SFB
823

Powerful generalized sign tests based on sign depth

Kevin Leckey, Dennis Malcherczyk,
Christine H. Müller

Discussion Paper

Nr. 12/2020

SFB
823

Powerful generalized sign tests based on sign depth

Kevin Leckey* Dennis Malcherczyk* Christine H. Müller*

April 29, 2020

Abstract

The classical sign test usually provides very bad power for certain alternatives. We present a generalization which is similarly easy to comprehend but much more powerful. It is based on K-sign depth, shortly denoted by K-depth. These so-called K-depth tests are motivated by simplicial regression depth, but are not restricted to regression problems. They can be applied as soon as the true model leads to independent residuals with median equal to zero. Moreover, general hypotheses on the unknown parameter vector can be tested. Since they depend only on the signs of the residuals, these test statistics are outlier robust. While the 2-depth test, i.e. the K-depth test for $K = 2$, is equivalent to the classical sign test, K-depth test with $K \geq 3$ turn out to be more powerful in many applications. As we will briefly discuss, these tests are also related to runs tests. A drawback of the K-depth test is its fairly high computational effort when implemented naively. However, we show how this inherent computational complexity can be reduced. In order to see why K-depth tests with $K \geq 3$ are more powerful than the classical sign test, we discuss the asymptotic behaviour of its test statistic for residual vectors with only few sign changes, which is in particular the case for some nonfits the classical sign test cannot reject. In contrast, we also consider residual vectors with alternating signs, representing models that fit the data very well. Finally, we demonstrate the good power of the K-depth tests for quadratic regression.

Keywords: K-sign depth, sign test, runs test, outlier robust, distribution free, quadratic regression

1 Introduction

We consider stochastic models where a parameter $\theta \in \Theta \subset \mathbb{R}^p$, $p \in \mathbb{N}$, is unknown and where residuals $R_1(\theta), \ldots, R_N(\theta)$ of N observations in $\mathbb{R}$ are independent with

$$P_\theta(R_n(\theta) > 0) = \frac{1}{2} = P_\theta(R_n(\theta) < 0). \tag{1}$$

This assumption is fulfilled by every continuous distribution with median zero. Examples of such models are linear and nonlinear regression models with additive errors E_n where the observations are of the form $Y_n = g(x_n, \theta) + E_n$ with $x_n \in \mathbb{R}^q$ leading to residuals $R_n(\theta) = Y_n - g(x_n, \theta)$. Generalized linear and nonlinear models are further examples if the link function can be expressed by the median

*Faculty of Statistics, TU Dortmund University, D-44227 Dortmund, Germany

of the observations Y_n, i.e. if $\text{med}(Y_n) = g(x_n, \theta)$. More examples are given by stochastic processes with i.i.d. increments such as AR(p) processes given by $Y_n = g(Y_{n-1}, \dots, Y_{n-p}, \theta) + E_n$.

In models given by (1), the classical sign test can be used for testing hypotheses $H_0 : \theta = \theta^0$ and for deriving confidence sets. This test counts the number of positive (or negative) residuals and rejects the null hypothesis if the number of positive signs is too small or too large. In particular, it does not reject the null hypothesis $H_0 : \theta = \theta^0$ if half of the residuals $R_n(\theta^0)$ are positive and half of them are negative. However, this can also happen for alternatives with parameters far away from θ^0, see Figure 1, p. 9. Hence this test has a very bad power for such alternatives. This can also be seen in the simulation studies of Kustosz et al. (2016a,b) where the classical sign test is compared to tests based on simplicial regression depth for linear and nonlinear regression and autoregression with two unknown regression parameters.

Simplicial regression depth is a modification of the regression depth introduced by Rousseeuw and Hubert (1999) to generalize the depth notion to regression. Originally, the halfspace depth of Tukey (1975) was used to obtain a generalization of the median for multivariate data. Liu (1988, 1990) extended this to simplicial depth. Simplicial depth can be expressed by counting the number of all $p + 1$-tupels of the p-dimensional data set with positive halfspace depth. Replacing halfspace depth by regression depth leads to simplicial regression depth.

For defining regression depth, Rousseeuw and Hubert (1999) introduced the concept of nonfit. They defined the nonfit via residuals. However, Mizera (2002) extended this for arbitrary quality functions and Mizera and Müller (2004); Müller (2005); Denecke and Müller (2011) used it for likelihood functions. This also led to depth concepts for scatter as proposed by Mizera and Müller (2004); Paindaveine and Van Bever (2018); Wang (2019) although most depth concepts concern central parts of a data set as those of Zuo and Serfling (2000); Mosler (2002); Agostinelli and Romanazzi (2011); Lok and Lee (2011); Paindaveine and Van Bever (2013); Dehghan and Faridrohani (2019). In this context, datasets consisting of functional data were considered as well, see e.g. López-Pintado and Romo (2009); Claeskens et al. (2014); López-Pintado et al. (2014); Cuesta-Albertos et al. (2017); Nagy and Ferraty (2018).

Simplicial depth has the advantage that it is a U-statistics although it is often a degenerated U-statistic so that more effort is necessary to derive the asymptotic distribution, see Dümbgen (1992); Müller (2005); Wellmann et al. (2009); Wellmann and Müller (2010). Moreover, for its calculation, Rousseeuw and Hubert (1999) and Müller (2005) noted that the regression depth of a p-dimensional parameter vector within $p + 1$ observations is greater than zero if and only if the residuals have alternating signs. Sufficient conditions for this equivalence and a proof of this property are given by Kustosz et al. (2016b). One of the sufficient conditions is that the observations are given by a natural order as this is the case for time series. This was the reason that the proof of the asymptotic distribution of the simplicial regression depth for $p = 2$ was given by Kustosz et al. (2016a) for AR(1) regression. However, the proof is not restricted to AR(1) regression since it uses only the alternating signs of $p + 1 = 3$ residuals. In particular, the derived asymptotic distribution can be used as soon as there is an approriate ordering of the observations and the median of the residuals is zero. This leads to the idea to define simplicial depth not via regression depth but via alternating signs of residuals.

We call this depth notion K-sign depth or shortly K-depth where K stands for the number of residuals used in the simplicial depth. In contrast to the regression depth, it is not necessary to choose $K = p + 1$ if the unknown parameter vector is p-dimensional. Tests based on this depth notion are called K-depth tests. They can be used to test arbitrary null hypotheses of the form $H_0 : \theta \in \Theta^0$. We show in this paper that the K-depth test with $K = 2$ is equivalent to the classical sign test, hence K-depth tests with $K \geq 2$ are indeed generalizations of this test. Moreover, we demonstrate that K-depth tests with $K > 2$ are much more powerful. In particular, in contrast to the classical sign test, they do not have the drawback of not rejecting alternatives for which (nearly) half of the residuals are positive. Furthermore, these tests are robust against outliers caused by heavy-tailed errors since they are based only on signs of residuals.

In Section 2, we introduce the K-depth and the K-depth tests, discuss a relationship to the runs test, and show how the computational complexity can be reduced by block implementation. Basic properties of the K-depth are derived in Section 3. This concerns a strong law of large numbers for the K-depth, the behaviour at alternating signs of residuals and the behaviour when only few sign changes occur. In particular, it is shown that the expected value is asymptotically an upper bound for K-depth and is not reached in situations of few sign changes for $K \geq 3$ which provides an explanation for the good power of the K-depth tests for $K \geq 3$ at alternatives that lead to only few sign changes. A comparison between the K-depth tests for different values of K is given in Section 4. At first, for $K = 2$, the equivalence of the K-depth test and the classical sign test is derived formally. Afterwards, the K-depth tests with $K = 3, 4, 5, 6$ are compared by p-values in some worst case scenarios with few sign changes which were derived in Section 3. Section 5 demonstrates the good power of the K-depth tests for $K = 3$ and $K = 4$ via simulations for quadratic regression. Finally, a discussion of the results and an outlook is given in Section 6. All proofs can be found in the appendix.

Notation. Throughout the article, $r_1(\theta), \ldots, r_N(\theta)$ denote realisations of $R_1(\theta), \ldots, R_N(\theta)$. If the choice of the parameter θ is clear, we also use the abbreviations $r_n := r_n(\theta)$ and $R_n := R_n(\theta)$ for $n = 1, \ldots, N$. The sign of a real number x is denoted by $\psi(x) = \mathbb{1}\{x > 0\} - \mathbb{1}\{x < 0\}$, where $\mathbb{1}\{\cdot\}$ denotes the indicator function. In some asymptotic calculations we make use of the O-Notation: For real-valued sequences $(a_n)_{n\geq 1}$ and $(b_n)_{n\geq 1}$, we write $a_n = O(b_n)$ if there is a constant $C > 0$ and an integer n_0 with $|a_n| \leq C|b_n|$ for all $n \geq n_0$. Furthermore, $a_n = \Theta(b_n)$ denotes that both $a_n = O(b_n)$ and $b_n = O(a_n)$.

2 K-depth tests and reduction of their computational complexity

In this section, we introduce the K-depth of a vector and how to use the K-depth notion as a test statistic. We also briefly discuss the issue of a fairly high computational complexity when working with K-depth tests. This issue can be resolved by using alternative representations of the original definition of the K-depth.

2.1 K-depth and K-depth tests

The **K-sign depth** or shortly **K-depth** $d_K(r_1, \ldots, r_N)$ of $r_1, \ldots, r_N$ is the relative number of K-element subsets with alternating signs, i.e. for $K \geq 2$,

$$d_K(r_1, \ldots, r_N) := \frac{1}{\binom{N}{K}} \sum_{1 \leq n_1 < n_2 < \ldots < n_K \leq N} \Big(\prod_{k=1}^{K} \mathbb{1}\left\{(-1)^k r_{n_k} > 0\right\} + \prod_{k=1}^{K} \mathbb{1}\left\{(-1)^k r_{n_k} < 0\right\} \Big). \tag{2}$$

Remark 2.1. Note that the definition of the K-sign depth depends on the chosen order and therefore this choice is a crucial aspect. If $x_n \in \mathbb{R}^q$ for $q > 1$ then various multivariate orderings can be used. Not all of them provide powerful tests. However, a paper is in preparation where we derive data based orderings which lead to powerful tests. The following arguments hold also for this multivariate case. In the applications of this paper, we focus on the univariate case $q = 1$ and choose the canonical order on $\mathbb{R}$.

In order to obtain a non-degenerate limit distribution, the **K-depth test** is based on the following test statistic:

$$\begin{aligned} T_K(\theta) &:= T_K(R_1(\theta), \ldots, R_N(\theta)) \\ &:= N\left(d_K(R_1(\theta), \ldots, R_N(\theta)) - \left(\frac{1}{2}\right)^{K-1}\right). \end{aligned} \tag{3}$$

A test based on (3) requires the α-quantiles of the distribution of the test statistic. If N is small, the finite sample distribution for any K can be easily simulated since the determination of the K-depth with an underlying C++ algorithm computing Formula (2) is fairly fast for small N. For larger N, see Subsection 2.2.

