_* = (z_1, \dots, z_N)^\top$ where $z_n = (y_n, y_{n-1})^\top$. Then, simplicial depth of a p-dimensional parameter $\theta \in \mathbb{R}^p$ in the sample z_* is in general (see Müller, 2005)

$$d_S(\theta, z_*) = \frac{1}{\binom{N}{p+1}} \sum_{1 \le n_1 < n_2 < \ldots < n_{p+1} \le N} \mathbb{1}\{d_T(\theta, (z_{n_1}, \dots, z_{n_{p+1}})) > 0\},$$

where $\mathbb{1}\{d_T(\theta, (z_1, \dots, z_{p+1})) > 0\}$ denotes the indicator function $\mathbb{1}_A(z_1, \dots, z_{p+1})$ with $A = \{(\overline{z}_1, \dots, \overline{z}_{p+1})^\top \in \mathbb{R}^{p+1};$ $d_T(\theta, (\overline{z}_1, \dots, \overline{z}_{p+1})) > 0\}$. The simplicial depth is a U-statistic. In the AR(1) model (1.1), it becomes

$$d_S(\theta, z_*) = \frac{1}{\binom{N}{3}} \sum_{1 \le n_1 < n_2 < n_3 \le N} \mathbb{1}\{d_T(\theta, (z_{n_1}, z_{n_2}, z_{n_3})) > 0\}.$$

If the regressors y_{n-1} satisfy $y_{n_1-1} < y_{n_2-1} < y_{n_3-1}$ for $n_1 < n_2 < n_3$, then $d_T(\theta, (z_{n_1}, z_{n_2}, z_{n_3})) > 0$ if and only if the residuals $r_{n_1}, r_{n_2}, r_{n_3}$ have alternating signs or at least one of them is zero (see Kustosz et al., 2015). Since Y_n is almost surely strictly increasing by assumption, we can always assume $y_{n_1-1} < y_{n_2-1} < y_{n_3-1}$ for

$n_1 < n_2 < n_3$ without loss of generality. Moreover, since $P(E_n = 0) = 0$, we can restrict ourselve to residuals with alternating signs so that almost surely

$$\begin{aligned} & d_S(\theta, z_*) \\ & = \frac{1}{\binom{N}{3}} \sum_{1 \le n_1 < n_2 < n_3 \le N} \left(\mathbb{1}\{r_{n_1}(\theta) > 0, r_{n_2}(\theta) < 0, r_{n_3}(\theta) > 0\} \right. \\ & \qquad \left. + \mathbb{1}\{r_{n_1}(\theta) < 0, r_{n_2}(\theta) > 0, r_{n_3}(\theta) < 0\} \right). \end{aligned}$$

If θ is the true parameter, then

$$\begin{aligned} & d_S(\theta, z_*) \\ & = \frac{1}{\binom{N}{3}} \sum_{1 \le n_1 < n_2 < n_3 \le N} \left(\mathbb{1}\{e_{n_1} > 0, e_{n_2} < 0, e_{n_3} > 0\} \right. \\ & \qquad \left. + \mathbb{1}\{e_{n_1} < 0, e_{n_2} > 0, e_{n_3} < 0\} \right), \end{aligned}$$

where e_n is the realization of the error E_n. However, this is not a U-statistic anymore since $\mathbb{1}\{e_{n_1} > 0, e_{n_2} < 0, e_{n_3} > 0\} + \mathbb{1}\{e_{n_1} < 0, e_{n_2} > 0, e_{n_3} < 0\}$ is not a symmetric kernel. Hence, we need the asymptotic distribution for this case which is derived in the next section.

Having the asymptotic distribution of $N(d_S(\theta, Z_*) - 1/2)$ under θ with α-quantile q_α, a simple asymptotic α-level test for the hypothesis $H_0 : \theta \in \Theta_0$, where Θ_0 is a subset of $[0, \infty) \times (1, \infty)$, is (see Müller, 2005):

$$\text{reject } H_0 \text{ if } \sup_{\theta \in \Theta_0} \left(N \left(d_S(\theta, z_*) - \frac{1}{4} \right) \right) \text{ is smaller than } q_\alpha, \tag{2.1}$$

i.e. the depths of all parameters of the hypotheses are too small.

3. Asymptotic distribution of the simplicial depth

Even though the statistic $d_S(\theta, z_*)$ is not an ordinary U-statistic due to the lack of symmetry of the kernel, its asymptotics can be obtained similarly to the derivation of of the limit distribution of first-order degenerate U-statistics. First, we define several functions related to the summands of the statistic $d_S(\theta, z_*)$ by

$$\begin{gathered} h(x, y, z) = \mathbb{1}\{x > 0, y < 0, z > 0\} + \mathbb{1}\{x < 0, y > 0, z < 0\}, \qquad (3.1) \\ h_1(x) = \mathrm{E}h(x, E_2, E_3), \quad h_2(y) = \mathrm{E}h(E_1, y, E_3), \quad h_3(z) = \mathrm{E}h(E_1, E_2, z), \\ h_{1,2}(x, y) = \mathrm{E}h(x, y, E_3), \quad h_{1,3}(x, z) = \mathrm{E}h(x, E_2, z), \\ h_{2,3}(y, z) = \mathrm{E}h(E_1, y, z). \end{gathered}$$

Note that

$$h_1(x) = \frac{1}{4} \left(\mathbb{1}\{x < 0\} + \mathbb{1}\{x > 0\} \right) = \frac{1}{4} = h_2(y) = h_3(z) \quad a.s.$$

which can be compared to first-order degeneracy of U-statistics since straight-forward calculations yield $var(h_{i,j}(E_1, E_2)) = 1/16 > 0$. The latter relations will be important auxiliary results for the derivation of the limit distribution of simplicial depth in the AR(1) setting. Moreover, we will make use of the following approximation.

Lemma 3.1. *Under the afore-mentioned assumptions,*

$$N\left(d_S(\theta, Z_*) - \frac{1}{4}\right)$$
$$= \frac{N}{\binom{N}{3}} \sum_{1\le n_1<n_2<n_3\le N} \Big(h_{1,2}(E_{n_1}, E_{n_2}) + h_{1,3}(E_{n_1}, E_{n_3})$$
$$+h_{2,3}(E_{n_2}, E_{n_3}) - \frac{3}{4}\Big) + \; o_P(1).$$

Proof. We have to show asymptotic negligibility of

$$\frac{N}{\binom{N}{3}} \sum_{1\le n_1<n_2<n_3\le N} \Bigg(h(E_{n_1}, E_{n_2}, E_{n_3}) - \frac{1}{4}$$
$$- \left[h_{1,2}(E_{n_1}, E_{n_2}) + h_{1,3}(E_{n_1}, E_{n_3}) + h_{2,3}(E_{n_2}, E_{n_3}) - \frac{3}{4}\right]\Bigg).$$

As the expectation of this quantity is equal to zero, it remains to show that its variance tends to zero as $N \to \infty$. The latter is given by

$$\frac{N^2}{\binom{N}{3}^2} \sum_{1\le n_1<n_2<n_3\le N} \; \sum_{1\le \bar n_1<\bar n_2<\bar n_3\le N}$$
$$\mathrm{E}\Bigg\{\Bigg(h(E_{n_1}, E_{n_2}, E_{n_3}) + \frac{1}{2} - [h_{1,2}(E_{n_1}, E_{n_2})$$
$$+h_{1,3}(E_{n_1}, E_{n_3}) + h_{2,3}(E_{n_2}, E_{n_3})]\Bigg)$$
$$\times \Bigg(h(E_{\bar n_1}, E_{\bar n_2}, E_{\bar n_3}) + \frac{1}{2} - [h_{1,2}(E_{\bar n_1}, E_{\bar n_2})$$
$$+h_{1,3}(E_{\bar n_1}, E_{\bar n_3}) + h_{2,3}(E_{\bar n_2}, E_{\bar n_3})]\Bigg)\Bigg\}.$$

The number of summands with $n_1 = \bar n_1, n_2 = \bar n_2, n_3 = \bar n_3$ is of order $O(N^3)$ and therefore the corresponding term asymptotically negligible as the factor in front of the sum is of order $O(N^{-4})$. Moreover, note that due to the increasing ordering of the indices it cannot happen that four or more indices coincide. Therefore only three cases remain:

(1) *All indices are different from each other.*
All these summands are equal to zero by the i.i.d. assumptions on the involved random variables.
(2) *Exactly two indices coincide.*
Examplarily, we consider the case $n_1 = \bar n_2$. All remaining pairs can be treated

in a similar manner and are therefore skipped here. We obtain

$$
\begin{aligned}
&\mathrm{E}\Bigg\{\Bigg(h(E_{n_1},E_{n_2},E_{n_3})+\frac{1}{2}-[h_{1,2}(E_{n_1},E_{n_2})\\
&+h_{1,3}(E_{n_1},E_{n_3})+h_{2,3}(E_{n_2},E_{n_3})]\Bigg)\\
&\times\Bigg(h(E_{\bar n_1},E_{n_1},E_{\bar n_3})+\frac{1}{2}-[h_{1,2}(E_{\bar n_1},E_{n_1})\\
&+h_{1,3}(E_{\bar n_1},E_{\bar n_3})+h_{2,3}(E_{n_1},E_{\bar n_3})]\Bigg)\Bigg\}\\
&=\mathrm{E}\Bigg\{\,(h(E_{n_1},E_{n_2},E_{n_3})-[h_{1,2}(E_{n_1},E_{n_2})+h_{1,3}(E_{n_1},E_{n_3})])\\
&\quad\times\Bigg(h(E_{\bar n_1},E_{n_1},E_{\bar n_3})+\frac{1}{2}-[h_{1,2}(E_{\bar n_1},E_{n_1})\\
&+h_{1,3}(E_{\bar n_1},E_{\bar n_3})+h_{2,3}(E_{n_1},E_{\bar n_3})]\Bigg)\Bigg\}\\
&=\mathrm{E}\Bigg\{h_1(E_{n_1})h_2(E_{n_1})+\frac{1}{8}-h_1(E_{n_1})h_2(E_{n_1})-\frac{1}{16}-h_1(E_{n_1})h_2(E_{n_1})\\
&\quad-h_1(E_{n_1})h_2(E_{n_1})-\frac{1}{8}+h_1(E_{n_1})h_2(E_{n_1})+\frac{1}{16}+h_1(E_{n_1})h_2(E_{n_1})\\
&\quad-h_1(E_{n_1})h_2(E_{n_1})-\frac{1}{8}+h_1(E_{n_1})h_2(E_{n_1})+\frac{1}{16}+h_1(E_{n_1})h_2(E_{n_1})\Bigg\}\\
&=0,
\end{aligned}
$$

where the second equality is obtained by conditioning on E_{n_1} and using the tower property of conditional expectation. The last equality follows from $h_1(E_{n_1})\equiv h_2(E_{n_1})\equiv 1/4$ *a.s.*.

