Electronic Journal of Statistics
Vol. 15 (2021) 5283–5309
ISSN: 1935-7524
https://doi.org/10.1214/21-EJS1906

Robust detection of abnormality in highly corrupted medical images*

Xiaoping Shi†

Irving K. Barber School of Arts and Sciences
University of British Columbia
Kelowna, BC V1V 1V7, Canada
e-mail: xiaoping.shi@ubc.ca

Shanshan Qin†,§ **and Yuehua Wu**‡

Department of Mathematics and Statistics
York University
Toronto, Ontario, Canada
e-mail: qinsslzu@gmail.com; wuyh@yorku.ca

Abstract: Medical imaging helps to detect and monitor internal irregularities in the human body. We leverage a block median filtering technique to model pixel-to-pixel differences between two images to develop automated detection of abnormalities in noisy medical images. We propose two robust detection methods, with the test statistic being the conventional maxima and the scale-invariant ratio of the medians from partitioned image grids. Theoretically, we investigate the asymptotic behaviors of two proposed tests. Numerically, we carry out simulation studies to investigate the type I error rate and the power of two tests. In addition, a real application in medical images with gastrointestinal bleeding demonstrates the outperformance and efficiency of the ratio test method. Besides, the developed tests can also be applied to problems in other scientific fields, e.g., air pollution detection using collected remote sensing hyperspectral images.

MSC2020 subject classifications: Primary 62P10, 62G10; secondary 68U10.
Keywords and phrases: Robust detection, median decomposition, maxima test, scale invariant ratio, image filtering.

Received January 2020.

Contents

*Shi's work was supported by NSERC Discovery Grant 2016-05694 and the work by Qin and Wu was supported by NSERC Discovery Grants RGPIN 2017-05720.
†Co-first author
‡First corresponding author
§Second corresponding author

1. Introduction

Medical imaging, collected using imaging technology such as radiology (e.g. ultrasound, CT, X-ray and MRI), nuclear medicine (e.g. PET and SPECT) and optical imaging (e.g. OCT), etc., helps to monitor disorders in the human body. For example, nuclear medicine can diagnose the severity of various diseases such as heart disease, gastrointestinal, endocrine, or neurological disorders, and cancer. Gastrointestinal bleeding is a major cause of death in the United States, with mortality rates ranging from 10% to 30% [18, 14].

Fig. 1 plots a sample of medical images of gastrointestinal bleeding originating from a branch of the superior mesenteric artery. Note that these figures are from Figure 4 of [14]. It can be seen in Fig. 1 that (a) these images, taken in chronological order, are noise-contaminated, (b) the bleeding starts in the 6th image and appears more severe after that (7th-11th images). Fig. 2 shows their histograms of the pixel-to-pixel differences of two consecutive images exhibited in Fig. 1. The histograms in subplots from Hist-1 to Hist-4 appear symmetric about zero, while the histograms in other subplots exhibit asymmetry. The asymmetric histograms correspond to instances in which the two consecutive images are different.

We can vastly improve the patient's prognosis if we can detect the bleeding by the 6th image in Fig. 1. The problem is how to detect abnormality in

FIG 1. *Eleven images with the same dimension 210 × 250 are taken from Figure 4 in [14] with removal of the 1st, 3rd, 8th, and 11th images that have additional marks not belonging to the original images. They were published in J. Nucl. Med. (2016), 57, 252-259,* © *by the Society of Nuclear Medicine and Molecular Imaging, Inc.*

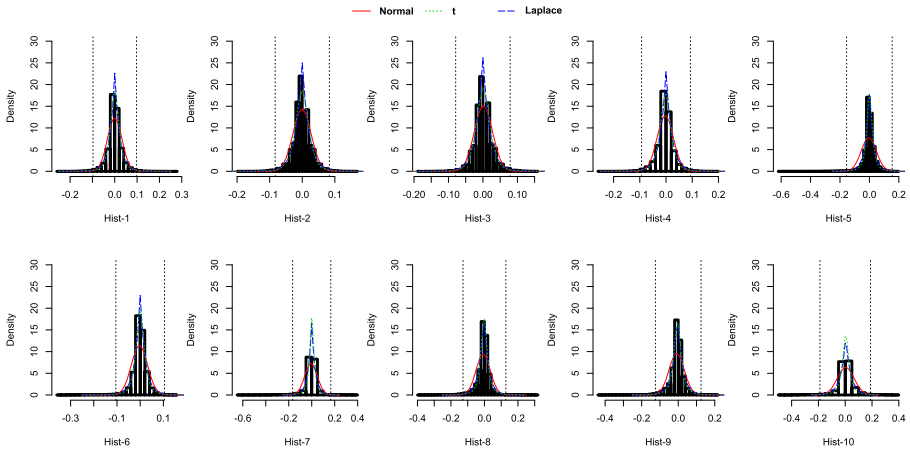

FIG 2. *Histograms of 10 pixel-to-pixel differences of consecutive images with medians being 0, 0, 0, 0, 0, 0, 0, -0.004, -0.004 and 0.004, respectively. The two vertical lines in each plot correspond to* ± 3 *sample standard deviations. We used Normal, t, and Laplace distributions to fit the data.*

noisy images timely and accurately. To tackle the problem as above, we leverage advanced statistical techniques to develop a fast and efficient robust detection method to test whether there is a statistical difference between two consecutive images. We propose to consider the pixel-to-pixel differences between the two images. Under such a framework, abnormality detection is equivalent to testing the symmetry/asymmetry about zero of the distribution for pixel-to-pixel differences. One should note that the images that we study herein do not have systematic change along the time. For example, the image that will change periodically based on the heart rate in ultrasound imaging and brain fMRI is not considered in this study.

As is shown in Fig. 2, the images are fairly noisy, and thus, Normal, t, or Laplace distribution may not be appropriate for modeling the pixel values. A possible way to the problem is to apply the censored weighted likelihood to down-weight the effect of observations with large variances [8, 2]. However, it might be difficult to obtain the true likelihood of the observations since a common error distribution, say, Normal, t or Laplace, can not be assumed. One can alternatively consider nonparametric approaches. Two classical nonparametric methods, i.e., the sign test [13] and the Wald-Wolfowitz runs test (also known

simply as the runs test) [21], are designed to test for symmetry of a distribution about zero. However, the type I error rates for both tests will not be small if the dataset contains a high level of noise. There exists a vast body of other nonparametric methods to test the hypothesis of symmetry of distribution, such as the bootstrap tests [10, 20, 26, 31, 25], the distribution-free tests [11, 22, 28], the kernel-based nonparametric testing methods [15], and also some other methods [9, 2, 4, 5]. For nonparametric tests, some recent literature also establishes the minimax property of the power [19, 30].

Besides the above nonparametric test methods, some researchers have applied a block median filtering technique to avoid specifying the noise distribution. This technique may be suitable for data analysis with positive/negative light/heavy-tailed distributed noise, likely due to sharp and sudden jumps in the image signals. For example, [6] first divided the observations into small bins and took the median values of data in each bin to form a new dataset, and then fit a Gaussian distribution to the new dataset. In light of [6], we propose a new test for testing whether the distribution of pixel values is symmetric about zero using maxima statistic. Beyond that, we propose another test using a ratio statistic, which is shown to have satisfactory performance in both simulation and a real example.

The major contributions of this study are as follows. Firstly, we leverage the block median filtering technique to model the pixel-to-pixel differences between two consecutive images. To the best of our knowledge, this technique is initially introduced into the problem of image abnormality detection. Secondly, we investigate the consistency of the proposed tests under mild conditions. Besides, The finite sample properties of the tests are enhanced via simulations, including the type I error rate and the power of the test. Finally, although this study is motivated by and developed for Gastrointestinal bleeding detection, it can also be potentially used to problems in quality control, satellite fire and smoke detection, and others.

The remaining paper is organized as follows. In Section 2, we introduce the median filtering method to address this abnormality detection problem and the model setup. In Section 3, we propose two nonparametric tests and the detection procedure. In Section 4, we approximate the p-value of the proposed tests and investigate their powers under mild conditions. In Section 5, we investigate the proposed tests by simulations and real data analysis. Section 6 concludes this paper.

2. Problem and model setup

2.1. Problem description

An image can be represented as a three-way tensor $(R, G, B) \in \mathbb{R}^{d_1 \times d_2 \times 3}$, where $R, G, B \in \mathbb{R}^{d_1 \times d_2}$ are two-dimensional pixel matrices. For simplity, we consider grayscale images, the pixel values of which are expressed as a matrix $X \in \mathbb{R}^{d_1 \times d_2}$. The elements of X can be accessed using two indices as in x_{i_1,i_2}, varying

i_j from one to d_j, $j = 1, 2$. Suppose that there are two images $X^{(1)}$ and $X^{(2)}$, and let $x^{(1)}_{i_1,i_2}$ and $x^{(2)}_{i_1,i_2}$ denote the respective pixel values in the grid location $\{i_1, i_2\}$ of these two images. Define $y_{i_1,i_2} = x^{(2)}_{i_1,i_2} - x^{(1)}_{i_1,i_2}$, $i_j \in \{1, \ldots, d_j\}$ for $j = 1, 2$. After the proper alignment, we consider that the differences y_{i_1,i_2} are independently distributed with the same distribution. We write y_{i_1,i_2} as

$$y_{i_1,i_2} = \mu_{i_1,i_2} + \epsilon_{i_1,i_2}, \tag{1}$$

where μ_{i_1,i_2} are the means, and ϵ_{i_1,i_2} are independently identically distributed (i.i.d.) random errors with median zero and standard deviation σ that may potentially be large due to limitations in current imaging technology; see Assumption A1 in Section 4. Under this assumption, the distribution of y_{i_1,i_2} is symmetric about zero if μ_{i_1,i_2} is 0, otherwise skewed to the right or left if μ_{i_1,i_2} is either positive or negative, respectively.