With the quantiles at hand, the K-depth test, $K \geq 2$, is defined as in Müller (2005): A hypothesis of the form $H_0 : \theta \in \Theta^0$ shall be rejected if the K-depth $d_K(r_1(\theta), \ldots, r_N(\theta))$ of θ or $T_K(\theta)$ is too small for all $\theta \in \Theta^0$. Hence, if q_α is the α-quantile of the distribution of $T_K(\theta)$ under θ then the K-depth test for $H_0 : \theta \in \Theta^0$ is given by

$$\text{reject } H_0 : \theta \in \Theta^0 \text{ if } \sup_{\theta \in \Theta^0} T_K(\theta) < q_\alpha. \tag{4}$$

Remark 2.2. The K-depth test can also be used in a two-sided version:

$$\text{reject } H_0 : \theta \in \Theta^0 \text{ if } \sup_{\theta \in \Theta^0} T_K(\theta) < q_{\frac{\alpha}{2}} \text{ or } \inf_{\theta \in \Theta^0} T_K(\theta) > q_{1-\frac{\alpha}{2}}.$$

This test also rejects H_0 if too many sign changes occur in the residual vector, which is an indicator for negatively correlated residuals. While the one-sided version is mostly focused on detecting deviations from 0 in the median and can detect only strong positive correlation in the residuals, the two-sided version is the preferable choice when testing simultaneously whether the residuals are independent and have median zero. The hypothesis of independent residuals can also be tested with the runs test of Wald and Wolfowitz (1940), see e.g. Gibbons and Chakraborti (2003), pp. 78-86. In fact, there is a simplified version of the K-depth which can be considered as a generalization of the runs test: This simplified K-depth uses only subsequent residuals and can be defined as in Kustosz et al. (2016b) for $K \geq 2$ by

$$\begin{aligned} d_K^S(r_1, \ldots, r_N) := \frac{1}{N-K+1} \sum_{n=1}^{N-K+1} \Big(&\prod_{k=1}^{K} \mathbb{1}\left\{(-1)^k r_{n+k-1} > 0\right\} \\ &+ \prod_{k=1}^{K} \mathbb{1}\left\{(-1)^k r_{n+k-1} < 0\right\} \Big). \end{aligned} \tag{5}$$

If $K = 2$ then this simplified K-depth counts the number of sign changes and thus the number of runs. Kustosz et al. (2016b) used the simplified versions because they are faster to compute and their asymptotic behaviour is easy to derive. However, since the simplified K-depth only considers $N-K+1$ subsets instead of $\binom{N}{K}$, tests based on it are usually less powerful than tests based on the full K-depth, in particular if the independence of the residuals is ensured. This can be clearly seen from the examples in Kustosz et al. (2016a) and Falkenau (2016) for AR(1)-models. Since our main interest lies in testing whether the median of the fitted residuals is zero, neither the two-sided version nor the simplified version is used here.

2.2 Runtime and block-implementation

A major drawback of the K-depth test is its slow runtime when using an algorithm based on the definition (2). This definition requires the consideration of all increasing K-tuples in $\{1,\ldots,N\}$, hence leading to an algorithm with runtime $\Theta(N^K)$. Such an algorithm is clearly impractical in applications with fairly large sample sizes. Fortunately, the derivation of a limit theorem of the test statistic $T_K(\theta)$ leads to an asymptotically equivalent form of (2) which can be computed in linear time for all $K\geq 3$. Such a limit theorem is given in Kustosz et al. (2016a) for $K=3$ and a generalization to arbitrary K is currently submitted elsewhere.

We will not go into detail on how this algorithm with runtime $\Theta(N)$ works since it requires a major part of the computation necessary to obtain the limit theorem and this is beyond this paper. Instead, we discuss a different approach which leads to an algorithm useful for residual vectors with only few sign changes. We refer to this approach as *block-implementation.* Aside from speeding up the implementation based on (2), this approach will be useful do derive some of the properties presented in Section 3.

Block-implementation. Let $r:=(r_1,\ldots,r_N)$ be a vector of residuals and let $\psi(x)$ denote the sign of a real number x, i.e. $\psi(x):=\mathbb{1}\{x>0\}-\mathbb{1}\{x<0\}$. The vector r is decomposed into *blocks* by letting a new block start at index j if and only if r_{j-1} and r_j have different signs. More formally, we define the number $B(r)$ of blocks and their starting positions $s_1(r),\ldots,s_{B(r)}(r)$ via $s_1(r):=1$ and

$$B(r) := 1+\sum_{n=2}^{N}\mathbb{1}\left\{\psi\left(r_{n-1}\right)\neq\psi\left(r_n\right)\right\},$$
$$s_b(r) := \min\left\{\ell>s_{b-1}(r);\,\psi\left(r_\ell\right)\neq\psi\left(r_{\ell-1}\right)\right\},\quad b=2,\ldots,B(r).$$

For convenience, we define $s_{B(r)+1}(r):=N+1$. The *block sizes* are defined as

$$q_b(r):=s_{b+1}(r)-s_b(r),\quad b=1,\ldots,B(r).$$

Example 2.1. The vector $r=(1,2,-1,3)$ consists of $B(r)=3$ blocks with starting positions $1,3,4$. The block sizes are $q_1(r)=2$ and $q_2(r)=q_3(r)=1$.

We say that the nth residual r_n belongs to block j if and only if $s_j(r)\leq n<s_{j+1}(r)$. The sign of a block j is defined as the sign of the first (and thus any) element $r_{s_j(r)}$ belonging to that block. Blocks $j_1<\ldots<j_k$ are called alternating if and only if the signs of the blocks are alternating, i.e. the signs of block j_i and j_{i+1} are different for all $i=1,\ldots,k-1$. Note that two blocks j_1 and j_2 have different signs if and only if j_1 is even and j_2 is odd or vice versa. In particular, the blocks $j_1<\ldots<j_k$ are alternating if and only if $j_{i+1}-j_i$ is odd for all $i=1,\ldots,k-1$.

The advantage of decomposing the residuals into blocks is that it helps to identify the K-tuples with alternating signs: A tuple $(r_{n_1},\ldots,r_{n_K})$ with $1\leq n_1<\ldots<n_K\leq N$ has alternating signs if

and only if $r_{n_1}, \ldots, r_{n_K}$ belong to alternating blocks $j_1 < \ldots < j_K$. In summary, we therefore have the following alternative representation of (2):

Lemma 2.2. *Let* $2\mathbb{N}_0 + 1$ *denote the set of all odd positive integers and let*

$$\mathcal{A}_{K,B} := \left\{(i_1, \ldots, i_K) \in \{1, \ldots, B\}^K;\ i_k - i_{k-1} \in 2\mathbb{N}_0 + 1 \text{ for } k = 2, \ldots, K\right\},$$

$$d_{K,N,B}(q_1, \ldots, q_B) := \frac{1}{\binom{N}{K}} \sum_{(i_1,\ldots,i_K)\in\mathcal{A}_{K,B}} \prod_{k=1}^{K} q_{i_k} \quad \text{for } B \in \mathbb{N},\ q_1, \ldots, q_B > 0.$$

Let $q_1(r), \ldots, q_{B(r)}(r)$ *be the block sizes of a residual vector* $r = (r_1, \ldots, r_N)$. *Then*

$$d_K(r_1, \ldots, r_N) = d_{K,N,B(r)}(q_1(r), \ldots, q_{B(r)}(r)). \tag{6}$$

Remark 2.3. Note that the size of $\mathcal{A}_{K,B}$ is $\Theta(B^K)$. Also note that the effort to compute the block sizes $q_1(r), \ldots, q_{B(r)}(r)$ of a vector $r = (r_1, \ldots, r_N)$ is $\Theta(N)$. Hence, an algorithm based on the expression in Lemma 2.2 has computational complexity $\Theta(N + B^K)$ if $B = B(r)$ is the number of blocks in r. This is a significant improvement to a direct implementation of (2) if the number of blocks is much smaller than N. Since, even in the worst case $B(r) = N$, the complexity of computing (6) never exceeds the complexity of (2), one should always use the block-implementation rather than (2).

3 Basic properties of the K-depth

This section contains some of the basic properties of the K-depth. In particular, we discuss the typical behaviour in terms of a law of large numbers in Section 3.1. Sections 3.2 and 3.3 contain extremal cases where the test statistic is close to its maximal or minimal value, respectively.

3.1 Law of large numbers

Let $R_1 := R_1(\theta), \ldots, R_N := R_N(\theta)$ be independent random variables satisfying (1). Then the expectation of the K-depth is given by

$$\begin{aligned}&\mathbb{E}_\theta\left(d_K(R_1(\theta), \ldots, R_N(\theta))\right)\\ &= \frac{1}{\binom{N}{K}} \sum_{1\le n_1<n_2<\ldots<n_K\le N} \left(\left(\frac{1}{2}\right)^K + \left(\frac{1}{2}\right)^K\right) = \left(\frac{1}{2}\right)^{K-1}.\end{aligned} \tag{7}$$

A convergence of the K-depth towards this expectation can be shown by rewriting the summands in (2) using the identity in the next lemma. In order to avoid triple indices, we write $i(j)$ instead of i_j.

Lemma 3.1. *If* $E_{n_1}, \ldots, E_{n_K}$ *are random variables with* $P(E_{n_i} \neq 0) = 1$ *for* $i = 1, \ldots, K$ *and* $K \in \mathbb{N} \setminus \{1\}$ *then we have*

$$\begin{aligned}&\prod_{k=1}^{K} \mathbb{1}\{E_{n_k}(-1)^k > 0\} + \prod_{k=1}^{K} \mathbb{1}\{E_{n_k}(-1)^k < 0\} - \left(\frac{1}{2}\right)^{K-1}\\ &= \frac{1}{2^{K-1}} \sum_{L=1}^{\lfloor \frac{K}{2} \rfloor} \sum_{1\le i(1)<\ldots<i(2L)\le K} \prod_{j=1}^{2L} (-1)^{i(j)} \psi\left(E_{n_{i(j)}}\right) \ P\text{-almost surely,}\end{aligned} \tag{8}$$

where $\psi(x) := \mathbb{1}\{x > 0\} - \mathbb{1}\{x < 0\}$.

Studying the variance of the expression (8) reveals that it converges to zero as $N \to \infty$, which is shown in detail in appendix. Hence Lemma 3.1 leads to a law of large numbers for K-sign depth:

Theorem 3.2. *Let* $K \geq 2$. *If* $R_1(\theta), \ldots, R_N(\theta)$ *are satisfying (1) then*

$$d_K(R_1(\theta), \ldots, R_N(\theta)) \longrightarrow \left(\frac{1}{2}\right)^{K-1}$$

P_θ*-almost surely as* $N \to \infty$.

3.2 K-depth for alternating signs

In this section we study the behaviour of the K-depth of residuals with alternating signs, i.e. of residuals $r_1, \ldots, r_N$ with $\psi(r_n) = -\psi(r_{n+1})$ for $n = 1, \ldots, N-1$. Alternating signs indicate a good fit and the K-depth attains its maximum value in this situation. Therefore it is of interest what exactly this maximum value is. This is given by the following theorem. As usual, we use the convention $\binom{n}{k} = 0$ for $n < k$.