(3) *Two pairs of indices coincide.*
Examplarily, we consider the case $n_1=\bar n_2,\ n_2=\bar n_3$. All remaining combinations can again be treated in a similar manner and are therefore skipped

here. We proceed as in the previous case and get

$$
\begin{aligned}
&\mathrm{E}\Bigg\{\Bigg(h(E_{n_1},E_{n_2},E_{n_3})+\frac{1}{2}-[h_{1,2}(E_{n_1},E_{n_2})\\
&+h_{1,3}(E_{n_1},E_{n_3})+h_{2,3}(E_{n_2},E_{n_3})]\Bigg)\\
&\quad\times\Bigg(h(E_{\bar{n}_1},E_{n_1},E_{n_2})+\frac{1}{2}-[h_{1,2}(E_{\bar{n}_1},E_{n_1})\\
&+h_{1,3}(E_{\bar{n}_1},E_{n_2})+h_{2,3}(E_{n_1},E_{n_2})]\Bigg)\Bigg\}\\
&=\mathrm{E}\Bigg\{h_{1,2}(E_{n_1},E_{n_2})h_{2,3}(E_{n_1},E_{n_2})+\frac{1}{8}\\
&\quad-h_1(E_{n_1})h_2(E_{n_1})-h_2(E_{n_2})h_3(E_{n_2})-h_{1,2}(E_{n_1},E_{n_2})h_{2,3}(E_{n_1},E_{n_2})\\
&\quad-h_{1,2}(E_{n_1},E_{n_2})h_{2,3}(E_{n_1},E_{n_2})-\frac{1}{8}\\
&\quad+h_1(E_{n_1})h_2(E_{n_1})+h_2(E_{n_2})h_3(E_{n_2})+h_{1,2}(E_{n_1},E_{n_2})h_{2,3}(E_{n_1},E_{n_2})\\
&\quad-h_1(E_{n_1})h_2(E_{n_1})-\frac{1}{8}+h_1(E_{n_1})h_2(E_{n_1})+\frac{1}{16}+h_1(E_{n_1})h_2(E_{n_1})\\
&\quad-h_2(E_{n_2})h_3(E_{n_2})-\frac{1}{8}+\frac{1}{16}+h_2(E_{n_2})h_3(E_{n_2})+h_2(E_{n_2})h_3(E_{n_2})\Bigg\}\\
&=0.
\end{aligned}
$$

Thus, the variance of the remainder term tends to zero which completes the proof.

□

In order to derive the asymptotics of $N\left(d_S(\theta,Z_*)-1/4\right)$ it remains to investigate

$$
\begin{aligned}
&\frac{N}{\binom{N}{3}}\sum_{1\le n_1<n_2<n_3\le N}\left(h_{1,2}(E_{n_1},E_{n_2})+h_{1,3}(E_{n_1},E_{n_3})+h_{2,3}(E_{n_2},E_{n_3})-\frac{3}{4}\right)\\
&=\frac{N}{2\binom{N}{3}}\Bigg\{\sum_{1\le n_1\ne n_2\le N}(N-\max\{n_1,n_2\})\left[h_{1,2}(E_{n_1},E_{n_2})-\frac{1}{4}\right]\\
&\quad+\sum_{1\le n_1\ne n_3\le N}(\max\{n_1,n_3\}-\min\{n_1,n_3\}-1)\left[h_{1,3}(E_{n_1},E_{n_3})-\frac{1}{4}\right]\\
&\quad+\sum_{1\le n_2\ne n_3\le N}(\min\{n_2,n_3\}-1)\left[h_{2,3}(E_{n_2},E_{n_3})-\frac{1}{4}\right]\Bigg\},
\end{aligned}
$$

where we used that $h_{1,2}$, $h_{1,3}$, and $h_{2,3}$, separately, are symmetric.
Now, invoking the fact that

$$\sum_{1\le n_1\ne n_2\le N} \min\{n_1,n_2\} h_{1,2}(E_{n_1},E_{n_2}) = \sum_{1\le n_2\ne n_3\le N} \min\{n_2,n_3\} h_{2,3}(E_{n_2},E_{n_3})$$

and $N^{-2}\sum_{1\le n_1\ne n_2\le N}(h_{1,2}(E_{n_1},E_{n_2})+h_{1,3}(E_{n_1},E_{n_2})-1/2) \longrightarrow 0$ almost surely with the SLLN, we obtain that the limit distribution of $N\left(d_S(\theta,Z_*)-1/4\right)$ is asymptotically equivalent with

$$\begin{aligned} U_n = {} & \frac{3}{N}\sum_{1\le n_1\ne n_2\le N}\left[h_{1,2}(E_{n_1},E_{n_2})-\frac{1}{4}\right] \\ & + \frac{3}{N^2}\sum_{1\le n_1\ne n_2\le N}(\max\{n_1,n_2\}-\min\{n_1,n_2\}) \\ & \qquad\qquad \cdot\left[h_{1,3}(E_{n_1},E_{n_2})-h_{1,2}(E_{n_1},E_{n_2})\right]. \end{aligned}$$

Now, we invoke a spectral decomposition of the remaining functions in order to separate the the variables E_{n_1} and E_{n_2} in a multiplicative manner. The spectral decomposition of $h_{1,2}-1/4$ is provided in Kustosz and Müller (2014, Proof of Theorem 2). The corresponding eigenfunction is given by $\Phi(x)=\mathbb{1}\{x<0\}-\mathbb{1}\{x>0\}$ and the eigenvalue is $-1/4$. That is $h_{1,2}(x,y)-1/4=-\Phi(x)\Phi(y)/4$. Similarly, we obtain $h_{1,3}(x,y)-h_{1,2}(x,y)=\Phi(x)\Phi(y)/2$. Since $\mathrm{E}[\Phi^2(E_1)]=1$, the SLLN implies

$$\begin{aligned} U_n &= \frac{3}{2N}\sum_{1\le n_1\ne n_2\le N}\left(\frac{|n_1-n_2|}{N}-\frac{1}{2}\right)\Phi(E_{n_1})\Phi(E_{n_2}) \\ &= \frac{3}{2N}\sum_{n_1,n_2=1}^{N}\left(\frac{|n_1-n_2|}{N}-\frac{1}{2}\right)\Phi(E_{n_1})\Phi(E_{n_2})+\frac{3}{4}+o_P(1). \end{aligned}$$

Even though we separated the involved random variables in a multiplicative manner, the application of a CLT to determine the limit of the first summand is not feasible yet because of the weighting factor $|n_1-n_2|/N-1/2$. We solve this problem by a convolution-based representation of the absolute value function on the interval [-1,1],

$$
\begin{aligned}
V_n &:= \frac{3}{2N} \sum_{n_1,n_2=1}^{N} \left(\frac{|n_1-n_2|}{N} - \frac{1}{2} \right) \Phi(E_{n_1})\Phi(E_{n_2}) \\
&= \frac{3}{2N} \sum_{n_1,n_2=1}^{N} \left(\frac{1}{2} - \int_{-\infty}^{\infty} \mathbb{1}_{(-0.5,0.5]}(t) \mathbb{1}_{(-0.5,0.5]} \left(\frac{n_1-n_2}{N} - t \right) dt \right) \\
&\qquad \cdot \Phi(E_{n_1})\Phi(E_{n_2}) \\
&= \frac{3}{2N} \sum_{n_1,n_2=1}^{N} \left(\frac{1}{2} - \int_{-\infty}^{\infty} \mathbb{1}_{(-0.5,0.5]} \left(\frac{n_1}{N} - t \right) \mathbb{1}_{(-0.5,0.5]} \left(\frac{n_2}{N} - t \right) dt \right) \\
&\qquad \cdot \Phi(E_{n_1})\Phi(E_{n_2}) \\
&= \frac{3}{4} \left(\frac{1}{\sqrt{N}} \sum_{n_1=1}^{N} \Phi(E_{n_1}) \right)^2 \\
&\quad - \frac{3}{2} \int_{-2}^{2} \left(\frac{1}{\sqrt{N}} \sum_{n_1=1}^{N} \mathbb{1}_{(-0.5,0.5]} \left(\frac{n_1}{N} - t \right) \Phi(E_{n_1}) \right)^2 dt
\end{aligned}
$$

To sum up

$$
N \left(d_S(\theta, Z_*) - \frac{1}{4} \right) \; = \; V_n + \frac{3}{4} + o_P(1) \tag{3.2}
$$

and the limit distribution of V_n can be deduced by the continuous mapping theorem if the bivariate process

$$
\begin{aligned}
X_N &= (X_{N,1}, X_{N,2})^\top \\
&= \left(\left(\frac{1}{\sqrt{N}} \sum_{n_1=1}^{N} \mathbb{1}_{(-0.5,0.5]} \left(\frac{n_1}{N} - t \right) \Phi(E_{n_1}), \frac{1}{\sqrt{N}} \sum_{n_1=1}^{N} \Phi(E_{n_1}) \right)^\top \right)_{t \in [-2,2]}
\end{aligned}
$$

converges in distribution to some continuous limiting process with respect to the uniform norm.