Let $\mathcal{G}_q, q \in \{1, \ldots, r\}$ be disjoint sets representing signals of abnormalities, and $\delta_q, q \in \{1, \ldots, r\}$ be their respectively unknown nonzero constants. Then $\mu_{i_1,i_2} \neq 0$ if $(i_1, i_2) \in \mathcal{G}_q$ for $q \in \{1, \ldots, r\}$, implying the existence of abnormalities at region $\{(i_1, i_2) \in \mathcal{G}_q, q \in \{1, \ldots, r\}\}$ in the second image $X^{(2)}$. This model describes the phenomenon that some abnormal parts are hidden in the pixel-to-pixel image differences. The problem is to identify the existence of these abnormalities.

Remark 1. *The medical images taken within a short period may be considered aligned. Hence, the differences in pixel values of the two images could be assumed to be i.i.d. under the null hypothesis. For example, the images of the artificial flower in [29] are aligned well since the wind has a rare impact on the camera and the artificial flower. In some cases, the images may be spatially localized. For example, if the artificial flower is replaced by a real one, its location in all taken images can no longer be considered fixed. Such a problem may also arise in the medical images. Ignoring these types of problems may presumably lead to a loss of power of the test. To ensure that the two images are properly aligned, we may rotate or relocate the two images by minimizing the sum of the two images' absolute differences.*

Remark 2. *Note that one may consider grayscale images (single channel: two-way tensor) or color images (multichannel: three-way tensor). One can also consider a p-way tensor, $\mathcal{X} \in \mathbb{R}^{d_1 \times \cdots \times d_p}$, where d_j is the dimension of mode-j of $\mathcal{X}$, $j = 1, \ldots, p$. It is thus an extension of the one-dimensional case presented in [6] to a p-dimensional one. However, for simple presentation, we only consider the grayscale images in this paper.*

2.2. Block median filtering

The major problem is that the distribution of noise is hard to specify. To tackle such a problem, we plan to use the block median filtering technique. Let n be the number of pixel values of the image, that is, $n = d_1 \times d_2$. Suppose that the

image pixels can be equally divided by κ_n disjoint partitions, which we label by $A_\ell \in \mathbb{R}^{m_1 \times m_2}$, $\ell = 1, 2, \ldots, \kappa_n$. Denote the number of pixels in A_ℓ by m_n, i.e., $|A_\ell| = m_n = m_1 m_2$ for $\ell = 1, 2, \ldots, \kappa_n$. Then, compute a sequence of minimizers x_ℓ, $\ell = 1, 2, \ldots, \kappa_n$, satisfying that,

$$x_\ell = \arg\min_{\gamma} \sum_{(i_1, i_2) \in A_\ell} \rho(y_{i_1, i_2} - \gamma), \tag{2}$$

where ρ is a given positive and symmetric loss function. Hereafter we suppress the subscript n in both κ_n and m_n for simplicity of notation.

Note that if ρ in (2) is the absolute value function, x_ℓ is a median of the partition A_ℓ, i.e.,

$$x_\ell = \text{median}\{y_{i_1, i_2},\ (i_1, i_2) \in A_\ell\}, \ell = 1, 2, \ldots, \kappa. \tag{3}$$

The block median transformations have been applied in the analysis of microarray data for data normalization in [27, 6]. In this study, we employ the median filtering as it is robust against a large collection of error distributions, including heavy-tailed, which is not unusual in literature.

Remark 3. *By changing ρ, x_ℓ can turn into the mean, minimum, and maximum of $\{y_{i_1, i_2}, (i_1, i_2) \in A_\ell\}$, all of which have different use in image filtering. Respectively, (a) the mean filter can be used to remove light-tailed noise; (b) the minimum filter that can remove positive impulses (white spots); (c) the maximum filter can effectively remove highly negative pixel values (black spots); additionally, (d) the median filter, applied in (3), can be used to remove not only the positive impulses and the highly negative pixel values, but also heavy-tailed noise (e.g. salt and pepper noise) caused by sharp and sudden disturbances in the image signals.*

2.3. Model setup

In light of [6], the sequence of block medians, $\{x_1, x_2, \ldots, x_\kappa\}$, is modelled as

$$x_\ell = \theta_\ell + \eta_\ell, \ell = 1, \ldots, \kappa, \tag{4}$$

where $\eta_\ell = \text{median}\{\epsilon_{i_1, i_2},\ (i_1, i_2) \in A_\ell\}$, $\theta_\ell = \delta_q \neq 0$ if there exists a q such that $A_\ell \subseteq \mathcal{G}_q$, otherwise $\theta_\ell = \delta_q^* \in [-\max_q(|\delta_q|), \max_q(|\delta_q|)]$ for $\mathcal{G}_q$, $q \in \{1, \ldots, r\}$, being disjoint sets of image pixels with convention that $\theta_\ell = 0$ if $r = 0$. One should note that δ_q^* is another unknown constant that may differ from δ_q for $q \in \{1, \ldots, r\}$.

Denote $\mathcal{U} = \{\mathcal{G}_1, \ldots, \mathcal{G}_r\}$. Our aim is to test

$$H_0 : \mathcal{U} = \emptyset \quad \text{v.s} \quad H_1 : \mathcal{U} \neq \emptyset, \tag{5}$$

i.e., to test if there exists one or more disjoint sets of image pixels $\mathcal{G}_q$, $q = 1, 2, \ldots, r, (r \geq 1)$, such that $\theta_\ell = \delta_q \neq 0$ for $A_\ell \subset \mathcal{G}_q$, against $\theta_\ell = 0$ for all $\ell \in \{1, \ldots, \kappa\}$. We remark that our proposed two tests will be more powerful for large r.

3. Nonparametric hypothesis tests

As displayed in the first four histograms (where there are no abnormalities) in Fig. 2, the null distribution of $\{y_{i_1,i_2}\}$ appears to be symmetric about 0. Further examinations reveal that when there exists an abnormality, the corresponding histogram is skewed. These observations lead us to test whether or not the null distribution of $\{y_{i_1,i_2}\}$ is symmetric about zero to detect images with abnormalities. Since such images are usually fairly noisy, we consider the medians $\{x_\ell\}$ and thus propose tests based on the maxima and ratio statistics.

3.1. Maxima test

As explained above, in light of [6], we propose the following maxima as the test statistic for testing H_0 against H_1,

$$M_\kappa = \max_{1\le\ell\le\kappa} |x_\ell|/\hat{\sigma}, \tag{6}$$

where x_ℓ is defined in (3) and $\hat{\sigma}$ is an estimate of the median absolute deviation (MAD), i.e.,

$$\hat{\sigma} = \text{median}_{1\le\ell\le\kappa}\{|x_\ell - \text{median}_{1\le k\le\kappa}\{x_k\}|\}/0.6745.$$

We reject H_0 when M_κ is larger than the corresponding critical value under a given significant level α. For convenience, we hereafter refer to this test as the "maxima test".

It is noted that there exists a nuisance parameter, $\hat{\sigma}$, in the maxima statistic. One can use the sample standard deviation in place of MAD for $\hat{\sigma}$ in (6). An inappropriate choice of the nuisance parameter may result in poor performance, say, a higher type I error rate or lower power of the test. This motivates us to consider another ratio statistic that does not involve the estimation of a nuisance parameter.

3.2. Ratio test

The estimate of the nuisance parameter in the maxima test may negatively affect the performance of the test, which may explain less optimal results in empirical examples. With the purpose of the practical applicability of the test, we may restrict our attention to a closed interval of block medians denoted as $[a, b]$, in view that the null distribution is assumed to be symmetric about zero. Assume that the observed pixel values of images are between 0 and 1. Under H_0, the interval may satisfy that $|a| = |b|$, with $a < 0, b > 0$. In contrast, under H_1, the interval may satisfy that $|a| \neq |b|$ with $a < 0, b > 0$. Next, we will show that how to construct the ratio-type statistic.

It is noted that the only information we have about the common error distribution is whether or not its probability density function is symmetric about

zero in our problem. We assume that the common distribution of ϵ_{i_1,i_2}, F, is continuous and has median 0, i.e., $F(0) = 1/2$. Denote its probability density function by f. Let $m_0 = (m-1)/2$. It is easy to show that the common density function of the block medians $x_1, \ldots, x_\kappa$ under H_0 (which we denote by $g(x)$) satisfies

$$g(x) \propto F^{m_0}(x)[1 - F(x)]^{m_0} f(x).$$

Fig. 2 visually shows that the histograms from Hist-1 and Hist-4 appear to be symmetric about zero, while others exhibit some asymmetry. Thus, we consider the following null and alternative hypotheses for testing symmetry of the distribution function F:

$$\tilde{H}_0 : \ a = -a_0, \ b = a_0 \text{ with } a_0 > 0 \ \text{ vs } \ \tilde{H}_1 : \ a = a_1 < 0, \ b = a_2 > 0 \text{ with } |a_1| \neq |a_2|.$$

The proposed test statistic described in (7) makes use of a first-order approximation of $F(x)$ or $f(x)$.