Theorem 3.3. *Suppose* $r_1, \ldots, r_N$ *have alternating signs. Then, for* $2 \leq K \leq N$,

$$d_K(r_1, \ldots, r_N) = \frac{1}{\binom{N}{K}} \left(\binom{\lfloor (N+K)/2 \rfloor}{K} + \binom{\lceil (N+K-2)/2 \rceil}{K} \right).$$

Note that Theorem 3.3 can also be used to determine the size of the index set $\mathcal{A}_{K,B}$ in the block-implementation:

Corollary 3.4. *Let* $B, K \geq 2$ *be integers and let* $\mathcal{A}_{K,B}$ *be as in Lemma 2.2. Then*

$$|\mathcal{A}_{K,B}| = \binom{\lfloor (B+K)/2 \rfloor}{K} + \binom{\lceil (B+K-2)/2 \rceil}{K},$$

where $|\mathcal{A}_{K,B}|$ *denotes the size of* $\mathcal{A}_{K,B}$.

Theorem 3.3 implies that the K-depth of residuals with alternating signs converges to the expected value $(1/2)^{K-1}$ as $N \to \infty$. In conjunction with Corollary 3.4, we may extend this property to the following more general class of alternating vectors:

Definition 3.5. *Let* $M \in \mathbb{N}$ *and let* $r = (r_1, \ldots, r_N)$ *be a vector of residuals. The residuals* $r_1, \ldots, r_N$ *are alternating in blocks of size* M *if* N *is a multiple of* M *and if*

$$q_j(r) = M \quad \text{for all } j = 1, \ldots, B(r),$$

where the number $B(r)$ *of blocks and the size* $q_j(r)$ *of block* j *are defined in Section 2.2. In particular, residuals have alternating signs if they are alternating in blocks of size* 1.

With Corollary 3.4, it is not hard to compute the K-depth of such residuals explicitly:

Lemma 3.6. *Let $M, N \in \mathbb{N}$ with $B := N/M \in \mathbb{N}$. Furthermore, let $\langle x \rangle_J = \prod_{j=0}^{J-1}(x-j)$ for $x \in \mathbb{N}$ and $x \geq J$. If $r_1, \dots, r_N$ are alternating in blocks of size M and if $B \geq K$, then*

(a) $d_K(r_1, \dots, r_N) = \dfrac{\langle \frac{B+K-2}{2} \rangle_{K-1}}{B^{K-1}} \cdot \dfrac{N^K}{\langle N \rangle_K}$ *if $K + B$ is even,*

(b) $d_K(r_1, \dots, r_N) = \dfrac{2\langle \frac{B+K-1}{2} \rangle_K}{B^K} \cdot \dfrac{N^K}{\langle N \rangle_K}$ *if $K + B$ is odd.*

An asymptotic analysis of the K-depth based on Lemma 3.6 reveals that the K-depth test statistic of residuals that alternate in blocks of size M converges to its maximal value:

Theorem 3.7. *Let M be a fixed integer. If the residuals $r_1, \dots, r_N$ are alternating in blocks of size M, then*

$$\lim_{N \to \infty} N \left(d_K(r_1, \dots, r_N) - \left(\frac{1}{2}\right)^{K-1} \right) = \frac{K(K-1)}{2^K}.$$

Remark 3.1.

(a) Theorem 3.7 yields that the maximal value of the test statistic (i.e. the value for residuals with alternating signs) is asymptotically $K(K-1)/2^K$. Since the minimal K-depth is zero, the minimal value of the test statistic is $-N/2^{K-1}$ which diverges as $N \to \infty$. Hence the (asymptotic) distribution of the test statistic $T_K(\theta)$ is bounded from above but unbounded from below. In particular, its distribution is not symmetric.

(b) Since the test statistic converges to its maximal value if the residuals are alternating in blocks of size $M \geq 1$, the (one-sided) K-depth test will not reject the model when such residuals are observed and N is sufficiently large. This can often be desirable in practice where alternating residuals indicate a good fit and a systematic alternation (in blocks of fixed size) can be caused by some vibration behaviour which is difficult to filter out.

(c) If the independence of the residuals is questionable and of additional interest then alternating residuals are indicating dependence. In such situations, the two-sided K-depth test as proposed in Remark 2.2 can be used. Since alternating residuals yield the maximal possible value, the two-sided test will always reject the model when such residuals are observed and N is sufficiently large.

3.3 Behaviour in situations of few sign changes

Residual vectors with only few sign changes usually indicate a bad choice for the modelling parameter, see, e.g., Figure 1 for so-called nonfits in a quadratic regression model. A nonfit is defined as in Rousseeuw and Hubert (1999):

Definition 3.8. *A parameter θ is called a nonfit if there exists another parameter $\tilde{\theta}$ such that $|r_n(\tilde{\theta})| < |r_n(\theta)|$ for all $n = 1, \dots, N$.*

The 2-depth test can struggle rejecting such bad choices since this test, as we will formally show in Section 4.1, is equivalent to the classical sign test. In particular, it does not reject the model if nearly half of the residuals are positive, regardless of how many sign changes the residuals have.

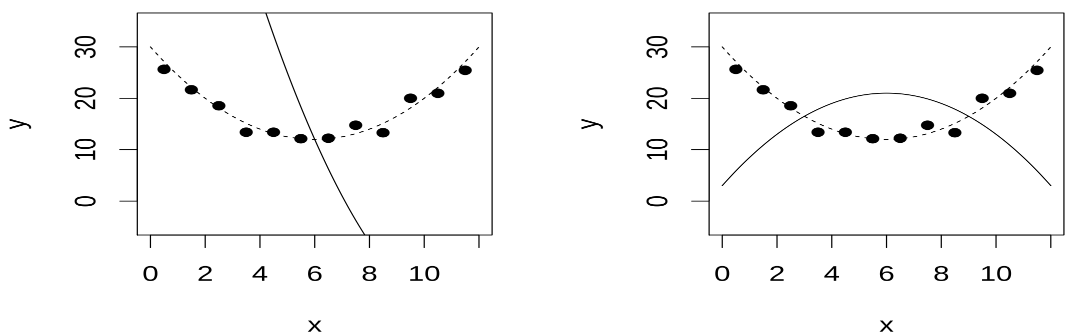

Figure 1: 12 observations generated by $Y_n = g(x_n, \theta^0) + E_n$ with $g(x, \theta^0) = 30 - 6\,x + 0.5\,x^2$ (dashed line, $\theta^0 = (30, -6, 0.5)^\top$) and $E_n \sim \mathcal{N}(0, 1.5^2)$. The solid lines correspond to parameters that yield nonfits with either one or two sign changes: $\theta^1 = (120, -24, 1)^\top$ yielding $g(x, \theta^1) = 120 - 24x + x^2$ on the left hand side and $\theta^2 = (3, 6, -0.5)^\top$ yielding $g(x, \theta^2) = 3 + 6\,x - 0.5\,x^2$ on the right hand side.

K-depth tests with $K \geq 3$ are much more powerful in this regard since they immediately reject models that lead to few sign changes. More precisely, the following lemma is easy to show for residuals vectors $r = (r_1, \ldots, r_N)$ where the number $B(r)$ of blocks (see Section 2.2) is small:

Lemma 3.9. *Let $K \geq 3$. Then $d_K(r_1, \ldots, r_N) = 0$ if and only if $B(r) \leq K - 1$.*

Note that a K-depth of zero is the smallest possible value of the K-depth. Hence this will always lead to a rejection of the null hypothesis by the K-depth test if the sample size is high enough that a rejection at level α is possible. Usually a nonfit of a p-dimensional parameter is expressed by at most $p - 1$ sign changes. Hence a K-depth test with $K = p + 1$ will protect against bad power at nonfits, see also Kustosz et al. (2016b). However, choices $K < p + 1$ can also lead to a good power of the K-depth test at alternatives for which the expected depth of $(1/2)^{K-1}$ is not reached. More precisely, since all α-quantiles of the asymptotic distribution of the K-depth test statistic $T_K(\theta)$ are fixed values greater than $-\infty$, we have the following property for growing sample size N: The strict inequality

$$\lim_{N\to\infty} \sup_{\theta\in\Theta^0} d_K(r_1(\theta), \ldots, r_N(\theta)) < \left(\frac{1}{2}\right)^{K-1} \tag{9}$$

implies $\lim_{N\to\infty} \sup_{\theta\in\Theta^0} T_K(\theta) = -\infty$ so that $H_0 : \theta \in \Theta^0$ is rejected if N is sufficiently large.

Condition (9) is in particular satisfied if the relative number of either the positive or negative residuals is tending to 1. This is often the case when the region of explanatory variables is growing to infinity as N converges to infinity. This was used in Kustosz et al. (2016a) to show the consistency of a test based on simplicial depth for explosive AR(1) regression.

Assuming a bounded, fixed support for the explanatory variables, the relative number of positive/negative residuals usually does not tend to one for alternatives, e.g. in polynomial regression.

However, one at least expects only few sign changes then; see Figure 1 for examples with only one or two sign changes. We therefore end the section with a discussion on the K-depth of residual vectors where the number of blocks/sign changes is bounded.

For the remainder of the section, we will use the alternative representation of the K-depth based on the block-implementation (see Section 2.2). Recall that the K-depth of residuals $r_1, \ldots, r_N$ with B blocks and block sizes $q_1, \ldots, q_B$ is given by

$$d_{K,N,B}(q_1, \ldots, q_B) = \frac{1}{\binom{N}{K}} \sum_{(i_1,\ldots,i_K) \in \mathcal{A}_{K,B}} \prod_{k=1}^{K} q_{i_k}.$$

Although $q_1, \ldots, q_B$ are integers in practice, it will be more convenient in the subsequent analysis to let $q_1, \ldots, q_B$ be positive real numbers. In order to see that the K-depth test always rejects the null hypothesis if B is sufficiently small, we need to consider the input $q_1, \ldots, q_B$ with maximal K-depth. While it is arguably quite intuitive to assume that this maximum is attained at $q_j = N/B$ for all $j = 1, \ldots, N$, a formal proof to determine the maximum is challenging. We therefore state the following conjecture which we only checked for some particular choices of K and B:

Conjecture 3.10. *Let $K \geq 3$, $B \geq K$ and $N \geq B$. Consider the set*

$$\mathcal{M}_{K,N,B} := \arg\max \left\{ d_{K,N,B}(q_1, \ldots, q_B);\ (q_1, \ldots, q_B) \in (0, N)^B, \sum_{b=1}^{B} q_b = N \right\}.$$

Then the following holds:

(a) If $K + B$ is even then

$$\mathcal{M}_{K,N,B} = \left\{ \left(\frac{N}{B}, \ldots, \frac{N}{B} \right) \right\}.$$

(b) If $K + B$ is odd then

$$\mathcal{M}_{K,N,B} = \left\{ \left(\frac{\beta N}{B-1}, \frac{N}{B-1}, \ldots, \frac{N}{B-1}, \frac{(1-\beta)N}{B-1} \right);\ \beta \in (0,1) \right\}.$$

The necessity of a case distinction between $K + B$ even/odd might be a bit surprising at first. But in fact it is not hard to check that the function $d_{K,N,B}$ has the following property:

Lemma 3.11. *Let $K \geq 2$ and $B \geq K$. If $K + B$ is odd then*

$$d_{K,N,B}(q_1, \ldots, q_B) = d_{K,N,B-1}(q_1 + q_B, q_2, \ldots, q_{B-1}).$$

Hence we may assume w.l.o.g. that $K + B$ is even and use Lemma 3.11 to cover the odd case. Before stating the general result, we consider the special cases $B = K$ and $B = K + 1$. In these cases, Conjecture 3.10 is easy to verify since, by definition,

$$d_{K,N,K}(q_1, \ldots, q_K) = \frac{1}{\binom{N}{K}} \prod_{j=1}^{K} q_j,$$

$$d_{K,N,K+1}(q_1, \ldots, q_{K+1}) = \frac{1}{\binom{N}{K}} (q_1 + q_{K+1}) \prod_{j=2}^{K} q_j.$$

In particular, we have the following theorem for the maximal K-depth among all valid block sizes $q_1, \ldots, q_B$. The set of these valid block sizes is denoted by

$$\mathcal{Q}_{N,B} := \left\{ (q_1, \ldots, q_B) \in \mathbb{N}^B;\ \sum_{j=1}^{B} q_j = N \right\}, \quad N, B \in \mathbb{N}. \tag{10}$$

Theorem 3.12. *Let $K \geq 2$, $B \in \{K, K+1\}$ and let $\mathcal{Q}_{N,B}$ be as above. Then*

$$\lim_{N\to\infty} \sup \{ d_{K,N,B}(q_1, \ldots, q_B);\ (q_1, \ldots, q_B) \in \mathcal{Q}_{N,B} \} = \frac{K!}{K^K} \leq \left(\frac{1}{2}\right)^{K-1}, \tag{11}$$

where the inequality in (11) is strict for $K \geq 3$.