Lemma 3.2. *Under the assumptions above*

$$
X_N \xrightarrow{d} X,
$$

where $X = (X_1, X_2)^\top$ is a centered Gaussian process on $[-2,2]$ with continuous paths and the covariance structure

$$
\begin{aligned}
&Cov(X(s), X(t)) \qquad (3.3) \\
&= \begin{pmatrix} \int_0^1 \mathbb{1}_{(-0.5,0.5]}(x-s) \mathbb{1}_{(-0.5,0.5]}(x-t) \; dx & \int_0^1 \mathbb{1}_{(-0.5,0.5]}(x-s) \; dx \\ \int_0^1 \mathbb{1}_{(-0.5,0.5]}(x-t) \; dx & 1 \end{pmatrix}.
\end{aligned}
$$

In Figure 1, a simulation of a path of this bivariate process is depicted. The solid line represents the variable $X_1(t)$ which starts in 0 at $t = -0.5$ and returns to 0 at $t = 1.5$. The dashed line is a simulation of $X_2(t)$. This process is a draw from a

$N(0,1)$ distribution and is constant over time. Note, that the two processes meet at $t = 0.5$, due to the underlying covariance structure.

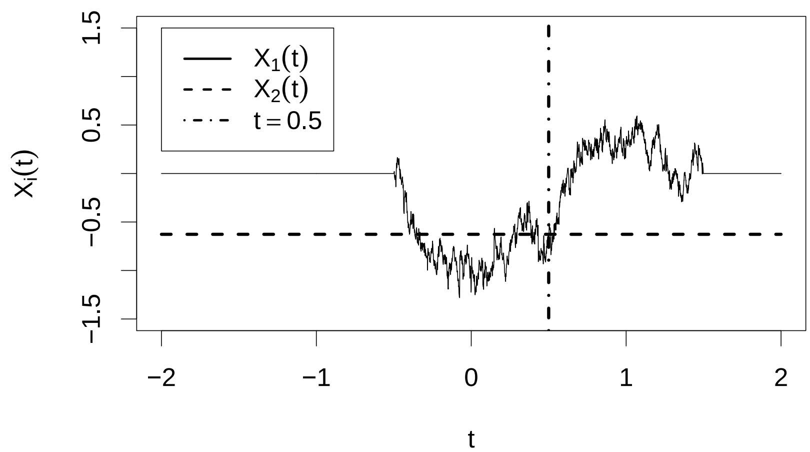


FIGURE 1. Simulation of a path of the limit process

Proof. To prove Lemma 3.2 we apply Theorem 1.5.4 and Problem 1.5.3 in van der Vaart and Wellner (2000) and proceed in several steps.

(1) *Convergence of the finite dimensional distributions.*
We apply the multivariate Lindeberg-Feller CLT to determine the asymptotics of $(X_{N,1}(s_1), \dots, X_{N,1}(s_k), X_{N,2}(t_1), \dots, X_{N,2}(t_l))$ for arbitrary $k, l \in \mathbb{N}$ and $s_1, \dots, s_k, t_1, \dots, t_l \in [-2, 2]$. Obviously, these variables are centered, have finite variance and satisfy the Lindeberg condition. Therefore, it remains to show convergence of the entries of the covariance matrices. For $i = 1, \dots, k$ and $j = 1, \dots, l$, we get

$$\frac{1}{N} \sum_{n_1=1}^{N} \operatorname{var}(\Phi(E_{n_1})) = 1,$$

$$\frac{1}{N} \sum_{n_1=1}^{N} \operatorname{cov}\left(\mathbb{1}_{(-0.5,0.5]}\left(\frac{n_1}{N} - t_j\right) \Phi(E_{n_1}), \Phi(E_{n_1})\right)$$

$$\xrightarrow[N\to\infty]{} \int_0^1 \mathbb{1}_{(-0.5,0.5]}\left(x - t_j\right)\, dx,$$

$$\frac{1}{N} \sum_{n_1=1}^{N} \operatorname{cov}\left(\mathbb{1}_{(-0.5,0.5]}\left(\frac{n_1}{N} - s_i\right) \Phi(E_{n_1}), \mathbb{1}_{(-0.5,0.5]}\left(\frac{n_1}{N} - t_j\right) \Phi(E_{n_1})\right)$$

$$\xrightarrow[N\to\infty]{} \int_0^1 \mathbb{1}_{(-0.5,0.5]}\left(x - s_i\right) \mathbb{1}_{(-0.5,0.5]}\left(x - t_j\right)\, dx,$$

in view of $\mathrm{E}\Phi(E_1) = 0$ and $\mathrm{E}\Phi^2(E_1) = 1$.

(2) *A useful moment bound.*
For $-2 \leq s \leq t \leq 2$, we obtain using $\mathrm{E}\Phi(E_{n_1}) = 0$, $\mathrm{E}\Phi^2(E_{n_1}) = 1 = \mathrm{E}\Phi^4(E_{n_1})$

$$\begin{aligned}
&\mathrm{E}\,\|X_N(s) - X_N(t)\|_1^4 \\
&= \mathrm{E}\left(\frac{1}{\sqrt{N}} \sum_{n_1=1}^{N} \Phi(E_{n_1}) \left[\mathbb{1}_{(-0.5,0.5]}\left(\frac{n_1}{N} - s\right) - \mathbb{1}_{(-0.5,0.5]}\left(\frac{n_1}{N} - t\right)\right]\right)^4 \\
&= \frac{1}{N^2}\left(\sum_{n_1=1}^{N} \left|\mathbb{1}_{(-0.5,0.5]}\left(\frac{n_1}{N} - s\right) - \mathbb{1}_{(-0.5,0.5]}\left(\frac{n_1}{N} - t\right)\right|\right)^2 \\
&+ \frac{1}{N^2} \sum_{n_1=1}^{N} \left[\mathbb{1}_{(-0.5,0.5]}\left(\frac{n_1}{N} - s\right) - \mathbb{1}_{(-0.5,0.5]}\left(\frac{n_1}{N} - t\right)\right]^4 .
\end{aligned}$$

Noting that

$$\left|\mathbb{1}_{(-0.5,0.5]}\left(\frac{n_1}{N} - s\right) - \mathbb{1}_{(-0.5,0.5]}\left(\frac{n_1}{N} - t\right)\right| = 1$$

for at most $2|t-s|N$ indices n_1, we end up with

$$\mathrm{E}\,\|X_N(s) - X_N(t)\|_1^4 \;\leq\; 4\;\left((t-s)^2 \;+\; N^{-1}\;|t-s|\right). \tag{3.4}$$

(3) *Existence and continuity of the limiting process.*
By step 1 and Kolmogorov's existence theorem, there exists a process X with the above-mentioned finite-dimensional distributions. Moreover, from (3.4) we obtain by Fatou's Lemma

$$\mathrm{E}\,\|X(s) - X(t)\|_1^4 \;\leq\; 4\;(t-s)^2.$$

Thus, the theorem of Kolmogorov and Chentsov implies that there exists a continuous modification of X that we also refer to as X in the sequel.

(4) *Tightness.*
By van der Vaart and Wellner (2000, Theorem 1.5.6) it remains to show that for any $\epsilon,\ \eta > 0$ there is a partition $-2 = t_0 < t_1 < \cdots < t_K = 2$ such that

$$\limsup_{N\to\infty} P\left(\sup_{k=1,\ldots,K}\ \sup_{s,t\in[t_{k-1},t_k]} |X_{N,1}(s) - X_{N,1}(t)| > \epsilon\right) \leq \eta$$

since tightness of $(X_{N,2})_N$ is trivial.

For $X_{N,1}$ we obtain

$$
\begin{aligned}
&P\left(\sup_{k=1,\dots,K}\sup_{s,t\in[t_{k-1},t_k]}|X_{N,1}(s)-X_{N,1}(t)|>\epsilon\right)\\
&\leq 2\sum_{k=1}^{K}P\left(\sup_{t\in[t_{k-1},t_k]}|X_{N,1}(t_{k-1})-X_{N,1}(t)|>\frac{\epsilon}{2}\right)\\
&= 2\sum_{k=1}^{K}P\Bigg(\sup_{t\in[t_{k-1},t_k]}\Bigg|\frac{1}{\sqrt{N}}\sum_{n_1=1}^{N}\phi(E_{n_1})\Bigg\{\mathbb{1}_{(-0.5+t_{k-1},-0.5+t]}\left(\frac{n_1}{N}\right)\\
&\qquad -\mathbb{1}_{(0.5+t_{k-1},0.5+t]}\left(\frac{n_1}{N}\right)\Bigg\}\Bigg|>\frac{\epsilon}{2}\Bigg)
\end{aligned}
$$

For symmetry reasons we only consider

$$
\frac{1}{\sqrt{N}}\sum_{n_1=1}^{N}\phi(E_{n_1})\mathbb{1}_{(-0.5+t_{k-1},-0.5+t]}\left(\frac{n_1}{N}\right)=Q_N(t)-Q_N(t_{k-1})
$$

with $Q_N(t):=\frac{1}{\sqrt{N}}\sum_{n_1=1}^{N}\phi(E_{n_1})\mathbb{1}_{(-2,-0.5+t]}\left(\frac{n_1}{N}\right)$. The process Q_N has càdlàg paths and independent increments. Moreover, note that it follows from the proof of (2) and Markov's inequality that for some $\alpha,\beta>0$, $P(|Q_N(t)-Q_N(s)|\leq\delta)\geq\beta$ whenever $|t-s|\leq\alpha$ and $N\geq N_0$. Therefore we can proceed as in the proof of Theorem V.19 in Pollard (1984) to obtain

$$
\limsup_{N\to\infty}\sum_{k=1}^{K}P\left(\sup_{t\in[t_{k-1},t_k]}|Q_N(t)-Q_N(t_{k-1})|>\frac{\epsilon}{4}\right)<\frac{\eta}{4}
$$

for a sufficiently fine equidistant grid.

□

To sum up, we get the following theorem on the asymptotics of simplicial depth by the continuous mapping theorem.

Theorem 3.1. *Under the assumptions above*

$$
N\left(d_S(\theta,Z_*)-\frac{1}{4}\right)\xrightarrow{d}\frac{3}{4}+\frac{3}{4}X_2^2(0)-\frac{3}{2}\int_{-2}^{2}X_1^2(t)dt.
$$

Note, that the asymptotic distribution of the simplicial depth is not restricted to the AR(1) model considered in this paper. It holds for all cases where depth of a two dimensional parameter at three data points is given by alternating signs of the three residuals. This holds in several other models as shown in Kustosz et al. (2015).