Lemma 1. *Given the first-order approximation to the density function $f(x)$ over the interval $[a, b]$, the following statistic*

$$\frac{x_{(\kappa)} - x_{(1)}}{\max\{x_{(\kappa)}, -x_{(1)}\}} \tag{7}$$

is proportional to the likelihood ratio statistic for testing $\tilde{H}_0$ against $\tilde{H}_1$, where $x_{(\kappa)} = \max\{x_\ell, \ \ell = 1, \ldots, \kappa\}$ and $x_{(1)} = \min\{x_\ell, \ \ell = 1, \ldots, \kappa\}$. The smaller the test statistic (7), the more evidence we have that supports the rejection of $\tilde{H}_0$.

The proof of Lemma 1 is given in Appendix A.1. Note that $x_{(\kappa)} - x_{(1)}$ can be expressed as $\max\{x_{(\kappa)}, -x_{(1)}\} + \min\{x_{(\kappa)}, -x_{(1)}\}$ and hence

$$\frac{x_{(\kappa)} - x_{(1)}}{\max\{x_{(\kappa)}, -x_{(1)}\}} = 1 + \frac{\min\{x_{(\kappa)}, -x_{(1)}\}}{\max\{x_{(\kappa)}, -x_{(1)}\}}.$$

It can be seen that the smaller the likelihood ratio, the smaller the value of $\min\{x_{(\kappa)}, -x_{(1)}\}/\max\{x_{(\kappa)}, -x_{(1)}\}$, or equivalently the larger the value of $\max\{x_{(\kappa)}, -x_{(1)}\}/\min\{x_{(\kappa)}, -x_{(1)}\}$ which, from above, has been shown to be equal to $(x_{(\kappa)} - x_{(1)})/\min\{x_{(\kappa)}, -x_{(1)}\} - 1$. It is obvious then that the smaller the test statistic (7), the larger the following statistic:

$$T_\kappa = \frac{x_{(\kappa)} - x_{(1)}}{\min\{x_{(\kappa)}, -x_{(1)}\}}. \tag{8}$$

Therefore, we reject the null hypothesis $\tilde{H}_0$ for large values of T_κ. For convenience, we call the test based on T_κ the "ratio test". The ratio statistic takes advantage of the median of block medians, which can be decomposed into the sum of a Gaussian random variable and a random error with median zero under some mild assumptions made in [6]. It is noted that the numerator of T_κ in (8) captures the range of the data, while the denominator of T_κ is neutral to the impact of the largest (unsigned) data point. In addition, T_κ is scale-invariant. Therefore, T_κ is the desired test statistic for testing H_0 against H_1.

Remark 4. *One can easily show that* $T_k \geq 2$ *as follows,*

$$T_\kappa = \begin{cases} \dfrac{x_{(\kappa)} - x_{(1)}}{-x_{(1)}}, & \text{if } x_{(\kappa)} \geq -x_{(1)} \\ \dfrac{x_{(\kappa)} - x_{(1)}}{x_{(\kappa)}}, & \text{if } x_{(\kappa)} \leq -x_{(1)} \end{cases} \geq \begin{cases} \dfrac{-x_{(1)} - x_{(1)}}{-x_{(1)}}, & \text{if } x_{(\kappa)} \geq -x_{(1)} \\ \dfrac{x_{(\kappa)} + x_{(\kappa)}}{x_{(\kappa)}}, & \text{if } x_{(\kappa)} \leq -x_{(1)} \end{cases}$$
$$\geq 2, \tag{9}$$

T_κ is expected to be close to 2 under H_0, but to be far away from 2 under H_1. For demonstration purpose, one can see the following example. Suppose that we have block medians $\{x_1, x_2, x_3, x_4\} = \{-1, 1, -2, 2\}$. Then, it is plausible that H_0 is true, as the value of T_κ is $\frac{2-(-2)}{\max\{2,-(-2)\}} = 2$. If $\{x_1, x_2, x_3, x_4\} = \{-1, 1, 2, 4\}$, it is plausible that H_1 holds as opposed to H_0 since the value of T_κ is $\frac{4-(-1)}{\min\{4,-(-1)\}} = 5$, which is much larger than 2. If $\{x_1, x_2, x_3, x_4\} = \{-1, 1, -2, -4\}$, it is still plausible that H_1 holds. The value of T_κ here is $\frac{1-(-4)}{\min\{1,-(-4)\}} = 5$ (again, larger than 2). These results are in line with our expectations.

3.3. Detection procedure

The above two proposed tests can detect abnormalities of medical images under the following detection procedures. Suppose that we have a sequence of medical images, say, $X^{(1)}, \ldots, X^{(K)} (K \geq 2)$, and the abnormal images with significant differences are $X^{(s)}, s \in D^{(0)}$, where $D^{(0)}$ is a subset of $\{1, \ldots, K\}$. Let $\hat{D}$ be an estimate of $D^{(0)}$. We present the following detection procedure to obtain $\hat{D}$. Set $\alpha = 0.05$.

Step 1. Set initial values: $k = 2$, $s = 1$, and $\hat{D} = \emptyset$.
Step 2. Compare the images $X^{(s)}$ and $X^{(k)}$. If the approximate p-value of the maxima test or the ratio test is less than α, then keep the kth image, i.e., $s = k$; otherwise, drop the kth image and $s = s$. And $\hat{D} = \hat{D} \cup \{s\}$.
Step 3. If $k = K$, stop. Otherwise, let $k = k + 1$ and return to Step 2.

Users can choose their preferred significant level α. The approximate p-values can be computed via Eq.(11) or Eq.(14) in Section 4. The proposed detection procedure is not intended to replace medical experts. Instead, the detection procedure aims to serve as an initial screening technique and alert medical experts timely when an abnormality occurs. Generally, the detection procedure can be applied in the following scenarios.

Remark 5. *The detection procedure can determine whether the most recent image is abnormal or not. For example, suppose that there is a patient who is seeing a doctor. A medical image of the patient shows that nothing is wrong. However, the patient continues to complain of discomfort. Thus, the doctor is likely to monitor the patient's condition and thus request more (say $K - 1$ with $K \geq 2$) images to be taken at scheduled times. Images are analyzed as they come in until perhaps, a significant difference is identified. Under such a scenario, the*

first image should be removed from $\hat{D}$, that is, $\hat{D} = \hat{D}/\{1\}$. If the purpose of the detection is to alert medical experts timely, we can stop the procedure when $|\hat{D}| = 1$.

Remark 6. *The detection method enables us to keep significantly different images for further examination. For example, suppose that a person has a history of a specific disease in his/her family, currently showing no signs of the inherited disease; however, he/she may develop the disease during some period of his/her life at high risk. Suppose that radiological imaging can be applied to diagnose the disease. A doctor would ask him/her to take several radiological images, say K, in a certain period. In this scenario, the detection procedure can be used to perform an initial screening of the images and delete these images that are not statistically different from ones taken just before them.*

Intuitively, an ideal estimate should contain important information, and exclude these informationless ones, i.e., $D^{(0)} \subset \hat{D}$ and the size of $\hat{D}$ (denoted as $|\hat{D}|$) is smallest. Thus, in order to measure the estimation efficiency, we define the criteria ω, i.e.,

$$\omega = a|D^{(0)}/\ \hat{D}| + (1-a)|\hat{D}/\ D^{(0)}|, \tag{10}$$

where $S_1/\ S_2$ denotes the intersection of S_1 and the complement of S_2, and $0 \le a \le 1$. The smaller the value of ω, the better the estimation. $|D^{(0)}/\ \hat{D}| > 0$ means that some important images are deleted, while $|\hat{D}/\ D^{(0)}| > 0$ implies that some informationless images are kept. Since $|D^{(0)}/\ \hat{D}| > 0$ is more serious than $|\hat{D}/\ D^{(0)}| > 0$, we thus assign a large weight on $|D^{(0)}/\ \hat{D}|$, say, $a = 0.95$.

4. Asymptotic properties of the proposed tests

Before proceeding, we make the following assumptions.

Assumption A1. $\{\epsilon_{i_1,i_2}\}$ are i.i.d. continuous random errors with the common distribution F symmetric about zero and its density function f satisfying $f(0) > 0$ and $|f(w) - f(0)| \le cw^2$ in an open neighborhood of zero, where c is an unknown constant that may vary with situation.

Assumption A2. $m \to \infty$, $\kappa/(m \log m) \to \infty$, and $\kappa \log \kappa / m^3 \to 0$, as $n \to \infty$.

As described in [6], Assumption A1 is satisfied by the Cauchy distribution, the Laplace distribution, the t-distribution, the Normal distribution, and other distributions. Under this assumption, we can approximate the distribution of the median of errors as Normal. Assumption A2 describes how m grows relative to κ. Simulation studies suggest that $m = 100$ would allow for a satisfactory Normal approximation.

4.1. Asymptotic properties of the maxima test

We have the following result on the approximation of the p-value of the maxima test.

Theorem 4.1.. *Suppose that H_0 with $r = 0$ given in (5) holds true. Under the Assumptions A1-A2, for any $v \in (-\infty, \infty)$, we have*

$$pr\left(\frac{M_\kappa - \mathtt{a}_\kappa}{\mathtt{b}_\kappa} > v\right) \to 1 - \exp\{-2\exp(-v)\}, \quad as\ n \to \infty \tag{11}$$

where M_κ is given in (6), $\mathtt{a}_\kappa = \sqrt{2\log\kappa} - [\log\log\kappa + \log(4\pi)]/[2\sqrt{2\log\kappa}]$, and $\mathtt{b}_\kappa = 1/\sqrt{2\log\kappa}$.

The following Proposition is needed for proving Theorem 4.1. (see Lemma 1 in [6]).