For the general case $B \geq K+2$, we will only consider the input $q_1 = \ldots = q_B = N/B$ since this is assumed to yield the maximal depth according to Conjecture 3.10 if $K + B$ is even. Lemma 3.6 yields the following result on the asymptotic K-depth.

Theorem 3.13. *Let $K \geq 2$ and $B \geq K$ be fixed. If $K + B$ is even then*

$$\lim_{N\to\infty} d_{K,N,B}\left(\frac{N}{B}, \ldots, \frac{N}{B}\right) = \frac{\prod_{k=1}^{K-1}\left(\frac{B+K}{2} - k\right)}{B^{K-1}} \leq \left(\frac{1}{2}\right)^{K-1}. \tag{12}$$

The inequality in (12) is strict for $K \geq 3$.

Remark 3.2. If $K + B$ is odd then Lemma 3.11 and Theorem 3.13 yield for all $\beta \in (0,1)$

$$\begin{aligned} &\lim_{N\to\infty} d_{K,N,B}\left(\frac{\beta N}{B-1}, \frac{N}{B-1}, \ldots, \frac{N}{B-1}, \frac{(1-\beta)N}{B-1}\right) \\ &= \frac{1}{2^{K-1}} \frac{\prod_{k=1}^{K-1}(B-1+K-2k)}{(B-1)^{K-1}} \leq \left(\frac{1}{2}\right)^{K-1} \end{aligned} \tag{13}$$

with a strict inequality for $K \geq 3$. Moreover, if we assume that Conjecture 3.10 is true, then (12) and (13) imply for any fixed number B of blocks

$$\begin{aligned} &\lim_{N\to\infty} \sup \{ d_{K,N,B}(q);\ q \in \mathcal{Q}_{N,B} \} \\ &= \frac{1}{2^{K-1}} \frac{\prod_{k=1}^{K-1}(B - \mathbb{1}\{K+B \text{ odd}\} + K - 2k)}{(B - \mathbb{1}\{K+B \text{ odd}\})^{K-1}} \leq \left(\frac{1}{2}\right)^{K-1} \end{aligned}$$

with $\mathcal{Q}_{N,B}$ defined as in (10). Moreover, the inequality above is strict for $K \geq 3$. Hence, $H_0 : \theta \in \Theta^0$ is rejected at an alternative for sufficiently large sample sizes N if the number of blocks in $(r_1(\theta), \ldots, r_N(\theta))$ is uniformly bounded for all $\theta \in \Theta^0$ as $N \to \infty$.

4 Comparison of K-depth tests for different K

A proper choice for K is a crucial aspect to obtain a K-depth test with high power. This section contains some basic observations for the cases $K \leq 6$, in particular in terms of power when only few sign changes are observed. A more profound comparison in applications will be done in Section 5.

As we will see in Section 4.1, the 2-depth test is usually a bad choice since it is equivalent to the classical sign test. This test struggles to reject the null hypothesis at alternatives that lead to a nearly equal amount of positive and negative residuals. K-depth tests with $K \geq 3$ can correctly identify and reject such alternatives as long as the number of sign changes in the residual vector is fairly low. A discussion on the p-values of the K-depth tests, $K = 3, \ldots, 6$, for several different sample sizes can be found in Section 4.2.

4.1 Equivalence of the 2-depth test and the classical sign test

The test statistic of the classical sign test is given by

$$T_{\text{sign}}(\theta) := \frac{N_+(\theta) - N/2}{\sqrt{N}/2} \quad \text{where} \quad N_+(\theta) := \sum_{n=1}^{N} \mathbb{1}\{R_n(\theta) > 0\}$$

denotes the number of residuals with positive signs among $(R_1(\theta), \ldots, R_n(\theta))$. Assuming (1), this test statistic converges in distribution to the standard normal distribution. Hence the classical sign test (in its asymptotic version) is defined via

$$\text{reject } H_0 : \theta \in \Theta^0 \text{ if for all } \theta \in \Theta^0 : T_{\text{sign}}(\theta) < u_{\frac{\alpha}{2}} \text{ or } T_{\text{sign}}(\theta) > u_{1-\frac{\alpha}{2}},$$

where u_α denotes the α-quantile of the standard normal distribution. Equivalently, one can define the classical sign test via

$$\text{reject } H_0 : \theta \in \Theta^0 \text{ if } \inf_{\theta \in \Theta^0} T_{\text{sign}}(\theta)^2 > \chi^2_{1,1-\alpha},$$

where $\chi^2_{1,\alpha}$ is the α-quantile of the χ^2_1 distribution. Note that $T_{\text{sign}}(\theta)^2$ is minimized if $N_+(\theta) = N/2$. Hence the test will not reject the null hypothesis if half of the residuals are positive.

To see the relationship to the 2-depth test, note that a pair of residuals has alternating signs if and only if one of them is positive and the other one is negative. Since we have $N_+(\theta)$ positive and $N - N_+(\theta)$ negative residuals (assuming $R_n(\theta) \neq 0$ P_θ-almost surely for all $n = 1, \ldots, N$), the 2-depth satisfies P_θ-almost surely:

$$d_2(R_1(\theta), \ldots, R_N(\theta)) = \frac{1}{\binom{N}{2}} N_+(\theta)\,(N - N_+(\theta)).$$

The 2-depth can be transformed into $T_{\text{sign}}(\theta)$ by using the identity

$$x(N - x) = -(x - N/2)^2 + N^2/4, \quad x \in \mathbb{R},$$

for $x = N_+(\theta)$. A straightforward calculation based on this identity reveals that the test statistic (3) satisfies for $K = 2$,

$$T_2(\theta) = \frac{N}{2(N-1)} - \frac{N}{2(N-1)} T_{\text{sign}}(\theta)^2 \quad P_\theta\text{-almost surely.}$$

Hence the 2-depth test and the classical sign test are equivalent.

4.2 Comparison of K-depth tests for $K \geq 3$

As we have seen in Section 3.3, K-depth tests with $K \geq 3$ are capable of rejecting nonfits that lead to a small number of sign changes, at least as long as the sample size N is sufficiently large. We will now take a closer look at the performance for small samples sizes up to $N = 160$.

Recall that, according to Conjecture 3.10, we assume that the maximal K-depth of a residual vector $r = (r_1, \ldots, r_N)$ with B blocks is given by

$$\eta_{K,N,B} := \begin{cases} d_{K,N,B}\left(\frac{N}{B}, \ldots, \frac{N}{B}\right), & \text{if } K + B \text{ is even,} \\ d_{K,N,B}\left(\frac{N}{2(B-1)}, \frac{N}{B-1}, \ldots, \frac{N}{B-1}, \frac{N}{2(B-1)}\right), & \text{if } K + B \text{ is odd.} \end{cases}$$

Hence, the test statistic (3) for a residual vector with B blocks can be at most

$$\widetilde{\eta}_{K,N,B} := N\left(\eta_{K,N,B} - \left(\frac{1}{2}\right)^{K-1}\right).$$

Figure 2 contains the p-values when observing a value of $\widetilde{\eta}_{K,N,B}$ for $B = 3, 4, 5, 6$ blocks or 2,3,4,5 sign changes, respectively, i.e. the probabilities

$$P_\theta\left(T_K(R_1(\theta), \ldots, R_N(\theta)) \leq \widetilde{\eta}_{K,N,B}\right)$$

are plotted for samples sizes N between 10 and 160 and $K = 3, 4, 5, 6$. Recall that if a residual vector has B block, i.e. $B - 1$ sign changes, then K-depth tests with $K > B$ will automatically reject the null hypothesis as soon as the sample size is large enough to make a rejection possible for the test. Figure 2 thus only contains K-depth tests with $K \leq 4$ for situations with two sign changes to highlight that the p-value of the 4-depth test indeed becomes 0 if N is sufficiently large. The same applies to the 5-depth test when three sign changes occur. The other two plots (four and five sign changes) do not contain the corresponding 6- and 7-depth tests since their p-values behave similarly.

All four subfigures of Figure 2 indicate that the p-values of all considered K-depth tests are decreasing to zero for growing sample size. They decrease more slowly for $K = 3, 4$ than for $K = 5, 6$, but even the p-value of the 3-depth test reaches 0.1 for a sample size greater than $N = 150$. It is remarkable that the p-values of the K-depth tests with $K = B - 1$ and $K = B$ are always very similar for all $B - 1 = 3, 4, 5$ sign changes we considered. However, this does not hold for $B - 1 = 2$ since the 2-depth test is the classical sign test which always has a p-value of 1 in the case of two blocks of equal size.

5 Applications

The high power of 3-depth tests in the case of two unknown parameters was already shown for explosive AR(1) models, namely in Kustosz et al. (2016a) for linear AR(1)-models given by $Y_n = \theta_0 + \theta_1 Y_{n-1} + E_n$ and in Kustosz et al. (2016b) for nonlinear AR(1)-models given by $Y_n = Y_{n-1} + \theta_1 Y_{n-1}^{\theta_2} + E_n$, see also Falkenau (2016). In particular these results showed for normally distributed errors E_n that 3-depth tests possess similarly high power compared to classical tests based on least squares.