4. Consistency of the test

Here we show consistency of the test given by (2.1) for hypotheses $H_0 : \theta = \theta^0$ and $H_0 : \theta_1 \geq \theta_1^0$ at all relevant alternatives $\theta^* = (\theta_0^*, \theta_1^*)^\top$ by using a large upper bound of the test statistic. Thereby, a test is called consistent at θ^* if the power of the test at θ^* is converging to one for growing sample size. Hence we have to proof here

$$\lim_{N\to\infty} P_{\theta^*}\left(\sup_{\theta\in\Theta_0}\left(N\left(d_S(\theta, Z_*) - \frac{1}{4}\right)\right) < q_\alpha\right) = 1,$$

where q_α is the α-quantile of the asymptotic distribution of $N\left(d_S(\theta, Z_*) - 1/4\right)$ given by Theorem 3.1.

Lemma 4.1. *If there exists $N_0 \in \mathbb{N}$, $\delta \in (0, \frac{1}{4})$, and a bounded function $H : \mathbb{R}^3 \to \mathbb{R}$ with*

$$\begin{aligned} &H(r_{n_1}(\theta^*), r_{n_2}(\theta^*), r_{n_3}(\theta^*)) \\ &\geq \sup_{\theta\in\Theta_0}\Big(\mathbb{1}\{r_{n_1}(\theta) > 0, r_{n_2}(\theta) < 0, r_{n_3}(\theta) > 0\} \\ &\qquad + \mathbb{1}\{r_{n_1}(\theta) < 0, r_{n_2}(\theta) > 0, r_{n_3}(\theta) < 0\}\Big) \end{aligned} \tag{4.1}$$

for all $n_1, n_2, n_3 > N_0$ and

$$E_{\theta^*}\Big(H(r_{n_1}(\theta^*), r_{n_2}(\theta^*), r_{n_3}(\theta^*))\Big) < \frac{1}{4} - \delta, \tag{4.2}$$

then the test given by (2.1) is consistent at θ^*.

Proof. Set $M_0 = \{(n_1, n_2, n_3); n_3 > n_2 > n_1 > N_0\}$, then

$$\begin{aligned} &\sup_{\theta\in\Theta_0}\left(N\left(d_S(\theta, Z_*) - \frac{1}{4}\right)\right) \\ \leq &\frac{N}{\binom{N}{3}}\Bigg(\sum_{1\leq n_1<n_2<n_3\leq N}\sup_{\theta\in\Theta_0}(\mathbb{1}\{r_{n_1}(\theta) > 0, r_{n_2}(\theta) < 0, r_{n_3}(\theta) > 0\} \\ &\qquad + \mathbb{1}\{r_{n_1}(\theta) < 0, r_{n_2}(\theta) > 0, r_{n_3}(\theta) < 0\}) - \frac{1}{4}\Bigg) \\ \leq &\frac{N}{\binom{N}{3}}\Bigg(\left[\binom{N}{3} - \binom{N-N_0}{3}\right] \\ &+ \sum_{(n_1,n_2,n_3)\in M_0}\left(H(r_{n_1}(\theta^*), r_{n_2}(\theta^*), r_{n_3}(\theta^*)) - \frac{1}{4}\right)\Bigg) =: T. \end{aligned}$$

Hence, we will work with T. To apply Chebyshev's inequality, we need upper bounds for the expectation and the variance of T. Since (4.2) holds on M_0 and the indicators are bounded on the remaining indices, we get

$$\mathrm{E}_{\theta^*}(T) \le \frac{N}{\binom{N}{3}} \left(\left[\binom{N}{3} - \binom{N-N_0}{3} + \binom{N-N_0}{3}(-\delta) \right] \right).$$

Then there exists $N_1 > N_0$ such that

$$\frac{1}{\binom{N}{3}} \binom{N-N_0}{3} \ge 1-\delta \text{ and } \frac{1}{\binom{N}{3}} \left[\binom{N}{3} - \binom{N-N_0}{3} \right] < \frac{\delta}{2}$$

for all $N \ge N_1$ implying

$$E_{\theta^*}(T) \le N \left(\frac{\delta}{2} + (1-\delta)(-\delta) \right) = -N\,\delta_0$$

with $\delta_0 = \frac{1}{2}\delta - \delta^2 > 0$ since $\delta < \frac{1}{4}$. Setting $H_{n_1,n_2,n_3} = H(r_{n_1}(\theta^*), r_{n_2}(\theta^*), r_{n_3}(\theta^*)) - E_{\theta^*}\left[H(r_{n_1}(\theta^*), r_{n_2}(\theta^*), r_{n_3}(\theta^*))\right]$, we obtain for the variance

$$\begin{aligned} \mathrm{var}_{\theta^*}(T) &= \frac{N^2}{\binom{N}{3}^2} \mathrm{E}_{\theta^*} \left\{ \left(\sum_{(n_1,n_2,n_3)\in M_0} H_{n_1,n_2,n_3} \right)^2 \right\} \\ &= \frac{N^2}{\binom{N}{3}^2} \sum_{(n_1,n_2,n_3)\in M_0} \sum_{(\overline{n}_1,\overline{n}_2,\overline{n}_3)\in M_0} \mathrm{E}_{\theta^*} \left(H_{n_1,n_2,n_3}\, H_{\overline{n}_1,\overline{n}_2,\overline{n}_3} \right). \end{aligned}$$

For $\binom{N-N_0}{3}\binom{N-N_0-3}{3}$ combinations, all n_1, n_2, n_3 are different from $\overline{n}_1, \overline{n}_2, \overline{n}_3$, so that the independence of the residuals $r_n(\theta^*)$ implies $\mathrm{E}_{\theta^*}(H_{n_1,n_2,n_3}\, H_{\overline{n}_1,\overline{n}_2,\overline{n}_3}) = 0$ for these cases. In all other cases, $\mathrm{E}_{\theta^*}(H_{n_1,n_2,n_3}\, H_{\overline{n}_1,\overline{n}_2,\overline{n}_3})$ is bounded by some b^2 so that

$$\mathrm{var}_{\theta^*}(T) \le \frac{N^2}{\binom{N}{3}^2} \left(\binom{N-N_0}{3}^2 - \binom{N-N_0}{3}\binom{N-N_0-3}{3} \right)\ b^2.$$

Hence for all $\epsilon > 0$, there exists $N_2 \ge N_1 > N_0$ such that $\delta_0 + \frac{q_\alpha}{N_2} > 0$ and

$$\mathrm{var}_{\theta^*}(T) \le \epsilon N^2 \left(\delta_0 + \frac{q_\alpha}{N_2} \right)^2$$

for all $N \geq N_2$. Finally, Chebyshev's inequality provides for all $N \geq N_2$ using $\frac{q_\alpha}{N} \geq \frac{q_\alpha}{N_2}$

$$
\begin{aligned}
& P_{\theta^*}\left(\sup_{\theta\in\Theta_0}\left(N\left(d_S(\theta, Z_*) - \frac{1}{4}\right)\right) \geq q_\alpha\right) \\
\leq \;& P_{\theta^*}\left(T \geq q_\alpha\right) \leq P_{\theta^*}\left(|T - E_{\theta^*}(T)| \geq q_\alpha - E_{\theta^*}(T)\right) \\
\leq \;& P_{\theta^*}\left(|T - E_{\theta^*}(T)| \geq q_\alpha + N\,\delta_0\right) \leq \frac{\epsilon N^2\left(\delta_0 + \frac{q_\alpha}{N_2}\right)^2}{N^2\;\left(\delta_0 + \frac{q_\alpha}{N}\right)^2} \\
\leq \;& \frac{\epsilon\left(\delta_0 + \frac{q_\alpha}{N_2}\right)^2}{\left(\delta_0 + \frac{q_\alpha}{N_2}\right)^2} = \epsilon,
\end{aligned}
$$

since $q_\alpha < 0$ (see Section 5). $\square$

The following Lemma is easy to see by induction.

Lemma 4.2. *If $Y_0 = y_0$ and the errors satisfy $E_n \geq y_0 - \theta_0 - \theta_1 y_0 + c$ for all n for some $y_0 > 0$ and $c > 0$, then Y_n is strictly increasing with*

$$
Y_n \geq \left(\sum_{k=0}^{n-1} \theta_1^k\right) c + y_0.
$$

If, for example, the errors have a shifted Fréchet distribution as used in the simulations in Section 5, then $E_n \geq y_0 - \tau y_0 + c$ and $\mathrm{med}(E_n) = 0$ is satisfied for some $y_0 > 0$, $c > 0$, and $\tau > 1$.

Theorem 4.1. *If the errors satisfy $E_n \geq y_0 - \tau y_0 + c$ for all n for some $y_0 > 0$, $c > 0$, $\tau > 1$, $Y_0 = y_0$, $\Theta = [0, \infty) \times [\tau, \infty)$, $\theta_0^0 \geq 0$, $\theta_1^0 > \tau$ and*

$$
\Theta_0 = \{(\theta_0, \theta_1)^\top \in \Theta;\ \theta_1 \geq \theta_1^0\} \quad \text{or} \quad \Theta_0 = \{(\theta_0^0, \theta_1^0)^\top\},
$$

then the test given by (2.1) is consistent at all $\theta^* \in \Theta \setminus \Theta_0$.

Proof. Set $\Theta_0^h = \{(\theta_0, \theta_1)^\top \in \Theta;\ \theta_1 \geq \theta_1^0\}$ for the half-sided null hypotheses and $\Theta_0^p = \{(\theta_0^0, \theta_1^0)^\top\}$ for the point null hypothesis and use h defined in (3.1).