Proposition 1. *Under the Assumption A1, η_ℓ defined in (4) can be written as*

$$\eta_\ell = \frac{1}{2f(0)\sqrt{m}} z_\ell + \frac{1}{\sqrt{m}}\zeta_\ell, \tag{12}$$

where z_ℓ are i.i.d. $N(0,1)$, ζ_s are independently distributed random variables which can be decomposed into $\zeta_s = \zeta_{s,1} + \zeta_{s,2}$ such that $E(\zeta_{s,1}) = 0$, all moments of $|\zeta_{s,1}|/m$ are finite and $pr(\zeta_{s,2} \neq 0) \leq c_1 \exp(-mc_2)$ for some constants $c_1 > 0$ and $c_2 > 0$.

The proof of Theorem 4.1. is given in the Appendix A.2. Assume that α is the significant level of the hypothesis test. Theorem 4.1. implies that the type I error rate goes to α under mild conditions.

Besides the type I error rate, we will also show that the maxima test is powerful. Before proceeding, we make an additional assumption.

Assumption A3. Under H_1, $\tilde{S} \neq \emptyset$, where $\tilde{S} = \{s : A_s \subseteq \mathcal{G}_q \text{ for some } \mathcal{G}_q, q \in \{1,\ldots,r\}, s \in \{1,\ldots,\kappa\}\}$. In addition, there is a constant $\varrho > 0$ such that $\min_{q\in\{1,\ldots,r\}}\{|\delta_q|\} > \varrho$.

The Assumption A3 implies that there exists at least one set $\mathcal{G}_q$ such that it includes A_s.

Theorem 4.2.. *Suppose that H_1 given in (5) holds true. Under the Assumptions A1-A3, we have $M_\kappa/\log\kappa \to \infty$ in probability as $n \to \infty$, where M_κ is given in (6). Furthermore, for any $v \in (-\infty, \infty)$, we have that*

$$pr\left(\frac{M_\kappa - \mathtt{a}_\kappa}{\mathtt{b}_\kappa} > v\right) \to 1, n \to \infty. \tag{13}$$

The proof of Theorem 4.2. is provided in the Appendix A.3. This theorem guarantees that the power of the maxima test converges to one as the number of pixels goes to infinity.

4.2. Asymptotic properties of the ratio test

We have the following result on the approximation of the p-value of the ratio test.

Theorem 4.3.. *Suppose that H_0 with $r = 0$ defined in (5) holds true. Under the Assumptions A1-A2, for $\nu > 2$, as $n \to \infty$, we have that*

$$\left| pr(T_\kappa \geq \nu) - 1 + 2\int_{-\infty}^{0} \kappa\phi(x)\{[\Phi((1-\nu)x) - \Phi(x)]^{\kappa-1} - [1-2\Phi(x)]^{\kappa-1}\}dx \right| \to 0, \tag{14}$$

where T_κ is given in (8), $\Phi(x)$ and $\phi(x)$ denote respectively the standard normal distribution and its density function.

We ask the readers to turn to Appendix A.4 for details on the proof. Similar to Theorem 4.1., Theorem 4.3. also controls the type I error rate of the ratio test. Next, we consider the power of the ratio test. Since the main purpose of this paper is to detect abnormalities in medical images, all $\delta_q, q \in \{1, \ldots, r\}$ should have the same sign if there are abnormalities in such images. Thus, we make the following assumption.

Assumption A4. $\delta_1, \ldots, \delta_r$ with $r > 1$ are either all greater than zero or all less than zero.

Then, we have the following theorem.

Theorem 4.4.. *Suppose that H_1 given in (5) holds true. Under the Assumptions A1-A4, we have $T_\kappa \to \infty$ in probability as $n \to \infty$, where T_κ is given in (8). Furthermore, for $\nu > 2$, we have that $pr(T_\kappa \geq \nu) \to 1$.*

One can refer to Appendix A.5 for the proof of Theorem 4.4.. Theorem 4.4. implies that the power of the ratio test converges to one under mild conditions.

5. Numerical studies

5.1. Simulations

In the following, we present simulation studies to examine the performance of the proposed tests.

5.1.1. Example 1: $K = 2$ and $r = 1$

As in (1)-(2), set $p = 2$, $d_1 = d_2 = \sqrt{n}$, $\kappa = 2500$, and $m = (5j)^2$ for $j = 1, 2, \ldots, 8$, $n = m\kappa$, and $\{\epsilon_{i_1,i_2}\}, i_1, i_2 = 1, \ldots, \sqrt{n}$, are generated from the following distributions:

- Case 1. $N(0, 0.16)$, where $N(\mu, \sigma^2)$ denotes the Normal distribution with mean μ and variance σ^2;
- Case 2. $0.4t(10)$, where $X \sim 0.4t(10)$ if and only if $X/0.4 \sim t(10)$, and $t(\gamma)$ denotes the t distribution with the degrees of freedom γ;
- Case 3. $Laplace(0, 0.3)$, where $Laplace(a, b)$ denotes the Laplace distribution with location a and scale b;

- Case 4. $0.5N(0, 0.09)+0.5N(0, 0.25)$, where $\alpha N(0, \sigma_1^2)+(1-\alpha)N(0, \sigma_2^2)$ denotes a contaminated-normal distribution with random errors being generated from $N(0, \sigma_1^2)$ with probability α and from $N(0, \sigma_2^2)$ with probability $1 - \alpha$.
- Case 5. $0.2N(-0.01, 0.09) + 0.6N(0, 0.25) + 0.2N(0.01, 0.09)$, a contaminated-normal distribution that is similar to $\alpha N(0, \sigma_1^2) + (1 - \alpha)N(0, \sigma_2^2)$.

The simulated type I errors for the proposed two tests are reported based on 10,000 repetitions with the p-value approximated via Theorem 4.1. or Theorem 4.3. (depending on the test). We set $\alpha = 0.05$. It can be seen from Table 1 that both tests perform well with the normal distributed random errors (Case 1), reflected by the type I errors abound 0.05 for the whole candidate m. When the random errors are distributed with the Laplace (Case 3) or contaminated-normal (Cases 4-5) distribution, the type I errors tends to be stable around 0.05 as $m \geq 10^2$. A special case one can see is Case 2, where the type I errors are larger than 0.05, and the values obtained via the ratio test are smaller than those via the maxima test, but their values decrease as m increases. Overall, both the proposed methods perform well in most cases in terms of type I errors as the partition size m is larger than a threshold, say, 10^2.

TABLE 1
Type I errors of both maxima test and ratio test

	Test	$m = 5^2$	$m = 10^2$	$m = 15^2$	$m = 20^2$	$m = 25^2$	$m = 30^2$	$m = 35^2$	$m = 40^2$
Case 1	Maxima	0.050	0.048	0.048	0.049	0.046	0.045	0.043	0.046
	Ratio	0.053	0.050	0.051	0.048	0.049	0.048	0.049	0.049
Case 2	Maxima	0.086	0.050	0.047	0.047	0.048	0.048	0.046	0.046
	Ratio	0.061	0.054	0.048	0.046	0.051	0.053	0.050	0.054
Case 3	Maxima	0.854	0.389	0.243	0.170	0.139	0.115	0.101	0.096
	Ratio	0.129	0.092	0.074	0.066	0.070	0.065	0.063	0.060
Case 4	Maxima	0.094	0.054	0.048	0.045	0.052	0.045	0.049	0.041
	Ratio	0.067	0.053	0.051	0.051	0.056	0.048	0.051	0.050
Case 5	Maxima	0.092	0.056	0.053	0.046	0.045	0.047	0.044	0.047
	Ratio	0.061	0.052	0.054	0.053	0.049	0.052	0.051	0.055

In addition to the type I error rates, we investigate powers of both maxima test and ratio test when H_1 holds. 10,000 simulations are carried out with $\delta_1 = -0.15$ or 0.15 in the first grid. Table 2 lists the powers of both tests. The powers of both tests are increasing with the partition size m either for negative or positive δ_1, across all Cases (1-5). Specifically, when $m = 25^2$, the minimum powers of the maxima test and ratio test increase to 0.999 and 0.994, respectively. This result agrees with the power consistency in Theorems.

TABLE 2
Powers of both maxima test and ratio test

Case	Test	$\delta = -0.15$					$\delta = 0.15$				
		$m = 5^2$	$m = 10^2$	$m = 15^2$	$m = 20^2$	$m = 25^2$	$m = 5^2$	$m = 10^2$	$m = 15^2$	$m = 20^2$	$m = 25^2$
Case 1	Maxima	0.053	0.141	0.599	0.958	0.999	0.053	0.141	0.584	0.955	0.999
	Ratio	0.056	0.117	0.503	0.914	0.996	0.049	0.121	0.496	0.910	0.996
Case 2	Maxima	0.091	0.137	0.551	0.939	0.999	0.086	0.128	0.551	0.943	0.999
	Ratio	0.065	0.110	0.468	0.885	0.994	0.062	0.107	0.459	0.893	0.992
Case 3	Maxima	0.859	0.821	0.998	1.000	1.000	0.862	0.821	0.998	1.000	1.000
	Ratio	0.126	0.429	0.977	1.000	1.000	0.131	0.429	0.977	1.000	1.000
Case 4	Maxima	0.094	0.179	0.702	0.982	1.000	0.093	0.190	0.702	0.983	1.000
	Ratio	0.067	0.144	0.607	0.957	0.998	0.065	0.150	0.606	0.955	0.999
Case 5	Maxima	0.098	0.159	0.617	0.963	1.000	0.097	0.155	0.625	0.964	1.000
	Ratio	0.067	0.124	0.514	0.922	0.997	0.067	0.123	0.529	0.922	0.998

5.1.2. Example 2: $K = 2$ and $r = 2$

In this subsection, we compare the power of both proposed tests when there is δ_1 shift in the first grid and a δ_2 shift in the second grid. Let $\{\epsilon_{i_1,i_2}\}$ be i.i.d. from the above five distributions given in Example 1. The critical values are computed by using Theorem 4.1. or Theorem 4.3. with $\alpha = 0.05$.