Here we will compare the 3-depth test, the 4-depth test, the classical sign test and the classical t-test in a quadratic regression model with three unknown parameters. Other results for a nonlinear

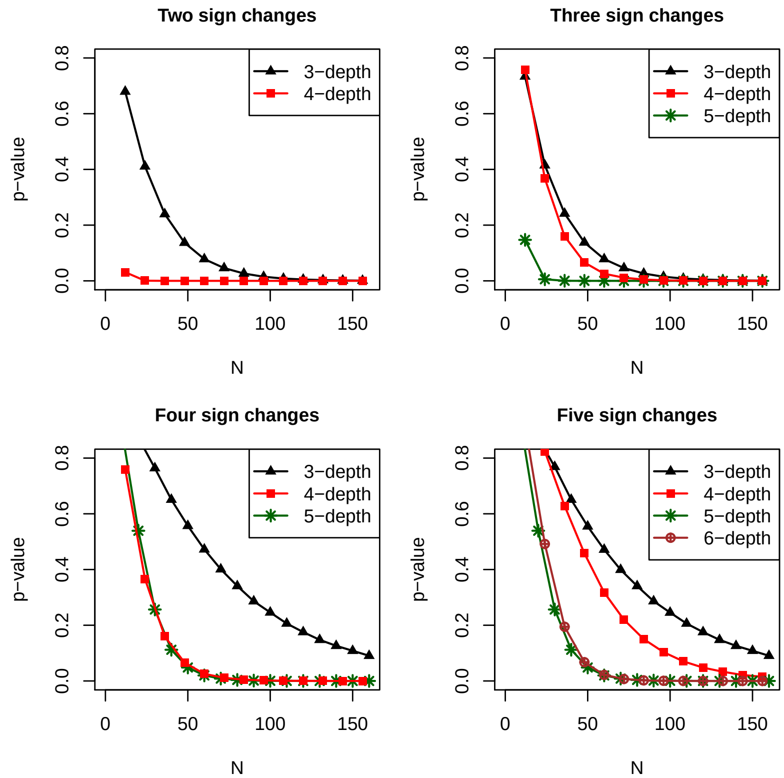

Figure 2: Simulated p-values of K-depth tests, $K = 3, 4, 5, 6$, for two to five sign changes (top left to bottom right) at different sample sizes N.

AR(1)-model and an explosive AR(2)-model, each with three unknown parameters, can be found in the supplementary material.

Note that the 3- and 4-depth tests require the (α-quantile of the) distribution of their test statistics. For all applications, we used the exact distribution for the small sample size $N = 12$ and a simulated distribution for the large sample size $N = 96$. The simulated distribution of both depth tests was obtained via 10 000 i.i.d. samples from the distribution. The power at each alternative is computed based on 100 samples. Note that we also considered an increased number of 500 repetitions for some cases which, however, did not lead to any visible changes in the plots. Hence we decided to stick to only 100 repetitions to speed up the computation.

Quadratic regression. In the quadratic regression model given by

$$Y_n = \theta_0 + \theta_1 \, x_n + \theta_2 \, x_n^2 + E_n, \;\; n = 1, \ldots, N, \;\; \theta = (\theta_0, \theta_1, \theta_2)^\top,$$

we consider the problem of testing the null hypothesis $H_0 : \theta = (1, 0, 1)^\top$ with a test with level $\alpha = 0.05$ and samples sizes $N = 12$ and $N = 96$. For each simulation, a 41×41 grid of alternatives is used. The explanatory variables $x_1, \ldots, x_N$ are chosen to be equidistant elements from the interval $[-6, 6]$.

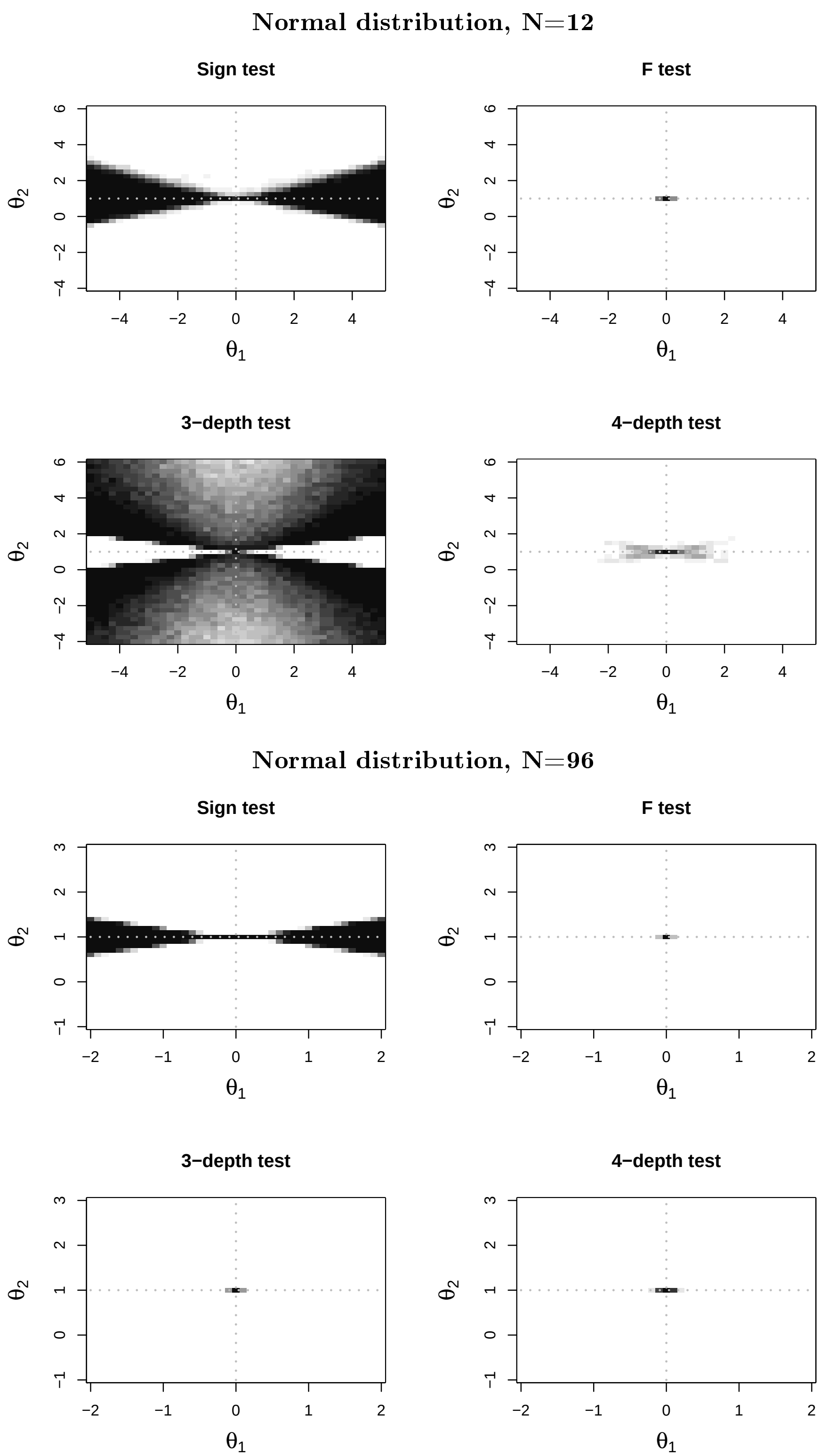

Figure 3: Simulated power of the sign test, the F test, the 3-depth test, and the 4-depth test for normally distributed errors for sample size $N = 12$ (upper part) and $N = 96$ (lower part) where the component θ_0 is fixed to 1 (20 grey levels were used, where black corresponds to $[0, 0.05]$ and white to $(0.95, 1]$).

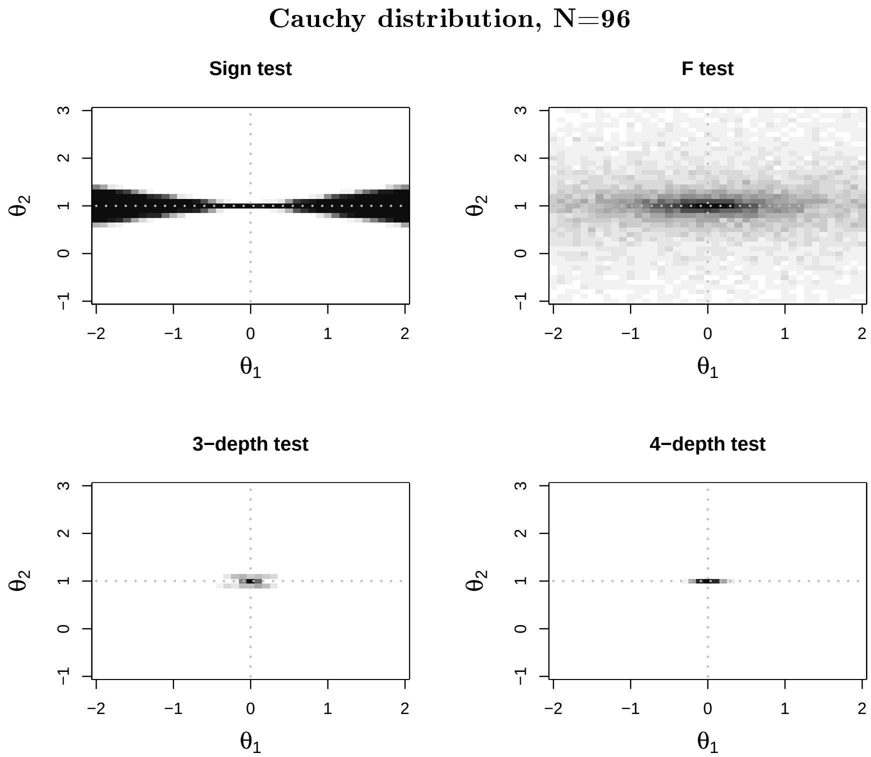

Figure 4: Simulated power of the sign test, the F test, the 3-depth test, and the 4-depth test for errors with Cauchy distributed for sample size $N = 96$ where the component θ_0 is fixed to 1 (20 grey levels were used, where black corresponds to $[0, 0.05]$ and white to $(0.95, 1]$).

Figure 3 shows the simulated power of the sign test, the F test, the 3-depth test, and the 4-depth test for the case where E_n has a standard normal distribution and the component θ_0 is fixed to 1. The parameters θ_1 and θ_2 of the null hypothesis are given by the intersection of the two dotted lines. The results for $N = 12$ are shown in the upper part of this figure. The sign test and the 3-depth test both possess an unbounded area of power below $\alpha = 0.05$, whereas only the F test and 4-sign test perform fairly well. The main reason for the bad performance of the 3-depth test (and the fairly good performance of the 4-depth test) is that for small N, the K-sign test mostly only rejects the hypothesis if the residual vector has less than K blocks (i.e. less than $K-1$ sign changes). More specifically, the p-value of the 3-depth for $N = 12$ and a residual vector with 2 sign changes can be up to 0.758 when considering the worst case scenario from Section 4.2. Now a quadratic model corresponding to $\theta_2 x^2 + \theta_1 x + 1$ and the true model $x^2 + 1$ intersect in $x = 0$ and in $x = \theta_1/(1 - \theta_2)$ if $\theta_2 \neq 1$. Since we consider explanatory variables in $[-6, 6]$, a residual vector for the parameter $(1, \theta_1, \theta_2)$ thus tend to have either one or two sign changes depending on whether $|\theta_1/(1 - \theta_2)| \leq 6$ or not. The sign test performs poorly when only one sign change occurs (since it typically occurs at a position n with $x_n \approx 0$ and hence nearly half of the residuals are positive), whereas the 3-depth test performs poorly if two sign changes occur. The lower part of Figure 3 contains the corresponding power levels for $N = 96$. In this case N is sufficiently large such that residual vectors with only few sign changes will usually lead to a rejection by the K-depth tests. For example, the p-value of the 3-depth test for vectors with two sign changes is at most 0.014 and thus will always lead to a rejection. In summary, Figure 3 indicates that both 3- and 4-depth test can compete with the F test in quadratic regression models where N is sufficiently large whereas the sign test always struggles at alternatives which lead to residual vectors where nearly half of the signs are positive.