If $\theta^* \in \Theta \setminus \Theta_0^h$ then

$$\begin{aligned}
&\sup_{\theta \in \Theta_0^h} h(r_{n_1}(\theta), r_{n_2}(\theta), r_{n_3}(\theta)) \\
&\quad\leq \sup_{\theta \in \Theta_0^h} \Big(\mathbb{1}\{r_{n_1}(\theta) > 0\} + \mathbb{1}\{r_{n_2}(\theta) > 0\} \Big) \\
&\quad= \sup_{\theta \in \Theta_0^h} \Big(\mathbb{1}\{r_{n_1}(\theta^*) > \theta_0 - \theta_0^* + (\theta_1 - \theta_1^*) Y_{n_1 - 1}\} \\
&\qquad\qquad + \mathbb{1}\{r_{n_2}(\theta^*) > \theta_0 - \theta_0^* + (\theta_1 - \theta_1^*) Y_{n_2 - 1}\} \Big) \\
&\quad\leq \mathbb{1}\{r_{n_1}(\theta^*) > -\theta_0^* + (\theta_1^0 - \theta_1^*) Y_{n_1 - 1}\} \\
&\qquad + \mathbb{1}\{r_{n_2}(\theta^*) > -\theta_0^* + (\theta_1^0 - \theta_1^*) Y_{n_2 - 1}\}
\end{aligned}$$

since $\theta_0 \geq 0$ and $\theta_1 \geq \theta_1^0$ for $\theta = (\theta_0, \theta_1)^\top \in \Theta_0^h$. According to Lemma 4.2, for all $\gamma > 0$ there exist N_0 such that

$$Y_n \geq \left(\sum_{k=0}^{n-1} \theta_1^k \right) c + y_0 \geq \gamma$$

for all $N \geq N_0$. In particular γ can be chosen such that $-\theta_0^* + (\theta_1^0 - \theta_1^*)\gamma > k$ and $P_{\theta^*}(r_n(\theta^*) > k) < \frac{1}{2}(\frac{1}{4} - \delta)$ with $\delta \in (0, \frac{1}{4})$ since $\theta_1^0 > \theta_1^*$. Setting

$$H(r_{n_1}(\theta^*), r_{n_2}(\theta^*), r_{n_3}(\theta^*)) = \mathbb{1}\{r_{n_1}(\theta^*) > k\} + \mathbb{1}\{r_{n_2}(\theta^*) > k\},$$

Conditions (4.1) and (4.2) are satisfied, so that consistency holds for all $\theta^* \in \Theta \setminus \Theta_0^h$.

If $\theta^* \in \Theta \setminus \Theta_0^p$ and $\theta_1^0 > \theta_1^*$, then consistency at θ^* follows as above. If $\theta_1^0 < \theta_1^*$, then there exists $\gamma > 0$ with $\theta_0^0 - \theta_0^* + (\theta_1^0 - \theta_1^*)\gamma < -k$ and $P_{\theta^*}(r_n(\theta^*) < -k) < \frac{1}{2}(\frac{1}{4} - \delta)$ so that

$$H(r_{n_1}(\theta^*), r_{n_2}(\theta^*), r_{n_3}(\theta^*)) = \mathbb{1}\{r_{n_1}(\theta^*) < -k\} + \mathbb{1}\{r_{n_2}(\theta^*) < -k\}$$

satisfies Conditions (4.1) and (4.2). If $\theta_1^0 = \theta_1^*$, then $k := \theta_0^0 - \theta_0^* \neq 0$ and

$$\begin{aligned}
&h(r_{n_1}(\theta), r_{n_2}(\theta), r_{n_3}(\theta)) \\
&\quad= \mathbb{1}\{r_{n_1}(\theta^*) > k, r_{n_2}(\theta^*) < k, r_{n_3}(\theta^*) > k\} \\
&\qquad + \mathbb{1}\{r_{n_1}(\theta^*) < k, r_{n_2}(\theta^*) > k, r_{n_3}(\theta^*) < k\} \\
&\quad=: H(r_{n_1}(\theta^*), r_{n_2}(\theta^*), r_{n_3}(\theta^*)).
\end{aligned}$$

Since $p(1-p)p + (1-p)p(1-p) = p(1-p) < \frac{1}{4}$ for all $p \neq \frac{1}{2}$ and $p = P_{\theta^*}(r_n(\theta^*) < k) \neq \frac{1}{2}$, Condition (4.2) is also satisfied. Hence consistency holds for all $\theta^* \in \Theta \setminus \Theta_0^p$ as well.

$\square$

5. Simulation of the power of the test

Since confidence sets can be constructed from point hypotheses, we consider the most important hypothesis $H_0 : \theta = \theta^0$. To simulate the power of the test based on the simplicial depth for this hypothesis, approximate quantiles of the asymptotic distribution given by Theorem 3.1 were determined, and an efficient algorithm for calculating the test statistic was developed.

5.1. **Quantiles of the asymptotic distribution.**

To calculate approximate quantiles, we use an equidistant partition of $[-2, 2]$ defined by $t_0 = -2$, $t_{i+1} = t_i + h$ and $t_K = 2$ with $h = 0.001$ and generate 200000 repetitions of the process $x(t_0), x(t_1), \ldots, x(t_K)$. According to Lemma 3.2, $(X(t_0)^\top, X(t_1)^\top, \ldots, X(t_K)^\top)^\top$ has a multivariate normal distribution with a degenerated $2K \times 2K$ covariance matrix. To avoid degeneracy in some cases and to reduce the computational costs, we simulate $(x_1(t_1), x_1(t_2), \ldots, x_1(t_K))^\top$ from the conditional distribution of $(X_1(t_1), X_1(t_2), \ldots, X_1(t_K))^\top$ given $(X_2(t_1), X_2(t_2), \ldots, X_2(t_K))^\top = x$,
where $x = (x_2(t_0), ..., x_2(t_0))^\top$ results from one realization of the standard normal distribution. Then an approximation of the integrated limit process is given by

$$W_K = \frac{3}{4} + \frac{3}{4} x_2(t_0)^2 - \frac{3}{2} \sum_{k=0}^{K-1} \frac{1}{2} (x_1(t_{k+1})^2 + x_1(t_k)^2)(t_{k+1} - t_k).$$

Taking the empirical quantiles of W_K, yields the approximate quantiles of the asymptotic distribution presented in Table 1. A complete quantile plot is depicted in Figure 2.

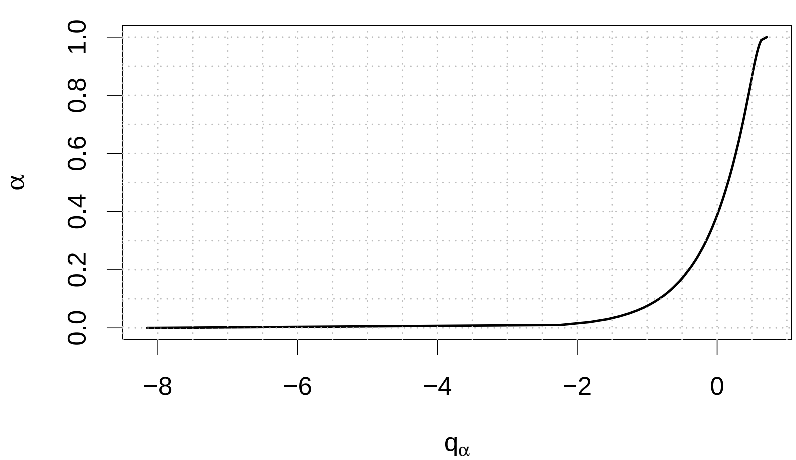


Figure 2. Quantiles of the approximate distribution of $N(d_S(\theta, Z_*) - 1/4)$

5.2. **Calculation of the test statistic.**
To evaluate the full depth statistic $d_S(\theta, z_*)$, we apply parallel computation and matrix based operations which reduce the computational costs to order N^2 compared to the naive calculation with order N^3. Therefore we use tilted triangular matrices and Hankel matrices to calculate depth with one residual fixed in an efficient way. In particular the new algorithm allows a parallel computation of the remaining loop. All computations are performed in R (see R Core Team, 2013) with the packages for multi-core computation (see Yu, 2002; Tierney et al., 2013; Venables and Ripley, 2002) and a package for fast matrix calculations (see Novomestky, 2012).

5.3. **Simulation study.** We compare the test based on simplicial depth for $H_0 : \theta = \theta^0$ using $\theta^0 = (\theta_0^0, \theta_1^0)^\top = (0.2, 1.01)^\top$ with five other tests. Three of them are simplified depth statistics using partial evaluation of the full simplicial depth and were proposed in Kustosz et al. (2015). This simplifies the calculation and allows a simple derivation of limit distributions which is the normal distribution. The simplified statistics are defined by

$$
\begin{aligned}
d_S^1(\theta, z_*) = \frac{1}{\lfloor \frac{N}{3} \rfloor} \sum_{n=1}^{\lfloor \frac{N}{3} \rfloor} & \mathbb{1}\{r_{3n-2}(\theta) > 0, r_{3n-1}(\theta) < 0, r_{3n}(\theta) > 0\} \\
& + \mathbb{1}\{r_{3n-2}(\theta) < 0, r_{3n-1}(\theta) > 0, r_{3n}(\theta) < 0\}, \\
d_S^2(\theta, z_*) = \frac{1}{\lfloor \frac{N-1}{2} \rfloor} \sum_{n=1}^{\lfloor \frac{N-1}{2} \rfloor} & \mathbb{1}\{r_n(\theta) > 0, r_{\lfloor \frac{N+1}{2} \rfloor}(\theta) < 0, r_{N-n+1}(\theta) > 0\} \\
& + \mathbb{1}\{r_n(\theta) < 0, r_{\lfloor \frac{N+1}{2} \rfloor}(\theta) > 0, r_{N-n+1}(\theta) < 0\}, \\
d_S^3(\theta, z_*) = \frac{1}{N-2} \sum_{n=1}^{N-2} & \mathbb{1}\{r_n(\theta) > 0, r_{n+1}(\theta) < 0, r_{n+2}(\theta) > 0\} \\
& + \mathbb{1}\{r_n(\theta) < 0, r_{n+1}(\theta) > 0, r_{n+2}(\theta) < 0\}.
\end{aligned}
$$

The other two tests are the simple sign test defined by

$$
T_s(\theta, z_*) = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} sign(r_n(\theta)),
$$

which is described in Huggins (1989) for stochastic processes, and an OLS test based on the limit distributions derived in Wang and Yu (2013). The sign test uses the exact error distribution of the residual signs when the $med(E_n) = 0$ assumption holds. For the OLS test we define critical values based on the asymptotic independence of the marginal estimators. Since we do not know the exact error distribution, we apply the OLS test assuming normal errors in all examples. We evaluate the power of the six tests on a grid defined by $\theta_0 \in [-0.15, 0.52]$ with mesh size 0.01 and $\theta_1 \in [1, 1.021]$ with mesh size 0.0003. For each grid point, we simulate $R = 100$ processes of length $N = 100$ with the underlying parameter combination and with starting value $y_0 = 15$.