TABLE 3
Power comparison for δ_1 and δ_2

Case	Test	$\delta_1 = \delta_2 = -0.15$					$\delta_1 = \delta_2 = 0.15$				
		$m = 5^2$	$m = 10^2$	$m = 15^2$	$m = 20^2$	$m = 25^2$	$m = 5^2$	$m = 10^2$	$m = 15^2$	$m = 20^2$	$m = 25^2$
Case 1	Maxima	0.059	0.364	0.965	1.000	1.000	0.059	0.365	0.965	1.000	1.000
	Ratio	0.057	0.289	0.888	0.997	1.000	0.055	0.296	0.891	0.998	1.000
Case 2	Maxima	0.094	0.337	0.950	1.000	1.000	0.093	0.342	0.949	1.000	1.000
	Ratio	0.065	0.256	0.861	0.997	1.000	0.065	0.260	0.865	0.997	1.000
Case 3	Maxima	0.874	0.995	1.000	1.000	1.000	0.869	0.995	1.000	1.000	1.000
	Ratio	0.129	0.795	1.000	1.000	1.000	0.130	0.804	1.000	1.000	1.000
Case 4	Maxima	0.101	0.472	0.990	1.000	1.000	0.103	0.471	0.989	1.000	1.000
	Ratio	0.071	0.362	0.945	0.999	1.000	0.066	0.362	0.943	1.000	1.000
Case 5	Maxima	0.101	0.401	0.971	1.000	1.000	0.104	0.385	0.971	1.000	1.000
	Ratio	0.066	0.313	0.901	0.998	1.000	0.068	0.293	0.897	0.999	1.000

In this example, δ_1 and δ_2 are of the same sign as required by assumption A4. We carry out 10,000 simulations. The power comparisons are given in Table 3. Clearly, as m increases, the powers of both tests are increasing, which agrees with the results in Example 1. When m is large enough, say, $m \geq 20^2$, both tests show strong powers (≥ 0.997) for all the cases.

5.1.3. Example 3: $K = 11$ and $r = 1$

We examine the performance of both maxima test and ratio test for abnormal images detection. Let the pixel-to-pixel differences between the ℓth and $\ell - 1$th images be generated independently from $\delta_1 + 0.4t(\gamma)$ distribution with $\gamma = \ell + 1$ for $\ell = 2, \ldots, 11$, where $\delta_1 = 0.25$ for the first two grids when $\ell = 5, 10$, and 0 otherwise.

We carried out 10,000 simulations. Fig. 3 displays the boxplots of ω (see Eq. (10)) obtained via the maxima test (a) and the ratio test (b) under different partition sizes. It can be seen that the performance of both tests becomes better as the partition size increases, reflected by the decreasing trends of ω. As m increases from 10^2 to 20^2, the medians of ωs drop significantly for both test methods. This result demonstrates the efficiency of the proposed tests in terms of abnormal image detection.

5.2. Real data example

Now we return to analyze the medical images with gastrointestinal bleeding described in Section 1. As is displayed in Fig. 1, from the 6th image onwards, bleeding intensifies gradually. And it is difficult to distinguish the images, $X_1, \ldots, X_5$

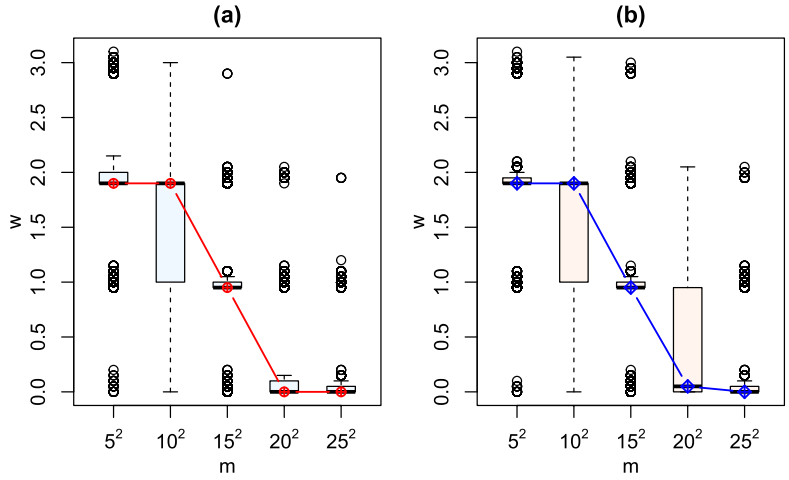

FIG 3. *Boxplots of ω obtained via the maxima test (a) and the ratio test (b) over 10,000 simulations. The red and blue lines represent the medians of ω for the maxima test and the ratio test, respectively. The x-axis denotes the partition size, while the y-axis denotes the w value.*

TABLE 4
The p-values and $\hat{D}$ for abnormality detection of the medical image with gastrointestinal bleeding.

		$\ell=2$	$\ell=3$	$\ell=4$	$\ell=5$	$\ell=6$	$\ell=7$	$\ell=8$	$\ell=9$	$\ell=10$	$\ell=11$	$\hat{D}$
$m=2^2$	Maxima	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	$\{2,\ldots,11\}$
	Ratio	0.4240	0.2731	0.4657	0.5664	**<0.001**	**0.0051**	**0.0066**	0.3801	0.1878	0.8169	$\{6,7,8\}$
$m=5^2$	Maxima	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	$\{2,\ldots,11\}$
	Ratio	0.2687	**0.0267**	0.2433	0.0943	**0.0000**	**0.0328**	0.1204	0.1515	0.0885	**0.0162**	$\{3,6,7,11\}$
$m=10^2$	Maxima	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	$\{2,\ldots,11\}$
	Ratio	**0.0037**	0.2515	0.7473	0.1130	**0.0000**	**0.0182**	0.9378	**0.0328**	**0.0000**	0.0822	$\{2,6,7,9,10\}$)

(before bleeding) and $X_8,\ldots,X_{11}$ (after bleeding) by eyeballing. This study aims to detect the initial point when the bleeding begins and the timepoints when the bleeding gets worse than that at the initial time. We consider the pixel-to-pixel differences between images by applying the detection procedure provided in Section 3.3 to solve this problem. The size of each partition is given equally, such as $m = 2^2, 5^2, 10^2$. We report the p-values for the ℓ-th image ($\ell = 2,\ldots,11$) and the final $\hat{D}$ obtained via both tests in Table 4.

Unfortunately, the maxima test fails to detect the abnormalities in these medical images. It can be seen from Table 4 that $\hat{D}$ via the maxima test contains all the images for different m; that is, there exists a significant difference between any two images. These results are contrary to the fact that the initial bleeding starts from the 6th image. The failure of the maxima test in real data may result from the poor estimate of $\hat{\sigma}$ in Eq. (6). When it comes to the ratio test, however, it outperforms the maxima test. An interesting finding is that the size of $\hat{D}$ is increasing as the partition size m increases from 2^2 to 10^2, resulting in the increased power of the test, which is consistent with the Theorem

4.4.. Specifically, the intersection of $\hat{D}$ is $\{6, 7\}$ for different m. The ratio test accurately detects the initial time point when it starts bleeding since $\{6\} \subset \hat{D}$ for different m. Intuitively, there exists a slight difference between the 6th and 7th images. The results of $\{7\} \subset \hat{D}$ for different m also demonstrate the outperformance of the ratio test in detecting a slight difference between two abnormal images.

6. Conclusion and discussion

This paper proposes two robust detection methods, the maxima test and the ratio test, by leveraging the block median technique. Theoretically, we show that the p-values for both tests can be approximated with satisfactory accuracy, and the consistency of both tests under mild assumptions. We develop the detection procedure and discuss its applications in detecting abnormalities in noisy medical images under two scenarios. Numerically, simulation studies and real data analysis are performed to examine the performance of the proposed tests. The maxima test performs well in simulations but falls short in the real data application. This phenomenon may be due to the presence of the nuisance parameter, i.e., the standard deviation of the median. Although the MAD estimator in (6) to improve the performance of the maxima test, the resultant test still does not perform satisfactorily in the real data application. However, the ratio test exhibits satisfactory performance and is generally powerful in both simulations and real data applications.

Our detection methods can figure out whether the most recent image is abnormal and enables us to keep significantly different images for further examination while removing large but informationless images regularly to save image storage space. In addition to detecting abnormal medical images, the methods may be used in air pollution detection. For example, some wildfires are human-caused due to open fire in the permitted parks. We could collect remote sensing images, especially hyperspectral images, acquired from Earth-orbiting satellites, which hold great promise in monitoring and categorizing wildfires. Hyperspectral images are characterized by having multiple layers or image planes at a specific wavelength or range of wavelengths. We may focus on some layers with fixed ranges of wavelengths that are sensitive to wildfires. The detection procedure can be applied for monitoring some special areas which may have a high risk of fires.

Although the proposed detection methods have wide applications, they have some limitations. The proposed pixel-based image comparison methods might not be reasonable when there is a systematic change along the time. For example, the image will change periodically based on the heart rate in ultrasound imaging and brain fMRI. Honestly, this limitation would hinder the application of the proposed methods in some medical imaging data sets.