Figure 4 shows what happens when the normal distribution for the errors is replaced by the Cauchy distribution. Then the power of the F test becomes very bad while the power functions of the 3-depth test and the 4-depth test change only slightly. Hence the 3-depth test and the 4-depth test are much more robust against outliers than the F test. Furthermore, the depth tests are more powerful than the sign test.

Similar results where either θ_1 or θ_2 is fixed can be found in the supplementary material.

Supplementary material. The further simulation results and the R-code can be found under `https://www.statistik.tu-dortmund.de/2273.html`.

6 Discussion and outlook

K-sign depths can be used to define robust tests which we refer to as K-depth tests. While the parameter choice $K = 2$ essentially leads to the classical sign test and thus has several limitations in rejecting nonfits, K-depth tests for $K \geq 3$ are fairly powerful. They can even outperform classical approaches such as the t-test, in particular in the presence of outliers. Note that those tests are not very well-suited for small sample sizes and models where the number of sign changes in the residual vector is likely to exceed $K - 1$ at alternatives. However, the K-depth tests perform very well in our examples once the sample size is sufficiently large. Moreover, it seems like there is no advantage in using the 4-depth test instead of the 3-depth test once the sample size is large. Note that this observation may not be true in general and requires further research.

This paper is mainly focused on the one-sided version of K-depth test to detect shifts in the medians of the residuals. A two-sided version of the K-depth test can also detect dependence structures within the residuals and may be useful for stationary AR-models and other stationary processes. Once again, further research is necessary to compare the two-sided K-depth test with other approaches when testing simultaneously whether residuals are independent and have medians equal to zero.

To reduce the runtime of $\Theta(N^K)$ of the definition of the K-depth, a faster block implementation is presented. A linear runtime of an asymptotically equivalent form can be obtained by the derivation of the asymptotic distribution of the K-depth for $K \geq 3$. However, the lengthy derivation will be published elsewhere.

Finally note that the K-depth depends on the order of the residuals, hence the chosen order is a crucial aspect to obtain powerful tests. The canonical order of $\mathbb{R}$ seems to yield good results for regression models with real-valued explanatory variables. For $\mathbb{R}^q$ with $q > 1$, several data driven orderings of the q-dimensional explanatory variables exists which lead to powerful tests as well. A detailed treatment is currently in preparation.

Acknowledgments

The authors gratefully acknowledge support from the Collaborative Research Center "Statistical Modeling of Nonlinear Dynamic Processes" (SFB 823, B5) of the German Research Foundation (DFG).

Appendix

Proof of Lemma 3.1. In order to simplify the notation, we assume $(n_1, \dots, n_K) = (1, \dots, K)$. Note for $x \neq 0$

$$\mathbb{1}\{x > 0\} = \frac{1}{2}\left(\psi\left(x\right) + 1\right), \qquad \mathbb{1}\{x < 0\} = \frac{1}{2}\left(-\psi\left(x\right) + 1\right).$$

It is straightforward to check $\prod_{i=1}^{K}(a_i+1) = \sum_{\ell=1}^{K} \sum_{1\le i(1)<\ldots<i(\ell)\le K} \prod_{j=1}^{\ell} a_{i(j)} + 1$ for arbitrary $a_1, \dots, a_K$. This implies P-almost surely

$$\prod_{k=1}^{K} \mathbb{1}\{E_k(-1)^k > 0\} = \frac{1}{2^K} \prod_{k=1}^{K} \left((-1)^k \psi\left(E_k\right) + 1\right)$$
$$= \frac{1}{2^K}\left(\sum_{\ell=1}^{K} \sum_{1\le i(1)<\ldots<i(\ell)\le K} (-1)^{i(1)+\cdots+i(\ell)} \prod_{j=1}^{\ell} \psi\left(E_{i(j)}\right) + 1\right)$$

Similarly

$$\prod_{k=1}^{K} \mathbb{1}\{E_k(-1)^k < 0\}$$
$$= \frac{1}{2^K}\left(\sum_{\ell=1}^{K} \sum_{1\le i(1)<\ldots<i(\ell)\le K} (-1)^{i(1)+\cdots+i(\ell)+\ell} \prod_{j=1}^{\ell} \psi\left(E_{i(j)}\right) + 1\right)$$
$$= \frac{1}{2^K}\left(\sum_{\substack{\ell=1,\ldots,K \\ \ell \text{ even}}} \sum_{1\le i(1)<\ldots<i(\ell)\le K} (-1)^{i(1)+\cdots+i(\ell)} \prod_{j=1}^{\ell} \psi\left(E_{i(j)}\right) + 1\right)$$
$$- \frac{1}{2^K} \sum_{\substack{\ell=1,\ldots,K \\ \ell \text{ odd}}} \sum_{1\le i(1)<\ldots<i(\ell)\le K} (-1)^{i(1)+\cdots+i(\ell)} \prod_{j=1}^{\ell} \psi\left(E_{i(j)}\right).$$

Therefore

$$
\begin{aligned}
&\prod_{k=1}^{K} \mathbb{1}\{E_k(-1)^k > 0\} + \prod_{k=1}^{K} \mathbb{1}\{E_k(-1)^k < 0\} \\
&= \frac{1}{2^{K-1}} \left(\sum_{\substack{\ell=1,\ldots,K \\ \ell \text{ even}}} \sum_{1 \leq i(1) < \ldots < i(\ell) \leq K} (-1)^{i(1)+\cdots+i(\ell)} \prod_{j=1}^{\ell} \psi\left(E_{i(j)}\right) + 1 \right) \\
&= \left(\frac{1}{2}\right)^{K-1} + \frac{1}{2^{K-1}} \sum_{L=1}^{\lfloor \frac{K}{2} \rfloor} \sum_{1 \leq i(1) < \ldots < i(2L) \leq K} (-1)^{i(1)+\cdots+i(2L)} \prod_{j=1}^{2L} \psi(E_{i(j)})
\end{aligned}
$$

and the assertion follows. □

Proof of Theorem 3.2. Set $R_n = R_n(\theta)$. Lemma 3.1 yields

$$
\begin{aligned}
&d_K(R_1, \ldots, R_N) - \left(\frac{1}{2}\right)^{K-1} \\
&= \frac{1}{\binom{N}{K}} \sum_{1 \leq n_1 < n_2 < \ldots < n_K \leq N} \frac{1}{2^{K-1}} \sum_{L=1}^{\lfloor \frac{K}{2} \rfloor} \sum_{1 \leq i(1) < \ldots < i(2L) \leq K} \prod_{j=1}^{2L} (-1)^{i(j)} \psi\left(R_{n_{i(j)}}\right)
\end{aligned}
$$

with $\psi(x) := \mathbb{1}\{x > 0\} - \mathbb{1}\{x < 0\}$. Set

$$
v := \sum_{L=1}^{\lfloor \frac{K}{2} \rfloor} \sum_{1 \leq i(1) < \ldots < i(2L) \leq K} 1
$$

for the number of summands in the representation of K alternating signs given by Lemma 3.1. This number depends only on K and not on N. First of all, we show that each of these v summands is converging in probability to zero.

To this end, let $L = 1, \ldots, \lfloor \frac{K}{2} \rfloor$ and $1 \leq i(1) < \ldots < i(2L) \leq K$ be arbitrary. We consider the summand multiplied by the factor 2^{K-1}. Because $\mathbb{E}_\theta\left(\psi\left(R_n\right)\right) = 0$ and $R_1, \ldots, R_N$ are independent, we get at once for this summand

$$
\mathbb{E}_\theta \left(\frac{1}{\binom{N}{K}} \sum_{1 \leq n_1 < n_2 < \ldots < n_K \leq N} (-1)^{i(1)+\ldots+i(2L)} \prod_{j=1}^{2L} \psi\left(R_{n_{i(j)}}\right) \right) = 0.
$$

Moreover, $\psi\left(R_n\right)^2 = 1$ P_θ-almost surely implies

$$
\mathbb{E}_\theta \left(\prod_{j=1}^{2L} \psi\left(R_{n_{i(j)}}\right) \prod_{j=1}^{2L} \psi\left(R_{\tilde{n}_{i(j)}}\right) \right) = \begin{cases} 1, & \text{if } n_{i(j)} = \tilde{n}_{i(j)} \text{ for } j = 1, \ldots, 2L, \\ 0, & \text{else.} \end{cases}
$$

Hence

$$\begin{aligned}
&\operatorname{var}_\theta\left(\frac{1}{\binom{N}{K}}\sum_{1\le n_1<\ldots<n_K\le N}(-1)^{i(1)+\ldots+i(2L)}\prod_{j=1}^{2L}\psi\left(R_{n_{i(j)}}\right)\right)\\
&= \frac{1}{\binom{N}{K}^2}\sum_{1\le n_1<\ldots<n_K\le N}\;\sum_{1\le \tilde n_1<\ldots<\tilde n_K\le N}\mathbb{E}_\theta\left(\prod_{j=1}^{2L}\psi\left(R_{n_{i(j)}}\right)\prod_{j=1}^{2L}\psi\left(R_{\tilde n_{i(j)}}\right)\right)\\
&= \frac{1}{\binom{N}{K}^2}\sum_{\substack{1\le n_1<\ldots<n_K\le N,\ 1\le \tilde n_1<\ldots<\tilde n_K\le N\\ n_{i(j)}=\tilde n_{i(j)}\text{ for } j=1,\ldots,2L}} 1\\
&\le \frac{1}{\binom{N}{K}^2}\sum_{1\le n_1<\ldots<n_{2L}\le N}\;\sum_{n_{2L+1},\ldots,n_K\in\{1,\ldots,N\}}\;\sum_{\tilde n_{2L+1},\ldots,\tilde n_K\in\{1,\ldots,N\}} 1\\
&= \frac{\binom{N}{2L}N^{K-2L}N^{K-2L}}{\binom{N}{K}^2}\le\frac{(K!)^2}{(2L)!}\frac{N^{2L+2K-4L}}{(N-(K+1))^{2K}}\\
&= \frac{(K!)^2}{(2L)!}\,\frac{1}{N^{2L}}\,\frac{1}{\left(1-\frac{K+1}{N}\right)^{2K}}\longrightarrow 0
\end{aligned}$$

for $N\to\infty$ so that Chebyshev inequality provides the convergence in probability to zero. Furthermore, the convergence in probability is sufficiently quick of order $O(N^{-2L})$ so that the Borel-Cantelli lemma implies the convergence to zero P_θ-almost surely. □

Proof of Theorem 3.3. Let $r_1,\ldots,r_N$ be residuals with alternating signs. First note that $r_{i_1},\ldots,r_{i_K}$ are alternating if and only if $(i_1,\ldots,i_K)\in\mathcal{A}_{K,N}$ with $\mathcal{A}_{K,N}$ defined as in Lemma 2.2. Hence

$$d_K(r_1,\ldots,r_N)=\frac{|\mathcal{A}_{K,N}|}{\binom{N}{K}}$$

where $|\mathcal{A}_{K,N}|$ denotes the size of $\mathcal{A}_{K,N}$. Thus it only remains to determine this size.