We use three different distributions for the errors: a normal distribution with mean zero and variance 0.01, a contaminated normal distribution given by $A_n + P_n \cdot B_n$, whereby $A_n \sim \mathcal{N}(0, 0.1), B_n \sim \mathcal{N}(5, 1)$ and $P_n \sim Pois(5/100)$ are independent random variables for each n, and a Fréchet distribution defined by the density

$$f_{\alpha,\beta,\gamma}(x) = \frac{\gamma}{\alpha}\left(\frac{x-\beta}{\alpha}\right)^{-1-\gamma} exp\left(-\left(\frac{x-\beta}{\alpha}\right)^{-\gamma}\right)$$

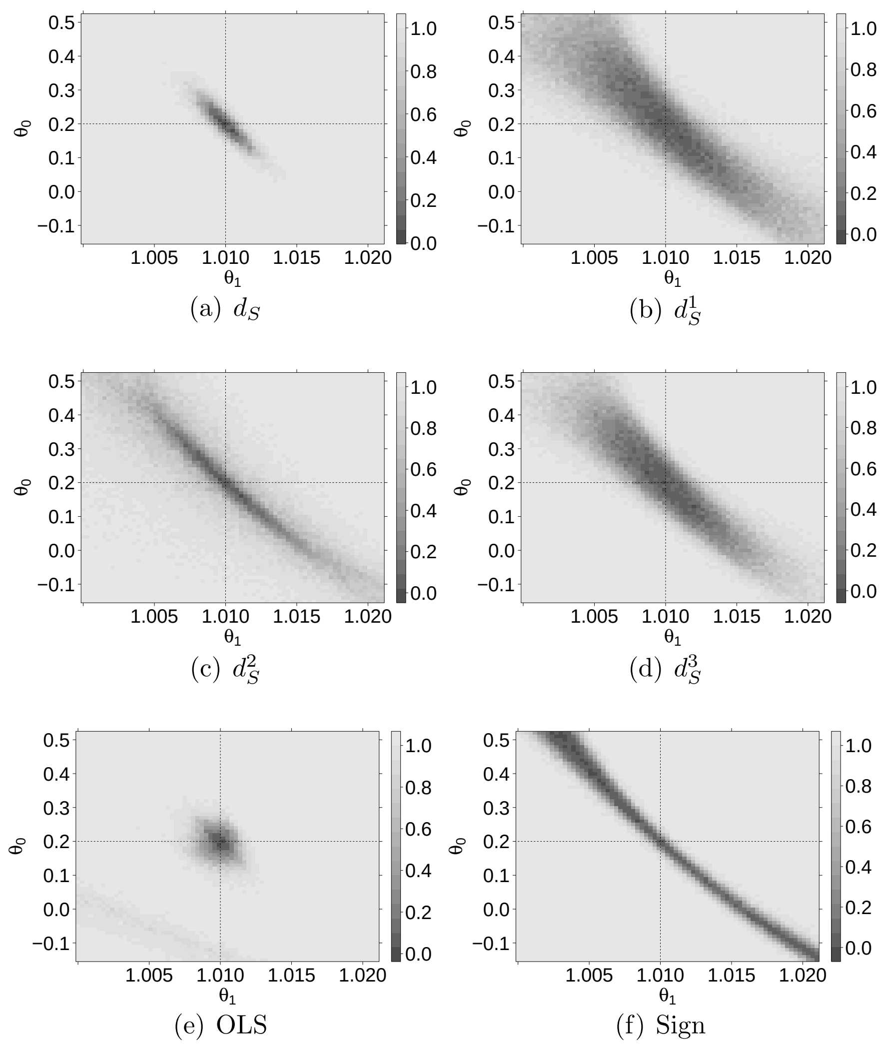
and parameters $\alpha = 1.928, \beta = -2, \gamma = 10$. Thereby, the normal distribution and the Fréchet distribution satisfy $med(E_n) = 0$, but only the Fréchet distribution also satisfies the conditions of Theorem 4.1 for consistency, if the starting value is large enough. The aim of the simulation study is to show how the proposed simplicial depth test behaves if the assumptions are not completely satisfied. In particular, the errors with contaminated normal distribution provide innovation outliers in the sense of Fox (1972).

Figure 3 shows the power functions for the normally distributed errors. Thereby, the horizontal and vertical lines denote the components of θ^0 so that their intersection is θ^0. One can clearly see, that the OLS test performs best under the normal distribution. This is not surprising, since it assumes the correct error distribution. The sign test behaves quite well close to the alternative. Unfortunately in case of explosive processes the power also decreases when θ_0 and θ_1 lead to residuals which have a poor fit but indicate an error median of zero. This for example happens, if the first half of residuals is positive and the second half is negative. As a result this test is very unstable in case of explosive AR(1) processes. The d_S test clearly outperforms the simplified depth tests. It also shows a better performance than the OLS test in direction of a diagonal with positive slope, but accepts a wider range of values on a diagonal with negative slope.

In Figure 4, the comparison for errors with the contaminated normal distribution is depicted, and Figure 5 provides the comparison for errors with the Fréchet distribution. Now the simplicial depth test performs clearly best. The OLS test suffers from heavy bias due to the skewed error distributions and the sign test still shows the identifiability problem.

In Figures 6, 7, 8, we compare the tests evaluated on the diagonal given by $\theta_0 = 50.7 - 50 \cdot \theta_1$, where the slope of the diagonal is negative. The straight line goes from $(-0.325, 1.0205)$ to $(0.725, 0.9995)$ through H_0 defined by $\theta = (0.2, 1.01)^T$. In the Figures the x-axis is defined by the parameter $\lambda \in [0, 1]$ from the parametric form of the straight line given by $(0.725, 0.9995)^T + \lambda \cdot (-1.05, 0.021)^T$. On this line $\lambda = 0.5$ coincides with H_0. Here, the main advantage of the full simplicial depth compared to the sign test is clearly visible. Additionally, these figures show how the new test outperforms the OLS test in the case of nonnormal errors where the OLS test in particular does not keep the level anymore.

Summarising we see, that the d_S test can be applied to explosive AR(1) processes under quite general conditions and does not suffer of systematic failure or heavy bias in case of skewed errors or outliers. Further, by the price of additional computational costs, the full simplicial depth statistic defines a test with higher power than the simplified statistics based on simplicial depth.