Appendix

A.1. Proof of Lemma 1

We will need the Taylor series expansions of $F(x)$ and $f(x)$ about zero. The leading-term approximations of them are

$$F(x) \approx F(0) = 1/2, \quad f(x) \approx f(0).$$

Recall that $g(x)$ denotes the common density function of the medians $x_1, \ldots, x_\kappa$ under H_0. A truncated approximated density function of $g(x)$ in the interval $[a, b]$ has the following form

$$\tilde{g}(x; -1 < a < 0,\ 0 < b < 1) = \begin{cases} 1/(b-a), & \text{if } a \le x \le b, \\ 0, & \text{otherwise.} \end{cases}$$

The likelihood functions under $\tilde{H}_0$ and $\tilde{H}_1$ respectively given by

$$L_0(-a_0, a_0 | x_1, \ldots, x_\kappa) = \begin{cases} 1/(2a_0)^\kappa, & \text{if } x_1, \ldots, x_\kappa \in [-a_0,\ a_0], \\ 0, & \text{ortherwise,} \end{cases}$$

and

$$L_1(a_1, a_2 | x_1, \ldots, x_\kappa) = \begin{cases} 1/(a_2 - a_1)^\kappa, & \text{if } x_1, \ldots, x_\kappa \in [a_1,\ a_2], \\ 0, & \text{ortherwise.} \end{cases}$$

Hence, a likelihood ratio statistic is

$$\begin{aligned} \Lambda(x_1, \ldots, x_\kappa) &= \frac{\sup_{a_0}\{L_0(-a_0, a_0 | x_1, \ldots, x_\kappa), a_0 > 0\}}{\sup_{a_1, a_2}\{L_1(a_1, a_2 | x_1, \ldots, x_\kappa), a_1 < 0, a_2 > 0\}} \\ &= \left[\frac{x_{(\kappa)} - x_{(1)}}{\max\{x_{(\kappa)}, -x_{(1)}\}}\right]^\kappa. \end{aligned}$$

A.2. Proof of Theorem 4.1.

Let $u_\kappa = \mathtt{a}_\kappa + \mathtt{b}_\kappa v$. Since $M_\kappa = \max_{1 \le \ell \le \kappa} |x_\ell|/\hat{\sigma}$ (see (6)), for any $\xi > 0$, we have

$$\begin{aligned} \mathrm{pr}(M_\kappa > u_\kappa) = \mathrm{pr}(&\max_{1 \le \ell \le \kappa} |x_\ell|/\hat{\sigma} > u_\kappa, \max_{1 \le \ell \le \kappa} |\zeta_\ell| < \xi) \\ &+ \mathrm{pr}(\max_{1 \le \ell \le \kappa} |x_\ell|/\hat{\sigma} > u_\kappa, \max_{1 \le \ell \le \kappa} |\zeta_\ell| \ge \xi), \end{aligned} \tag{15}$$

where ζ_ℓ is defined in Proposition 1. Let $\xi = \kappa^{1/2}/m^{3/2}$ and $\tilde{\xi} = 2f(0)\xi$. By Proposition 1, for some positive constant c,

$$\mathrm{pr}(\max_{1 \le \ell \le \kappa} |x_\ell|/\hat{\sigma} > u_\kappa, \max_{1 \le \ell \le \kappa} |\zeta_\ell| \ge \xi) \le \mathrm{pr}(\max_{1 \le \ell \le \kappa} |\zeta_\ell| \ge \xi)$$

$$\leq \mathrm{pr}(\max_{1\leq\ell\leq\kappa}|\zeta_{\ell,1}|\geq\xi/2)+\mathrm{pr}(\max|\zeta_{\ell,2}|\geq\xi/2)$$
$$\leq c\exp\{\log(\kappa)-\xi^2m^2/8-\log(\xi m)\}+c_1\exp(-c_2m), \tag{16}$$

where $\zeta_{\ell,1}$ and $\zeta_{\ell,2}$ are defined in Proposition 1. Thus, by Assumption A2, $\exp(-c_2m)\to 0$. As $\log(\kappa)-\log(\xi m)=\log(\kappa/m)-\log(\xi)$ and $-\log(\xi)=-\frac{1}{2}\log(\kappa)+\frac{3}{2}\log(m)$, by Assumption A2, we have

$$\exp\{\log(\kappa)-\xi^2m^2/8-\log(\xi m)\}\to 0.$$

Hence,

$$\mathrm{pr}(\max_{1\leq\ell\leq\kappa}|\zeta_\ell|\geq\xi)\to 0. \tag{17}$$

We now find the limit of $\mathrm{pr}(\max_{1\leq\ell\leq\kappa}|z_\ell|>u_\kappa)$, where z_ℓ, $1\leq\ell\leq\kappa$, defined in Proposition 1, are i.i.d. $N(0,1)$.

For given $\mathtt{b}_\kappa$ and $\mathtt{a}_\kappa$ as in (11), by Davison [7, eq. 6.35], $m\{1-\Phi(\mathtt{a}_\kappa+\mathtt{b}_\kappa v)\}\to\exp(-v)$. Since $z_1,\ldots,z_\kappa$ are i.i.d. $N(0,1)$, by Davison [7, eq. 6.32], it follows that $\mathrm{pr}\{\mathtt{b}_\kappa^{-1}(\max_{1\leq\ell\leq\kappa}z_\ell-\mathtt{a}_\kappa)>v\}\to 1-\exp\{-\exp(-v)\}$. Therefore, $m\{1-\Phi(\mathtt{a}_\kappa+\mathtt{b}_\kappa v)+\Phi(-\mathtt{a}_\kappa-\mathtt{b}_\kappa v)\}=2m\{1-\Phi(\mathtt{a}_\kappa+\mathtt{b}_\kappa v)\to 2\exp(-v)$, and

$$\mathrm{pr}\{\max_{1\leq\ell\leq\kappa}|z_\ell|>u_\kappa)=\mathrm{pr}\{\mathtt{b}_\kappa^{-1}(\max_{1\leq\ell\leq\kappa}|z_\ell|-\mathtt{a}_\kappa)>v\}\to 1-\exp\{-2\exp(-v)\}. \tag{18}$$

By Assumption A2, it can be shown that $\xi=o(1/\sqrt{\log\kappa})$, $2f(0)\sqrt{m}\hat{\sigma}-1=o_p(1/\sqrt{m})=o_p(1/\log\kappa)$, which, jointly with $u_\kappa=O(\sqrt{\log\kappa})$, yield that $2f(0)\sqrt{m}\hat{\sigma}u_\kappa-2f(0)\xi-u_\kappa=o_p(1/\sqrt{\log\kappa})$. Hence, by Slutsky's theorem and (18), it follows that

$$\mathrm{pr}(\max_{1\leq\ell\leq\kappa}|x_\ell|/\hat{\sigma}>u_\kappa,\max_{1\leq\ell\leq\kappa}|\zeta_\ell|<\xi)-\mathrm{pr}(\max_{1\leq\ell\leq\kappa}|z_\ell|>u_\kappa)$$
$$\leq\mathrm{pr}(\max_{1\leq\ell\leq\kappa}|z_\ell|>2f(0)\sqrt{m}\hat{\sigma}u_\kappa-2f(0)\xi)-\mathrm{pr}(\max_{1\leq\ell\leq\kappa}|z_\ell|>u_\kappa)\to 0. \tag{19}$$

Combining (15)-(19), we obtain that

$$\overline{\lim_{n\to\infty}}\ \mathrm{pr}(M_\kappa>u_\kappa)\leq 1-\exp\{-2\exp(-v)\}. \tag{20}$$

We now show that $\underline{\lim}_{n\to\infty}\mathrm{pr}(M_\kappa>u_\kappa)\geq 1-\exp\{-2\exp(-v)\}$. By (15)

$$\mathrm{pr}(M_\kappa>u_\kappa)\geq\mathrm{pr}(\max_{1\leq\ell\leq\kappa}|x_\ell|/\hat{\sigma}>u_\kappa,\max_{1\leq\ell\leq\kappa}|\zeta_\ell|<\xi)$$
$$\geq\mathrm{pr}(\max_{1\leq\ell\leq\kappa}|z_\ell|>2f(0)\sqrt{m}\hat{\sigma}u_\kappa+2f(0)\xi,\max_{1\leq\ell\leq\kappa}|\zeta_\ell|<\xi)$$
$$=\mathrm{pr}(\max_{1\leq\ell\leq\kappa}|z_\ell|>2f(0)\sqrt{m}\hat{\sigma}u_\kappa+2f(0)\xi)$$
$$-\mathrm{pr}(\max_{1\leq\ell\leq\kappa}|z_\ell|>2f(0)\sqrt{m}\hat{\sigma}u_\kappa+2f(0)\xi,\max_{1\leq\ell\leq\kappa}|\zeta_\ell|\geq\xi). \tag{21}$$

Similar to (19), we can obtain that

$$\mathrm{pr}(\max_{1\leq\ell\leq\kappa}|z_\ell|>2f(0)\sqrt{m}\hat{\sigma}u_\kappa+2f(0)\xi)-\mathrm{pr}(\max_{1\leq\ell\leq\kappa}|z_\ell|>u_\kappa)\to 0. \tag{22}$$

It is easy to see that

$$\mathrm{pr}(\max_{1\le\ell\le\kappa}|z_\ell|>2f(0)\sqrt{m}\hat{\sigma}u_\kappa+2f(0)\xi,\max_{1\le\ell\le\kappa}|\zeta_\ell|\ge\xi)\le\mathrm{pr}(\max_{1\le\ell\le\kappa}|\zeta_\ell|\ge\xi),$$

which, jointly with (17), implies that

$$\mathrm{pr}(\max_{1\le\ell\le\kappa}|z_\ell|>2f(0)\sqrt{m}\hat{\sigma}u_\kappa+2f(0)\xi,\max_{1\le\ell\le\kappa}|\zeta_\ell|\ge\xi)\to 0. \tag{23}$$

Hence, by (21)–(23) and (18), we have $\underline{\lim}_{n\to\infty}\mathrm{pr}(M_\kappa > u_\kappa)\ge 1-\exp\{-2\exp(-v)\}$, which, jointly with (20), concludes the theorem.