In the subsequent analysis, we write $\mathbb{O}$ for the set of all odd positive integers, i.e. $\mathbb{O}=2\mathbb{N}_0+1$. For a vector $(i_1,\ldots,i_K)$ let $\Delta_1:=i_1$ and $\Delta_k:=i_k-i_{k-1}$ for $k=2,\ldots,K$. Note that $(i_1,\ldots,i_K)\in\mathcal{A}_{K,N}$ if and only if $(\Delta_1,\ldots,\Delta_K)$ is part of the set

$$\mathcal{D}_{K,N}:=\left\{(\Delta_1,\ldots,\Delta_K)\in\mathbb{N}\times\mathbb{O}^{K-1};\ \sum_{k=1}^{K}\Delta_k\le N\right\}.$$

Hence $|\mathcal{A}_{K,N}|=|\mathcal{D}_{K,N}|$. In order to remove the additional condition $\Delta_k\in\mathbb{O}$ for $k\ge 2$, we will use the transformation $\widetilde{\Delta}_k=(\Delta_k+1)/2$. Since this transformation for $k=1$ only provides an integer if Δ_1 is odd, we additionally split the set into the two parts

$$\begin{aligned}
\mathcal{D}^-_{K,N}&:=\{(\Delta_1,\ldots,\Delta_K)\in\mathcal{D}_{K,N};\ \Delta_1\in\mathbb{O}\},\\
\mathcal{D}^+_{K,N}&:=\{(\Delta_1,\ldots,\Delta_K)\in\mathcal{D}_{K,N};\ \Delta_1\notin\mathbb{O}\}.
\end{aligned}$$

The elements of $\mathcal{D}^-_{K,N}$ can be counted by noting that $(\Delta_1, \ldots, \Delta_K) \in \mathcal{D}^-_{K,N}$ if and only if

$$\left(\widetilde{\Delta}_1, \ldots, \widetilde{\Delta}_K\right) \in \widetilde{\mathcal{D}}^-_{K,N} := \left\{(n_1, \ldots, n_K) \in \mathbb{N}^K;\ \sum_{k=1}^{K} n_k \leq \frac{N+K}{2}\right\}$$

with $\widetilde{\Delta}_k = (\Delta_k + 1)/2$ for $k = 1, \ldots, K$. Similarly, $(\Delta_1, \ldots, \Delta_K) \in \mathcal{D}^+_{K,N}$ if and only if

$$\left(\frac{\Delta_1}{2}, \widetilde{\Delta}_2, \ldots, \widetilde{\Delta}_K\right) \in \widetilde{\mathcal{D}}^+_{K,N} := \left\{(n_1, \ldots, n_K) \in \mathbb{N}^K;\ \sum_{k=1}^{K} n_k \leq \frac{N+K-1}{2}\right\}$$

with $\widetilde{\Delta}_k$ as above. In summary, the (bijective) transformations discussed above yield

$$|\mathcal{A}_{K,N}| = \left|\widetilde{\mathcal{D}}^-_{K,N}\right| + \left|\widetilde{\mathcal{D}}^+_{K,N}\right|. \tag{14}$$

The sizes of the remaining sets can easily be determined by noting that each element $(n_1, \ldots, n_K)$ in $\widetilde{\mathcal{D}}^-_{K,N}$ corresponds to a K-element subset $\{m_1, \ldots, m_K\}$ of the set $\{1, 2, \ldots, \lfloor (N+K)/2 \rfloor\}$ by letting

$$m_k := \sum_{i=1}^{k} n_i \quad \text{for } k = 1, \ldots, K.$$

Hence

$$\left|\widetilde{\mathcal{D}}^-_{K,N}\right| = \binom{\lfloor (N+K)/2 \rfloor}{K}.$$

Essentially the same arguments yield

$$\left|\widetilde{\mathcal{D}}^+_{K,N}\right| = \binom{\lfloor (N+K-1)/2 \rfloor}{K}.$$

The assertion follows after rewriting $\lfloor (N+K-1)/2 \rfloor = \lceil (N+K-2)/2 \rceil$ and by plugging the sizes of the sets back into (14). □

Proof of Lemma 3.6. First note that if $(r_1, \ldots, r_N)$ consists of B blocks and each block has size $M = N/B$, then Lemma 2.2 and Corollary 3.4 yield

$$\begin{aligned} d_K(r_1, \ldots, r_N) &= d_{K,N,B}\left(\frac{N}{B}, \ldots, \frac{N}{B}\right) \\ &= \frac{\left(\frac{N}{B}\right)^K}{\binom{N}{K}} \left(\binom{\lfloor (B+K)/2 \rfloor}{K} + \binom{\lceil (B+K)/2 \rceil - 1}{K}\right). \end{aligned}$$

Since binomial coefficients satisfy $\displaystyle\binom{x}{K} = \frac{\langle x \rangle_K}{K!}$ for $x \geq K$, this can be simplified to

$$d_K(r_1, \ldots, r_N) = \frac{N^K}{B^K \langle N \rangle_K} \left(\langle \lfloor (B+K)/2 \rfloor \rangle_K + \langle \lceil (B+K)/2 \rceil - 1 \rangle_K\right). \tag{15}$$

If $K + B$ is odd, the assertion follows since $\lfloor (B+K)/2 \rfloor = (B+K-1)/2 = \lceil (B+K)/2 \rceil - 1$. It only remains to consider $K + B$ even. For this case, let $x = (B+K)/2$. Then

$$\begin{aligned}\langle \lfloor x \rfloor \rangle_K + \langle \lceil x \rceil - 1 \rangle_K &= x \langle x-1 \rangle_{K-1} + \langle x-1 \rangle_{K-1}(x-K) \\ &= (2x-K)\langle x-1 \rangle_{K-1}.\end{aligned}$$

Since $2x - K = B$, the assertion follows after plugging this equality back into (15). □

Proof of Theorem 3.7. The proof is based on the formula given in Lemma 3.6. Let $B = N/M$ be the number of blocks and recall that M is fixed and thus $B = \Theta(N)$. The key observation to derive the asymptotic value of the test statistic for residuals which alternate in blocks of size M is the following asymptotic expansion: For any fixed a, J and as $x \to \infty$,

$$\langle x+a \rangle_J = x^J + J\left(a - \frac{J-1}{2}\right) x^{J-1} + O(x^{J-2}). \tag{16}$$

This equality is based on expanding the product in the definition of the falling factorial:

$$\langle x+a \rangle_J = \prod_{j=0}^{J-1} (x+a-j) = x^J + \sum_{j=0}^{J-1} (a-j) x^{J-1} + O(x^{J-2}),$$

which yields (16) using the well-known formula $\sum_{j=0}^{J-1} j = J(J-1)/2$. Hence, Lemma 3.6(a) and (16) with $x = B/2$, $a = (K-2)/2$, $J = K-1$ yield for even $K+B$ that

$$d_K(r_1, \ldots, r_N) = \frac{\langle \frac{B+K-2}{2} \rangle_{K-1}}{B^{K-1}} \cdot \frac{N^K}{\langle N \rangle_K} = \left(\left(\frac{1}{2}\right)^{K-1} + O(N^{-2}) \right) \frac{N^K}{\langle N \rangle_K}. \tag{17}$$

Applying (16) for $x = N$, $a = 0$ and $J = K$ yields

$$\frac{N^K}{\langle N \rangle_K} = \frac{1}{1 - \frac{K(K-1)}{2N} + O(N^{-2})} = 1 + \frac{K(K-1)}{2N} + O(N^{-2}),$$

where the second equality holds since $1/(1-x) = \sum_{j=0}^{\infty} x^j = 1 + x + O(x^2)$ as $x \to 0$. Plugging this asymptotic expansion back into (17) yields for even $K + B$ that

$$\begin{aligned} d_K(r_1, \ldots, r_N) &= \left(\left(\frac{1}{2}\right)^{K-1} + O(N^{-2}) \right) \left(1 + \frac{K(K-1)}{2N} + O(N^{-2}) \right) \\ &= \left(\frac{1}{2}\right)^{K-1} + \left(\frac{1}{2}\right)^{K-1} \frac{K(K-1)}{2N} + O(N^{-2}). \end{aligned}$$

The case that $K + B$ is odd can be treated in a similar fashion and leads to the same asymptotic expansion. Hence the K-depth of $r_1, \ldots, r_N$ satisfies

$$N \cdot \left(d_K(r_1, \ldots, r_N) - \left(\frac{1}{2}\right)^{K-1} \right) = \frac{K(K-1)}{2^K} + O(N^{-1})$$

and the assertion follows by taking the limit $N \to \infty$. □

Proof of Lemma 3.11. For $x \in \mathbb{R}$ and $w = (w_1, \ldots, w_J) \in \mathbb{R}^J$ let $(x, w) = (x, w_1, \ldots, w_J)$ and let $(w, x) = (w_1, \ldots, w_J, x)$. Recall the definition of $\mathcal{A}_{K,B}$ and $d_{K,N,B}(q_1, \ldots, q_K)$ in Lemma 2.2. The key observations to prove Lemma 3.11 are the following: If $K + B$ is odd then, for every $i \in \{2, \ldots, B-1\}^{K-1}$,

(a) $(1, i) \in \mathcal{A}_{K,B}$ if and only if $(i, B) \in \mathcal{A}_{K,B}$,

(b) there is no vector $j \in \{2, \ldots, B-1\}^{K-2}$ with $(1, j, B) \in \mathcal{A}_{K,B}$.

Both (a) and (b) are not hard to check, details are given at the end of the proof. Based on these properties, we can split the sum in $d_{K,N,B}(q_1, \ldots, q_K)$ in the following way: Let

$$\mathcal{B}_{K,B} = \left\{ i \in \{2, \ldots, B-1\}^{K-1};\ (1, i) \in \mathcal{A}_{K,B} \right\},$$
$$\mathcal{C}_{K,B} = \mathcal{A}_{K,B} \cap \{2, \ldots, B-1\}^K.$$

We may now split $\mathcal{A}_{K,B}$ into three parts: The first one contains vectors $(v_1, \ldots, v_K)$ in $\mathcal{A}_{K,B}$ with $v_1 = 1$, the second one contains vectors with $v_K = B$ and the third part contains vectors with $v_1 \neq 1$ and $v_K \neq B$ (vectors with $v_1 =$ and $v_K = B$ are impossible according to (b)). Then (a) implies $\mathcal{A}_{K,B} = (\{1\} \times \mathcal{B}_{K,B}) \cup (\mathcal{B}_{K,B} \times \{B\}) \cup \mathcal{C}_{K,B}$. Hence

$$\sum_{(i_1,\ldots,i_K) \in \mathcal{A}_{K,B}} \prod_{k=1}^{K} q_{i_k} = (q_1 + q_B) \sum_{(i_1,\ldots,i_{K-1}) \in \mathcal{B}_{K,B}} \prod_{k=1}^{K-1} q_{i_k} + \sum_{(i_1,\ldots,i_K) \in \mathcal{C}_{K,B}} \prod_{k=1}^{K} q_{i_k}.$$