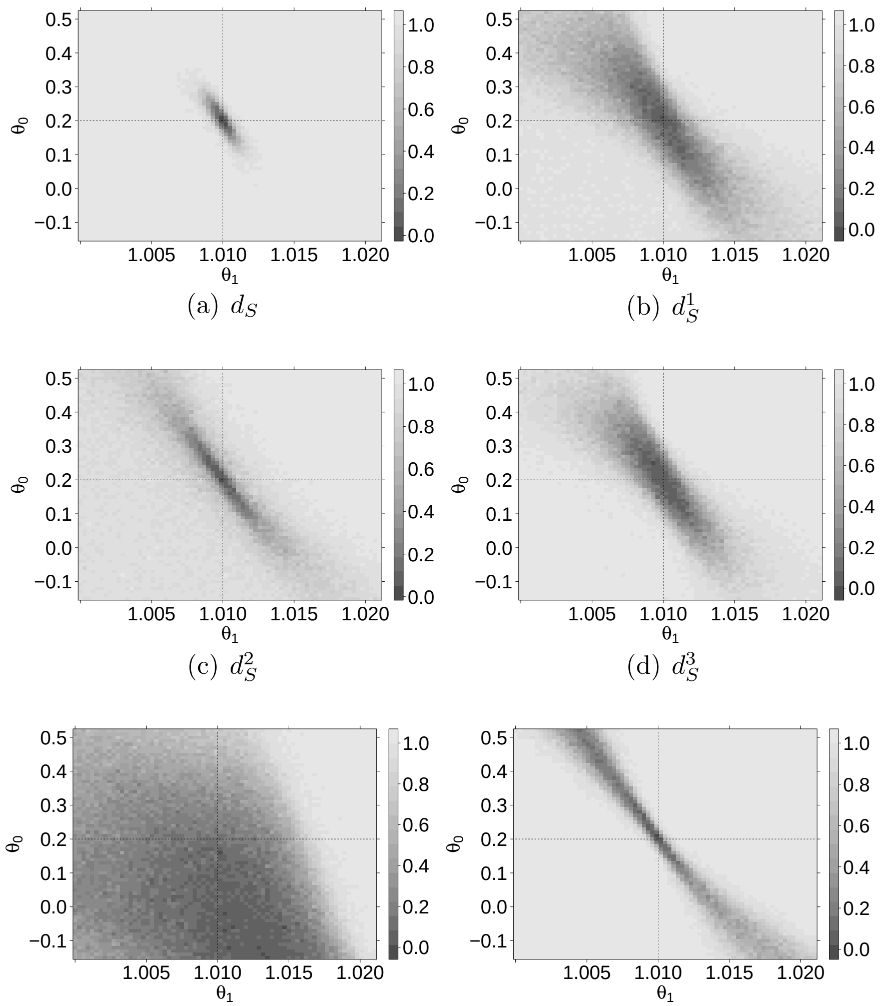


(a) d_S (b) d_S^1

(c) d_S^2 (d) d_S^3

(e) OLS (f) Sign

FIGURE 3. Power of the tests based on normally distributed errors

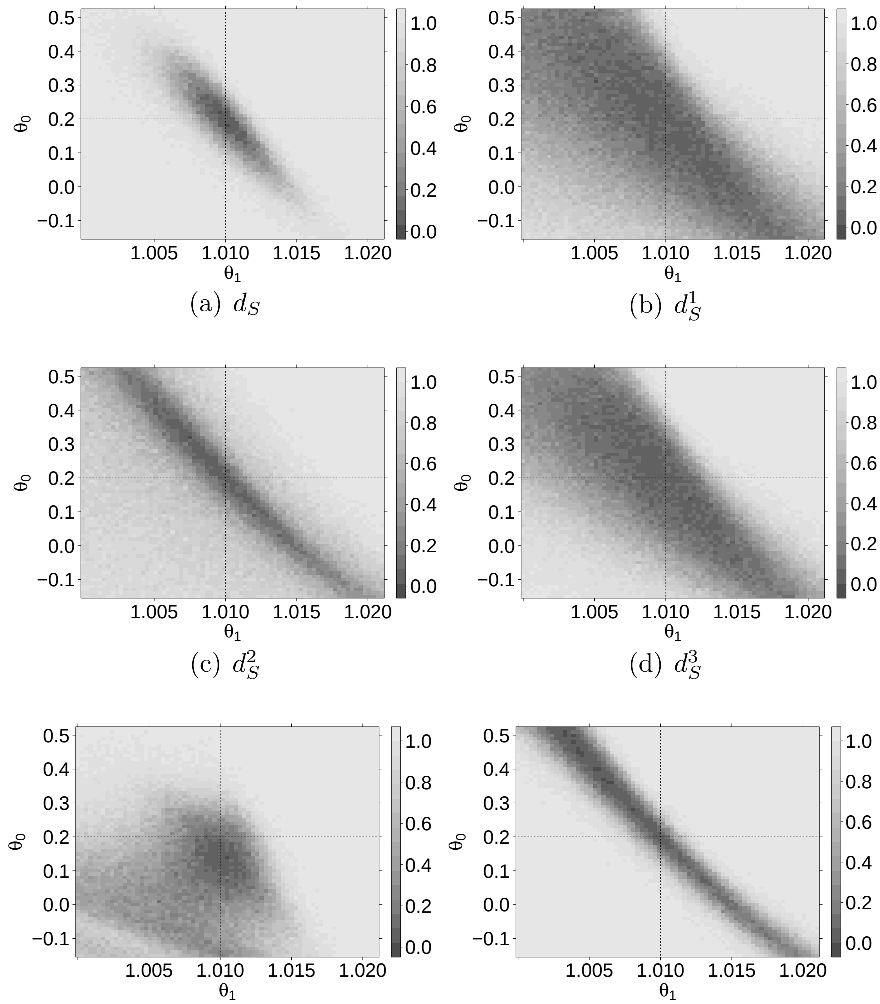


(a) d_S (b) d_S^1

(c) d_S^2 (d) d_S^3

(e) OLS (f) Sign

FIGURE 4. Power of the tests based on errors with contaminated normal distribution



(a) d_S (b) d_S^1

(c) d_S^2 (d) d_S^3

(e) OLS (f) Sign

Figure 5. Power of the tests based on errors with Fréchet distribution

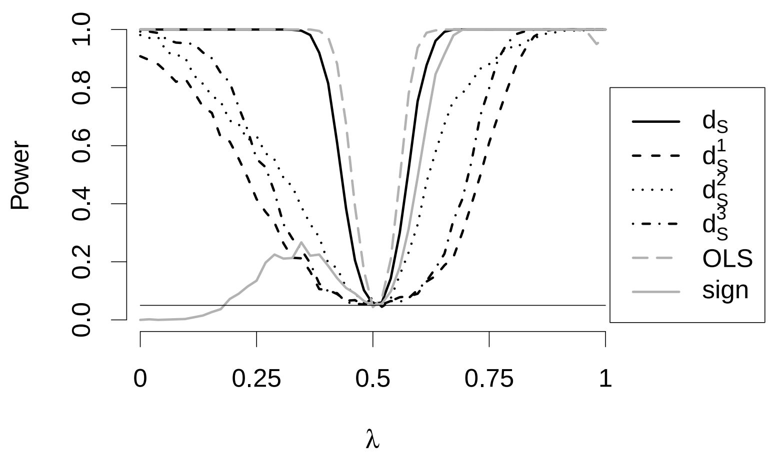


FIGURE 6. Power evaluated along $\theta = (0.725, 0.9995)^\top + \lambda(-1.05, 0.021)^\top$ for normally distributed errors

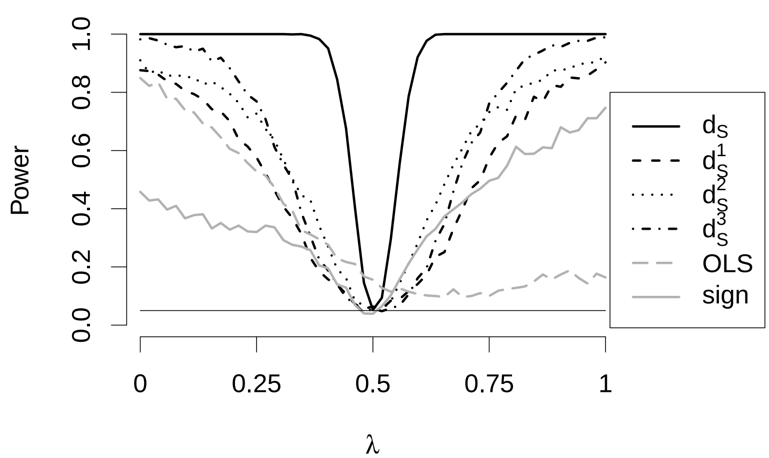


FIGURE 7. Power evaluated along $\theta = (0.725, 0.9995)^\top + \lambda(-1.05, 0.021)^\top$ for errors with contaminated normal distribution

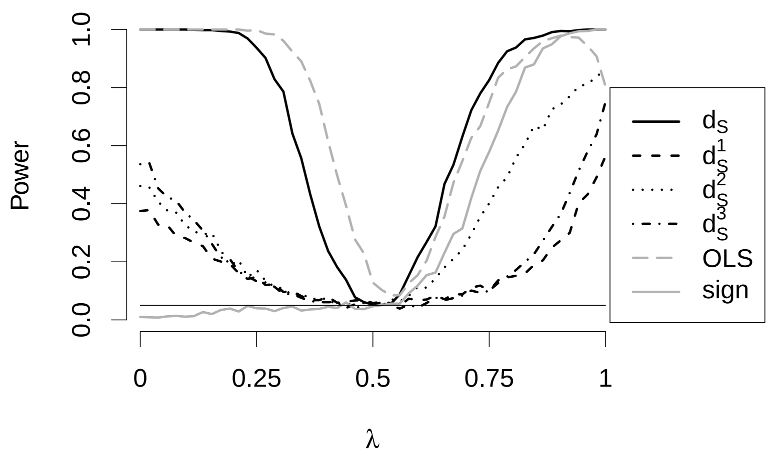


FIGURE 8. Power evaluated along $\theta = (0.725, 0.9995)^\top + \lambda(-1.05, 0.021)^\top$ for errors with Fréchet distribution

6. Application

Figure 9 shows the growth of the crack width y_n in prestressed concrete in an experiment conducted by Maurer and Heeke (2010), where $n = 1, ..., 75$ is the discrete observation index recorded at each 2256 load cycles.

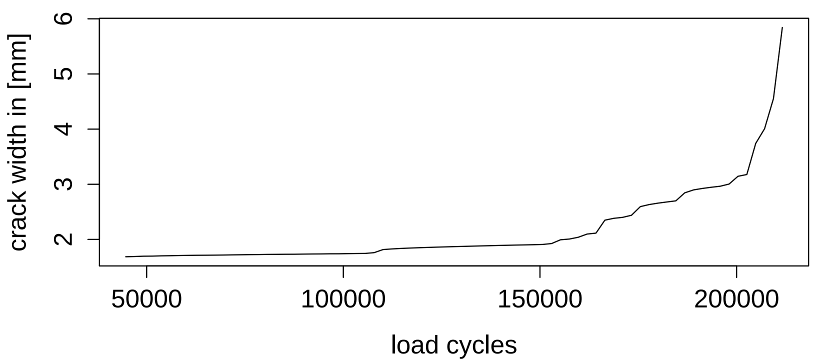


Figure 9. Observed crack growth processes in mm recorded at load cycles ranging from 44761 to 211624 measured in steps of 2256.

Crack growth is usually modelled by the so called deterministic Paris-Erdogan equation (see e.g. Pook, 2000, Chapter 2.1) given by

$$\frac{dy}{dt} = \theta y^k,$$

where y is the crack length or crack width, t the time usually measured by the number of load cycles, k is a material constant often assumed to satisfy $k = 1$, and θ is an unknown parameter. Since crack growth is not a deterministic process, a stochastic differential equation given by

$$dY_t = \theta Y_t^k \, dt + dZ_t$$

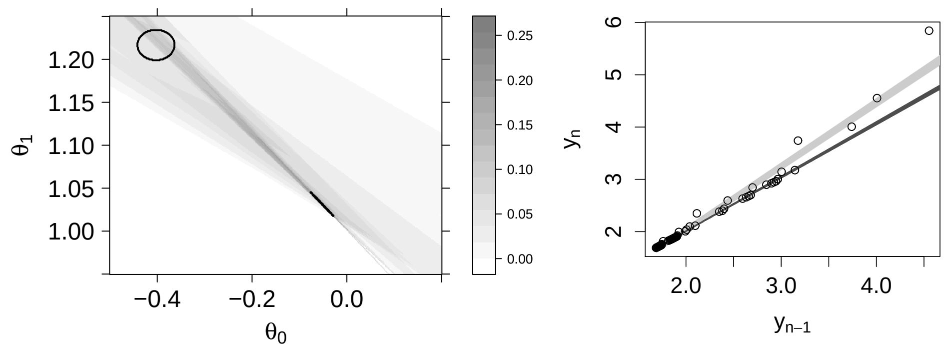
is more adequate, where Z_t is an error process which could be the Wiener process but also some nonnegative process like a Gamma or Gamma-Poisson process. Since the process Y_t cannot be observed continuously, a discretization must be used to estimate θ. The simplest discretization is given by the Euler-Maruyama approximation (see e.g. Iacus, 2008, Chapter 2.1) leading to