A.3. Proof of Theorem 4.2.

Our aim is to show that the lower bound of M_κ diverges to infinity in probability. It is easy to see that

$$M_\kappa>\min_{1\le\ell\le\kappa}|E(x_\ell)|/\hat{\sigma}-\max_{1\le\ell\le\kappa}|x_\ell-E(x_\ell)|/\hat{\sigma}. \tag{24}$$

By Assumption A3, $\min_{1\le\ell\le\kappa}|E(x_\ell)|>\varrho>0$, which, jointly with the fact that $2f(0)\sqrt{m}\hat{\sigma}\to 1$ in probability, yields that

$$\min_{1\le\ell\le\kappa}|E(x_\ell)|/\hat{\sigma}\to\infty,\quad\text{in probability} \tag{25}$$

In light of Theorem 4.1., it can be shown that

$$\max_{1\le\ell\le\kappa}|x_\ell-E(x_\ell)|/(\hat{\sigma}\sqrt{\log\kappa})=O_p(1). \tag{26}$$

By Assumption A2, it is easy to see that $\sqrt{m}/\log\kappa\to\infty$, which, jointly with (24)-(26), concludes the theorem.

A.4. Proof of Theorem 4.3.

Write $\xi=\kappa^{1/2}/m^{3/2}$ and $\tilde{\xi}=2f(0)\xi$. Recall the random variables ζ_ℓ and z_ℓ defined in Proposition 1 in which z_ℓ, $1\le\ell\le\kappa$, are i.i.d. $N(0,1)$. Recall that $z_{(\kappa)}=\max\{z_\ell,\ \ell=1,\ldots,\kappa\}$, and $z_{(1)}=\min\{z_\ell,\ \ell=1,\ldots,\kappa\}$. We have

$$\begin{aligned}
&\mathrm{pr}\left[\frac{x_{(\kappa)}-x_{(1)}}{\min\{x_{(\kappa)},-x_{(1)}\}}\ge\nu\right]\\
&=\mathrm{pr}\left[\frac{x_{(\kappa)}-x_{(1)}}{\min\{x_{(\kappa)},-x_{(1)}\}}\ge\nu,\ \max_{1\le\ell\le\kappa}|\zeta_\ell|<\xi,\ z_{(\kappa)}>\sqrt{\xi},\ -z_{(1)}>\sqrt{\xi}\right]\\
&+\mathrm{pr}\left[\frac{x_{(\kappa)}-x_{(1)}}{\min\{x_{(\kappa)},-x_{(1)}\}}\ge\nu,\ \max_{1\le\ell\le\kappa}|\zeta_\ell|\ge\xi\text{ or }z_{(\kappa)}\le\sqrt{\xi}\ \text{ or }\ -z_{(1)}\le\sqrt{\xi}\right]\\
&\hat{=}W_1+W_2.
\end{aligned} \tag{27}$$

It is obvious that

$$W_2 \le \mathrm{pr}\left(\max_{1\le\ell\le\kappa}|\zeta_\ell| \ge \xi\right) + \mathrm{pr}\left(z_{(\kappa)} \le \sqrt{\xi}\right) + \mathrm{pr}\left(-z_{(1)} \le \sqrt{\xi}\right).$$

Note that it has been shown in (17) that $\mathrm{pr}(\max_{1\le\ell\le\kappa}|\zeta_\ell| \ge \xi) \to 0$. By Assumption A2, $\xi \to 0$ and $\kappa \to \infty$, which yield that $\mathrm{pr}(z_{(\kappa)} \le \sqrt{\xi}) = \Phi^\kappa(\sqrt{\xi}) \to 0$ and $\mathrm{pr}(-z_{(1)} \le \sqrt{\xi}) = \Phi^\kappa(\sqrt{\xi}) \to 0$. Thus, as $n \to \infty$,

$$W_2 \to 0. \tag{28}$$

We now find the limit of W_1. By (4) and (12), it is easy to see that if $\max_{1\le\ell\le\kappa}|\zeta_\ell| < \xi$, then $z_{(\kappa)} - \tilde{\xi} \le 2f(0)\sqrt{m}x_{(\kappa)} \le z_{(\kappa)} + \tilde{\xi}$ and $-z_{(1)} - \tilde{\xi} \le -2f(0)\sqrt{m}x_{(1)} \le -z_{(1)} + \tilde{\xi}$. Further, if $z_{(\kappa)} > \sqrt{\xi}$ and $-z_{(1)} > \sqrt{\xi}$, it can be shown that $\sqrt{\xi} - \tilde{\xi} \le \min\{z_{(\kappa)}, -z_{(1)}\} - \tilde{\xi} < 2f(0)\sqrt{m}\min\{x_{(\kappa)}, -x_{(1)}\}$. By Assumption A2, it follows that $\sqrt{\xi} - \tilde{\xi} > 0$ and $\tilde{\xi}/\sqrt{\xi} \to 0$. Write

$$\mathbf{z}_{1,\kappa} = \frac{z_{(\kappa)} - z_{(1)}}{\min\{z_{(\kappa)}, -z_{(1)}\}}.$$

Thus, we have,

$$\begin{aligned}
W_1 \le & \mathrm{pr}\left[\frac{z_{(\kappa)} - z_{(1)} + 2\tilde{\xi}}{\min\{z_{(\kappa)}, -z_{(1)}\} - \tilde{\xi}} \ge \nu,\ \max_{1\le\ell\le\kappa}|\zeta_\ell| < \xi,\ z_{(\kappa)} > \sqrt{\xi},\ -z_{(1)} > \sqrt{\xi}\right] \\
\le & \mathrm{pr}\left[\mathbf{z}_{1,\kappa} \ge \nu - \frac{(2+\nu)\tilde{\xi}}{\min\{z_{(\kappa)}, -z_{(1)}\}},\ z_{(\kappa)} > \sqrt{\xi},\ -z_{(1)} > \sqrt{\xi}\right] \\
\le & \mathrm{pr}\left[\mathbf{z}_{1,\kappa} \ge \nu - \frac{(2+\nu)\tilde{\xi}}{\sqrt{\xi}},\ z_{(\kappa)} > 0,\ z_{(1)} < 0\right], \\
= & \mathrm{pr}\left[z_{(\kappa)} > 0, z_{(1)} < 0\right] - \mathrm{pr}\left[\mathbf{z}_{1,\kappa} < \nu - \frac{(2+\nu)\tilde{\xi}}{\sqrt{\xi}}, z_{(\kappa)} > 0, z_{(1)} < 0\right].
\end{aligned} \tag{29}$$

Since z_ℓ, $1 \le \ell \le \kappa$, are i.i.d. $N(0,1)$, it is obvious that as $n \to \infty$,

$$\mathrm{pr}\left[z_{(\kappa)} > 0,\ z_{(1)} < 0\right] \to 1. \tag{30}$$

We now show that

$$\mathrm{pr}\left[\mathbf{z}_{1,\kappa} \le \nu - \frac{(2+\nu)\tilde{\xi}}{\sqrt{\xi}},\ z_{(\kappa)} > 0,\ z_{(1)} < 0\right]$$

has the same limit as

$$\mathrm{pr}\left[\mathbf{z}_{1,\kappa} < \nu,\ z_{(\kappa)} > 0,\ z_{(1)} < 0\right],$$

or equivalently, as $n \to \infty$,

$$\mathrm{pr}\left[\nu - \frac{(2+\nu)\tilde{\xi}}{\sqrt{\xi}} < \mathtt{z}_{1,\kappa} < \nu,\ z_{(\kappa)} > 0,\ z_{(1)} < 0\right] \to 0. \tag{31}$$

Let $\gamma > 1$. Then we have

$$\begin{aligned}
&\mathrm{pr}\left[\nu - \frac{(2+\nu)\tilde{\xi}}{\sqrt{\xi}} < \mathtt{z}_{1,\kappa} < \nu,\ z_{(\kappa)} > 0,\ z_{(1)} < 0\right] \\
&= \mathrm{pr}\left[\nu - \frac{(2+\nu)\tilde{\xi}}{\sqrt{\xi}} < \mathtt{z}_{1,\kappa} < \nu, z_{(\kappa)} > \gamma(-z_{(1)}),\ z_{(\kappa)} > 0,\ z_{(1)} < 0\right] \\
&+ \mathrm{pr}\left[\nu - \frac{(2+\nu)\tilde{\xi}}{\sqrt{\xi}} < \mathtt{z}_{1,\kappa} < \nu, -z_{(1)} > \gamma(z_{(\kappa)}),\ z_{(\kappa)} > 0,\ z_{(1)} < 0\right] \\
&+ \mathrm{pr}\left[\nu - \frac{(2+\nu)\tilde{\xi}}{\sqrt{\xi}} < \mathtt{z}_{1,\kappa} < \nu, -z_{(1)} \le z_{(\kappa)} \le \gamma(-z_{(1)}),\ z_{(\kappa)} > 0,\ z_{(1)} < 0\right] \\
&+ \mathrm{pr}\left[\nu - \frac{(2+\nu)\tilde{\xi}}{\sqrt{\xi}} < \mathtt{z}_{1,\kappa} < \nu, z_{(\kappa)} < -z_{(1)} \le \gamma(z_{(\kappa)}),\ z_{(\kappa)} > 0,\ z_{(1)} < 0\right] \\
&= \mathrm{II}_1(\gamma) + \mathrm{II}_2(\gamma) + \mathrm{II}_3(\gamma) + \mathrm{II}_4(\gamma).
\end{aligned}$$