Furthermore, note that $\mathcal{A}_{K,B-1} = (\{1\} \times \mathcal{B}_{K,B}) \cup \mathcal{C}_{K,B}$ once again by splitting the set into two parts based to whether $v_1 = 1$ or not. In particular, if $\widetilde{q}_1 = q_1 + q_B$ and $\widetilde{q}_j = q_j$ for $j = 2, \ldots, B-1$, then

$$\sum_{(i_1,\ldots,i_K) \in \mathcal{A}_{K,B-1}} \prod_{k=1}^{K} \widetilde{q}_{i_k} = \widetilde{q}_1 \sum_{(i_1,\ldots,i_{K-1}) \in \mathcal{B}_{K,B}} \prod_{k=1}^{K-1} \widetilde{q}_{i_k} + \sum_{(i_1,\ldots,i_K) \in \mathcal{C}_{K,B}} \prod_{k=1}^{K} \widetilde{q}_{i_k}$$

Hence $d_{K,N,B}(q_1, \ldots, q_B) = d_{K,N,B-1}(\widetilde{q}_1, \ldots, \widetilde{q}_{B-1})$, which is the assertion.
Proof of (a) and (b). For simplicity, we will subsequently assume that K is odd and B is even. The other case can be treated similarly. For (a) note that $(1, i) \in \mathcal{A}_{K,B}$ requires $i = (i_1, \ldots, i_{K-1})$ to start with an even index i_1 and continue alternating between odd and even in the subsequent indices. Since the length $K-1$ of i is even, the last index i_{K-1} of the vector has to be odd. Since B is even, this means that i_{K-1} and B indeed alternate between odd and even. Hence $(i, B) \in \mathcal{A}_{K,B}$. Similarly, $(i, B) \in \mathcal{A}_{K,B}$ requires i_{K-1} to be odd and subsequent indices in the vector to alternate between odd/even. Hence i_1 has to be even and thus $(1, i) \in \mathcal{A}_{K,B}$. For part (b) assume for the sake of contradiction that $(1, j, B) \in \mathcal{A}_{K,B}$ for a vector $j = (j_1, \ldots, j_{K-2}) \in \{2, \ldots, B-1\}^{K-2}$. Since 1 is odd, this in particular means that j_1 is even. Since $K-2$ is odd, j_1 and j_{K-2} have the same parity in a vector j with entries that alternate between even/odd. Hence j_{K-2} is also even. However, since B is even, j_{K-2} has to be odd in order to have $(1, j, B) \in \mathcal{A}_{K,B}$, which leads to a contradiction. □

Before proving Theorem 3.12, we start with a Lemma that yields the inequalities in Theorem 3.12 and Theorem 3.13.

Lemma 6.1. *Let K, B be integers with $B \geq K \geq 2$. Then*

$$\frac{\prod_{k=1}^{K-1}\left(\frac{B+K}{2}-k\right)}{B^{K-1}} \leq \left(\frac{1}{2}\right)^{K-1}$$

with equality if and only if $K = 2$.

Proof of Lemma 6.1. First note that by rearranging the order of the product one obtains

$$\prod_{k=1}^{K-1}\left(\frac{B+K}{2}-k\right) = \varepsilon_{K,B} \prod_{k=1}^{\lfloor (K-1)/2 \rfloor}\left(\frac{B+K}{2}-k\right)\left(\frac{B+K}{2}-(K-k)\right), \tag{18}$$

$$\text{with} \quad \varepsilon_{K,B} = \begin{cases} 1, & \text{if } K \text{ is odd,} \\ B/2, & \text{if } K \text{ is even.} \end{cases}$$

Next note that the quadratic function $g(x) = ((B+K)/2 - x)((B-K)/2 + x)$ has a unique global maximum at $x = K/2$ and that $g(K/2) = B^2/4$. Hence

$$\prod_{k=1}^{\lfloor (K-1)/2 \rfloor}\left(\frac{B+K}{2}-k\right)\left(\frac{B-K}{2}+k\right) \leq \left(\frac{B^2}{4}\right)^{\lfloor (K-1)/2 \rfloor}$$

in which the inequality is strict if there is at least one factor with $k \neq K/2$, i.e. if $K \geq 3$. In combination with (18), this upper bound yields

$$\frac{\prod_{k=1}^{K-1}\left(\frac{B+K}{2}-k\right)}{B^{K-1}} \leq \frac{\varepsilon_{K,B}}{B^{K-1}}\left(\frac{B^2}{4}\right)^{\lfloor (K-1)/2 \rfloor} = \left(\frac{1}{2}\right)^{K-1}$$

where the last equality can easily be checked by a case distinction between K even/odd. The assertion follows since this inequality is strict for $K \geq 3$. □

Proof of Theorem 3.12. We first consider the case $K = B$, i.e. the aim is to compute the maximum of the function

$$(q_1, \ldots, q_K) \mapsto d_{K,K}(q_1, \ldots, q_K) = \frac{1}{\binom{N}{K}} \prod_{k=1}^{K} q_k$$

under the side condition $(q_1, \ldots, q_K) \in \mathcal{Q}_{N,K}$, i.e. $q_1, \ldots, q_K \in \mathbb{N}$ and $\sum_{k=1}^{K} q_k = N$. When disregarding the condition $q_1, \ldots, q_K \in \mathbb{N}$, this can easily be done, e.g., by using Lagrange multipliers (considering the function $\ln(d_{K,K}(\cdot))$ instead of $d_{K,K}(\cdot)$ simplifies the calculations), which reveals a global maximum at

$$q_1 = \ldots = q_K = \frac{N}{K}.$$

Hence,

$$\begin{aligned} &\sup\left\{d_{K,K}(q_1, \ldots, q_K);\ (q_1, \ldots, q_K) \in \mathcal{Q}_{N,K}\right\} \\ \leq\, &d_{K,K}\left(\frac{N}{K}, \ldots, \frac{N}{K}\right) = \frac{1}{\binom{N}{K}}\left(\frac{N}{K}\right)^{K} \end{aligned}$$

with equality if $N/K \in \mathbb{N}$. Thus the limit values of the maximal depth of residual vectors with K blocks is given by

$$\lim_{N\to\infty} \frac{1}{\binom{N}{K}} \left(\frac{N}{K}\right)^K = \frac{K!}{K^K}.$$

The case $B = K + 1$ can be treated in a similar fashion or can be deduced from $B = K$ and Lemma 3.11. In particular, the maximal value is attained at $q_1 + q_{K+1} = q_2 = \ldots = q_K$ and its limit value remains $K!/K^K$. The remaining inequality

$$\frac{K!}{K^K} < \left(\frac{1}{2}\right)^{K-1} \quad \text{for all } K \geq 3$$

follows from Lemma 6.1 with $B = K$. Hence the assertion follows. □

Proof of Theorem 3.13. The identity for the limit of the test statistic follows from Lemma 3.6 since $N^K/\langle N\rangle_K \to 1$ for fixed K as $N \to \infty$. The inequality in (12) follows from Lemma 6.1. □

References

Agostinelli, C. and Romanazzi, M. (2011). Local depth. *Journal of Statistical Planning and Inference*, 141:817–830.

Claeskens, G., Hubert, M., Slaets, L., and Vakili, K. (2014). Multivariate functional halfspace depth. *Journal of the American Statistical Association*, 109:411–423.

Cuesta-Albertos, J. A., Febrero-Bande, M., and Oviedo de la Fuente, M. (2017). The DD^G-classifier in the functional setting. *TEST*, 26(1):119–142.

Dehghan, S. and Faridrohani, M. R. (2019). Affine invariant depth-based tests for the multivariate one-sample location problem. *TEST*, 28(3):671–693.

Denecke, L. and Müller, C. H. (2011). Robust estimators and tests for copulas based on likelihood depth. *Computational Statistics and Data Analysis*, 55:2724–2738.

Dümbgen, L. (1992). Limit theorems for the simplicial depth. *Statistics and Probability Letters*, 14:119–128.

Falkenau, C. P. (2016). *Depth Based Estimators and Tests for Autoregressive Processes with Application on Crack Growth and Oil Prices*. Dissertation, TU Dortmund.

Gibbons, J. and Chakraborti, S. (2003). *Nonparametric Statistical Inference*. Statistics, textbooks and monographs. Marcel Dekker Incorporated.

Kustosz, C. P., Leucht, A., and Müller, C. H. (2016a). Tests based on simplicial depth for AR(1) models with explosion. *Journal of Time Series Analysis*, 37:763–784.

Kustosz, C. P., Müller, C. H., and Wendler, M. (2016b). Simplified simplicial depth for regression and autoregressive growth processes. *Journal of Statistical Planning and Inference*, 173:125–146.

Liu, R. Y. (1988). On a notion of simplicial depth. *Proceedings of the National Academy of Sciences of the United States of America*, 85:1732–1734.

Liu, R. Y. (1990). On a notion of data depth based on random simplices. *The Annals of Statistics*, 18:405–414.

Lok, W. S. and Lee, S. M. S. (2011). A new statistical depth function with application to multimodal data. *Journal of Nonparametric Statistics*, 23:617–631.

López-Pintado, S. and Romo, J. (2009). On the concept of depth for functional data. *Journal of the American Statistical Association*, 104(486):718–734.

López-Pintado, S., Sun, Y., Lin, J. K., and Genton, M. G. (2014). Simplicial band depth for multivariate functional data. *Advances in Data Analysis and Classification*, 8(3):321–338.

Mizera, I. (2002). On depth and deep points: A calculus. *The Annals of Statistics*, 30(6):1681–1736.

Mizera, I. and Müller, C. H. (2004). Location-scale depth (with discussion). *Journal of the American Statistical Association*, 99:949–966.

Mosler, K. (2002). *Multivariate Dispersion, Central Regions and Depth. The Lift Zonoid Approach.* Lecture Notes in Statistics, 165, Springer, New York.

Müller, C. H. (2005). Depth estimators and tests based on the likelihood principle with application to regression. *Journal of Multivariate Analysis*, 95(1):153–181.

Nagy, S. and Ferraty, F. (2018). Data depth for measurable noisy random functions. *Journal of Multivariate Analysis*, doi.org/10.1016/j.jmva.2018.11.003.

Paindaveine, D. and Van Bever, G. (2013). From depth to local depth: A focus on centrality. *Journal of the American Statistical Association*, 108:1105–1119.

Paindaveine, D. and Van Bever, G. (2018). Halfspace depths for scatter, concentration and shape matrices. *The Annals of Statistics*, 46(6B):3276–3307.

Rousseeuw, P. J. and Hubert, M. (1999). Regression depth. *Journal of the American Statistical Association*, 94(446):388–402.

Tukey, J. W. (1975). Mathematics and the picturing of data. *Proceedings of the International Congress of Mathematicians*, 2:523–531.

Wald, A. and Wolfowitz, J. (1940). On a test whether two samples are from the same population. *The Annals of Mathematical Statistics*, 11(2):147–162.

Wang, J. (2019). Asymptotics of generalized depth-based spread processes and applications. *Journal of Multivariate Analysis*, 169:363–380.

Wellmann, R., Harmand, P., and Müller, C. H. (2009). Distribution-free tests for polynomial regression based on simplicial depth. *Journal of Multivariate Analysis*, 100(4):622 – 635.

Wellmann, R. and Müller, C. H. (2010). Tests for multiple regression based on simplicial depth. *Journal of Multivariate Analysis*, 101(4):824 – 838.

Zuo, Y. and Serfling, R. (2000). General notions of statistical depth function. *The Annals of Statistics*, 28:461–482.