$$Y_n = Y_{n-1} + \theta Y_{n-1}^k (t_n - t_{n-1}) + Z_n - Z_{n-1},$$

where Y_n and Z_n are the processes at t_n. The usual assumption is that the increments $\tilde{E}_n := Z_n - Z_{n-1}$, $n = 1, \ldots, N$, are independent. In Kustosz and Müller (2014), the AR(1) model

$$Y_n = Y_{n-1} + \theta Y_{n-1}(t_n - t_{n-1}) + \tilde{E}_n$$

was used, where the jumps visible in the crack growth process in Figure 9 were considered as outliers since they are caused by the breaking of the prestressing wires. To ensure that the crack width Y_n is increasing, the error E_n must be bounded by below. However, $\tilde{E}_n$ should possess an infinitely divisible distribution to be the increment of a continuous process Z_t. Such distributions are either unbounded or bounded by zero, see e.g. Klenke (2006), Chapter 16. But if the distribution is bounded by zero its median cannot be zero. To ensure an infinitely divisible distribution with median equal to zero, we can subtract the median of $\tilde{E}_n$ from $\tilde{E}_n$ by setting $E_n := \tilde{E}_n - \text{med}(\tilde{E}_n)$. This leads to an AR(1) model of the form (1.1), where θ_0 is the median of the increment $\tilde{E}_n$ and θ_1 is the autoregressive parameter θ multiplied with the stepwidth introduced by discrete observations. In Kustosz and Müller (2014), only the drift parameter θ_1 was estimated and tested in a model without θ_0 and the requirement of an infinitely divisible distribution was neglected. With the results presented here we are able to estimate and test θ_1 as well as θ_0 in a more adequate model.



(a) Empirical simplicial depth for the observed crack growth process. The black line around $\theta_0 = -0.1$ is the 95% parameter confidence region derived by d_S. The region marked by the circle is the 95% parameter region interval resulting from an ordinary least squares estimator.

(b) Fit of the residuals of the crack growth process. The dots represent the pairs (y_{n-1}, y_n) from the observed data. The light gray region are fits based on the 95% OLS parameter confidence sets. The dark gray region is defined by fits from the 95% confidence regions based on the d_S test.

FIGURE 10. Empirical results for crack growth data

Figure 10(a) provides the depth contours and the 95% confidence set for $\theta = (\theta_0, \theta_1)^\top$ based on d_S and the ordinary least squares estimator. The robustness of the proposed method can be reviewed by consideration of the fits compared to a plot of y_n against y_{n-1} presented in Figure 10(b). While the OLS confidence region clearly follows the large increments and overestimates the central observations, the d_S confidence regions are more concentrated on the main part of the residuals. Further the OLS confidence region is remarkably wider than the depth based region, demonstrating a better performance of the depth based method.

Acknowledgement

We gratefully thank Prof. Dr.-Ing. Reinhard Maurer and Dipl.-Ing. Guido Heeke form the Chair of Concrete Construction of the TU Dortmund University for providing the data for the application in Section 6.

References

Agostinelli, C. (2003). Robust time series estimation via weighted likelihood. In *Developments in Robust Statistics*, eds. R. Dutter, P. Filzmoser, U. Gather, P. J. Rousseeuw, Physica-Verlag, Heidelberg, 266-276.

Anderson, T. (1959). On asymptotic distributions of estimates of parameters of stochastic difference equations. *Ann. Math. Statist.* 30, 676-687.

Basawa, I., Mallik, A., McCormick, W. and Taylor, R. (1989). Bootstrapping explosive autoregressive processes. *Ann. Statist.* 17, 1479-1486.

Bazarova, A., Berkes, I. and Horváth, L. (2014). Trimmed stable AR(1) processes. To appear in *Stochastic Process. Appl.*

Boldin, M. V. (2011). Local robustness of sign tests in AR(1) against outliers. *Math. Methods Statist.* 20(1), 1-13.

Denecke, L., and Müller, Ch.H. (2011). Robust estimators and tests for bivariate copulas based on likelihood depth. *Comput. Statist. Data Anal* 55, 2724-2738.

Denecke, L., and Müller, Ch.H. (2012). Consistency and robustness of tests and estimators based on depth. *J. Statist. Plann. Inference* 142, 2501-2517.

Denecke, L., and Müller, Ch.H. (2013). New robust tests for the parameters of the Weibull distribution for complete and censored data. To appear in *Metrika*. DOI:10.1007/s00184-013-0454-8.

Denecke, L., and Müller, Ch.H. (2014). Consistency of the likelihood depth estimator for the correlation coefficient. *Statist. Papers* 55, 3-13.

Fox, J. (1972). Outliers in time series. *J. Roy. Statist. Soc.* 43, 350-363.

Fried, R. and Gather, U. (2005). Robust trend estimation for AR(1) disturbances. *Austrian J. Statist.* 34, 139-151.

Gelper, S., Fried, R. and Croux, C. (2009). Robust forecasting with exponential and Holt-Winters smoothing. *J. Forecast.* 29, 285-300.

Grillenzoni, C. (2009). Robust non-parametric smoothing of non-stationary time series. *J. Stat. Comput. Simul.* 79, 379-393.

Grossi, L. and Riani, M. (2002). Robust time series analysis through the forward search. In *Proc. Comput. Statist.*, eds. W. Härdle, B. Rönz, Physica-Verlag, Heidelberg, 521-526.

Huggins, R. (1989). The sign test for stochastic processes. *Aust. N. Z. J. Stat.* 31, 153-165.

Hwang, S.Y. (2013). Arbitrary initial values and random norm for explosive AR (1) processes generated by stationary errors. *Statist. Probab. Lett.*, 83, 127-134.

Hwang, S.Y. and Basawa, I.V. (2005). Explosive random coefficient AR(1) processes and related asymptotics for least squares estimation. *J. Time Series Anal.* 26, 807-824.

Hwang, S.Y., Kim, S., Lee, S.D. and Basawa, I.V. (2007). Generalized least squares estimation for explosive AR(1) processes with conditionally heteroscedastic errors. *Statist. Probab. Lett.* 77, 1439-1448.

Iacus, S. (2008), *Simulation and Inference for Stochastic Differential Equations*, Springer, New York.
Liu, R.Y. (1988). On a notion of simplicial depth. *Proc. Nat. Acad. Sci. USA* 85, 1732-1734.
Liu, R.Y. (1990). On a notion of data depth based on random simplices. *Ann. Statist.* 18, 405-414.
Klenke, A. (2006). *Probability Theory. A Comprehensive Course.* Springer, Berlin.
Kustosz, C. P. and Müller, Ch. H. (2014). Analysis of crack growth with robust distribution-free estimators and tests for nonstationary autoregressive processes. *Statist. Papers* 55, 125-140.
Kustosz, C. P., Müller, Ch. H. and Wendler, M. (2015). Simplified simplicial depth for regression and autoregressive growth processes. *Submitted.*
Maronna, R. (2006). *Robust Statistics: Theory and Methods.* Wiley, New York.
Maurer, R. and Heeke, G. (2010). Ermüdungsfestigkeit von Spannstählen aus einer älteren Spannbetonbrücke. *Bericht für Landesbetrieb Straßenbau NRW.* Technical report, TU Dortmund University.
Mizera, I. (2002). On depth and deep points: a calculus. *Ann. Statist.* 30, 1681-1736.
Müller, Ch.H. (2005). Depth estimators and tests based on the likelihood principle with application to regression. *J. Multivariate Anal.* 95, 153-181.
Novomestky, F. (2012). matrixcalc: collection of functions for matrix calculations. R package version 1.0-3. http://CRAN.R-project.org/package=matrixcalc
Paulaauskas, V. and Rachev, T. (2003). Maximum likelihood estimators in regression models with infinite variance innovations. *Statist. Papers* 44, 47-65.
Pollard, D. (1984). *Convergence of Stochastic Processes.* Springer, New York.
Pook, L. (2000), *Linear Elastic Fracture Mechanics for Engineers: Theory and Application*, WIT Press, Southampton.
R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.
Rousseeuw, P.J. and Hubert, M. (1999). Regression depth. *J. Amer. Statist. Assoc.* 94, 388-433.
Stute, W. and Gründer, B. (1993). Nonparametric prediction intervals for explosive AR(1)-processes. *J. Nonparametr. Stat.* 2, 155-167.
Tierney, L., Rossini, A.J., Li, N. and Sevcikova, H. (2013). snow: simple network of workstations. R package version 0.3-13. http://CRAN.R-project.org/package=snow
Tukey, J.W. (1975). Mathematics and the picturing of data. In *Proc. Int. Congress Math.*, Vancouver 1974, 2, 523- 531.
Van der Vaart, A. W. and Wellner, J. A. (2000). *Weak Convergence and Empirical Processes. With Applications to Statistics.* Springer, New York.
Venables, W. N. and Ripley, B. D. (2002). *Modern Applied Statistics with S.* Fourth Edition. Springer, New York.
Wang, X. and Yu, J. (2013). Limit theory for an explosive autoregressive process. *Working Paper*, No. 08-2013. Singapore Management University, School of Economics.
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.

Wellmann, R. and Müller, Ch.H. (2010a). Tests for multiple regression based on simplicial depth. *J. Multivariate Anal.* 101, 824-838.

Wellmann, R. and Müller, Ch.H. (2010b). Depth notions for orthogonal regression. *J. Multivariate Anal.* 101, 2358-2371.

Yu, H. (2002). Rmpi: parallel statistical computing in R. R News 2/2:10-14.

α	0.005	0.01	0.025	0.05	0.1	0.2
q_α	-2.6543	-2.2404	-1.6792	-1.2545	-0.8271	-0.4032

Table 1. Quantiles of the integrated Gaussian process