By Davison [7, eq. 6.35], $(z_{(\kappa)} - \mathtt{a}_\kappa)/\mathtt{b}_\kappa$ converges to a Gumbel random variable, which, jointly with the fact that $\mathtt{b}_\kappa \to 0$, and $\mathtt{a}_\kappa \to \infty$, yields that $z_{(\kappa)} - \mathtt{a}_\kappa$ converges to zero in probability. Hence, we obtain that for any $\gamma > 1$, as $n \to \infty$,

$$\mathrm{II}_1(\gamma) \le \mathrm{pr}\left[z_{(\kappa)} > \gamma(-z_{(1)}), z_{(\kappa)} > 0,\ z_{(1)} < 0\right] \to 0.$$

Similarly, it can be shown that $\mathrm{II}_2(\gamma) \to 0$, as $n \to \infty$. Note that $-z_{(1)} \le z_{(\kappa)} \le \gamma(-z_{(1)}),\ z_{(\kappa)} > 0,\ z_{(1)} < 0$ imply that

$$\mathtt{z}_{1,\kappa} = \frac{z_{(\kappa)} - z_{(1)}}{-z_{(1)}} \le 1 + \gamma.$$

Choose a γ such that $1 \le \gamma < \nu - 1$ and denote it by γ_0. Since by Assumption A2, $\tilde{\xi}/\sqrt{\xi} \to 0$ as $n \to \infty$, it follows that

$$\mathrm{II}_3(\gamma_0) \le \mathrm{pr}\left[\nu - \frac{(2+\nu)\tilde{\xi}}{\sqrt{\xi}} < 1 + \gamma_0\right] \to 0, \quad \text{as } n \to \infty.$$

In a similar way, we can show that $\mathrm{II}_4(\gamma_0) \to 0$, as $n \to \infty$. Thus, (31) is proved.

Since z_ℓ, $1 \le \ell \le \kappa$, are i.i.d. $N(0,1)$, it is obvious that

$$\mathrm{pr}\left[\mathtt{z}_{1,\kappa} \le \nu, z_{(\kappa)} \le 0 \text{ or } z_{(1)} \ge 0\right] \to 0, \quad \text{as } n \to \infty.$$

Note that $\nu > 2$, $z_{(\kappa)} > 0$, and $z_{(1)} < 0$, and hence $\{z_{(\kappa)} - z_{(1)}\}/\min\{z_{(\kappa)}, -z_{(1)}\} \le \nu$ means that $z_{(\kappa)} \le (1-\nu)z_{(1)}$ if $z_{(\kappa)} > -z_{(1)}$, otherwise $(1-\nu)z_{(\kappa)} \le$

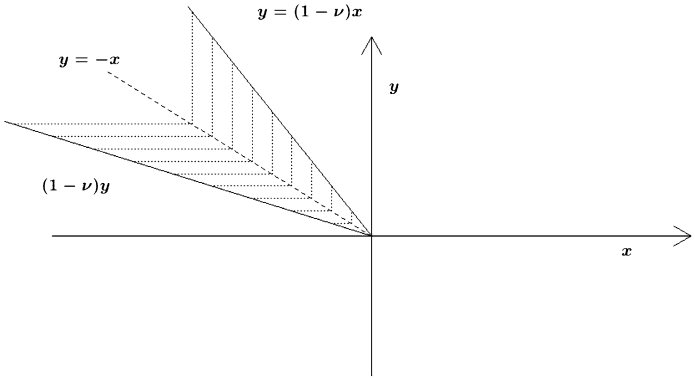

FIG 4. *Illustration of the integration domain with* $\nu = 3$ *of* (32).

$z_{(1)}$. As z_ℓ, $1 \le \ell \le \kappa$, are i.i.d. $N(0,1)$, the joint density of $z_{(1)}$ and $z_{(\kappa)}$ is $\kappa(\kappa-1)[\Phi(y)-\Phi(x)]^{\kappa-2}\phi(x)\phi(y)$ with $x \le y$. By the integration domain illustrated in Figure 4, we have

$$\begin{aligned}
&1 - \operatorname{pr}\left[\mathbf{z}_{1,\kappa} \le \nu,\ z_{(\kappa)} > 0,\ z_{(1)} < 0\right] \\
=&1 - \operatorname{pr}\left[z_{(\kappa)} \le (1-\nu)z_{(1)},\ z_{(\kappa)} > 0,\ z_{(1)} < 0,\ z_{(\kappa)} > -z_{(1)}\right] \\
&- \operatorname{pr}\left[(1-\nu)z_{(\kappa)} \le z_{(1)},\ z_{(\kappa)} > 0,\ z_{(1)} < 0,\ z_{(\kappa)} \le -z_{(1)}\right] \\
=&1 - \int_{-\infty}^{0} dx \int_{-x}^{(1-\nu)x} \kappa(\kappa-1)[\Phi(y)-\Phi(x)]^{\kappa-2}\phi(x)\phi(y)dy \\
&- \int_{0}^{\infty} dy \int_{(1-\nu)y}^{-y} \kappa(\kappa-1)[\Phi(y)-\Phi(x)]^{\kappa-2}\phi(x)\phi(y)dx \qquad (32) \\
=&1 - \int_{-\infty}^{0} \kappa\phi(x)\{[\Phi((1-\nu)x)-\Phi(x)]^{\kappa-1} - [\Phi(-x)-\Phi(x)]^{\kappa-1}\}dx \\
&+ \int_{0}^{\infty} \kappa\phi(y)\{[\Phi(y)-\Phi(-y)]^{\kappa-1} - [\Phi(y)-\Phi((1-\nu)y)]^{\kappa-1}\}dy.
\end{aligned}$$

Thus,

$$\begin{aligned}
\overline{\lim_{n\to\infty}} \Bigg\{ &\operatorname{pr}(T_\kappa \ge \nu) - 1 + \int_{-\infty}^{0} \kappa\phi(x)\{[\Phi((1-\nu)x)-\Phi(x)]^{\kappa-1} \\
&- [\Phi(-x)-\Phi(x)]^{\kappa-1}\}dx + \int_{0}^{\infty} \kappa\phi(y)\left\{[\Phi(y)-\Phi(-y)]^{\kappa-1}\right. \\
&\left.-[\Phi(y)-\Phi((1-\nu)y)]^{\kappa-1}\right\} dy \Bigg\} \le 0. \qquad (33)
\end{aligned}$$

We want to find out if we have

$$\underline{\lim_{n\to\infty}} \Bigg\{ \operatorname{pr}(T_\kappa \ge \nu) - 1 + \int_{-\infty}^{0} \kappa\phi(x)\{[\Phi((1-\nu)x)-\Phi(x)]^{\kappa-1}$$

$$- [\Phi(-x) - \Phi(x)]^{\kappa-1}\}dx + \int_0^{\infty} \kappa\phi(y) \left\{[\Phi(y) - \Phi(-y)]^{\kappa-1} - [\Phi(y) - \Phi((1-\nu)y)]^{\kappa-1}\right\} dy \Bigg\} \geq 0. \tag{34}$$

Similar to the proof of Theorem 4.1., we have

$$\begin{aligned} W_1 \geq & \operatorname{pr}\left[\frac{z_{(\kappa)} - z_{(1)} - 2\tilde{\xi}}{\min\{z_{(\kappa)}, -z_{(1)}\} + \tilde{\xi}} \geq \nu,\ \max_{1\leq\ell\leq\kappa} |\zeta_\ell| < \xi,\ z_{(\kappa)} > \sqrt{\xi},\ -z_{(1)} > \sqrt{\xi}\right] \\ = & \operatorname{pr}\left[\mathbf{z}_{1,\kappa} \geq \nu + \frac{(2+\nu)\tilde{\xi}}{\min\{z_{(\kappa)}, -z_{(1)}\}},\ \max_{1\leq\ell\leq\kappa} |\zeta_\ell| < \xi, z_{(\kappa)} > \sqrt{\xi},\ -z_{(1)} > \sqrt{\xi}\right] \\ = & \operatorname{pr}\left[\max_{1\leq\ell\leq\kappa} |\zeta_\ell| < \xi, z_{(\kappa)} > \sqrt{\xi}, -z_{(1)} > \sqrt{\xi}\right] \\ & - \operatorname{pr}\left[\mathbf{z}_{1,\kappa} < \nu + \frac{(2+\nu)\tilde{\xi}}{\min\{z_{(\kappa)}, -z_{(1)}\}}, \max_{1\leq\ell\leq\kappa} |\zeta_\ell| < \xi, z_{(\kappa)} > \sqrt{\xi}, -z_{(1)} > \sqrt{\xi}\right] \\ \geq & \operatorname{pr}\left[\max_{1\leq\ell\leq\kappa} |\zeta_\ell| < \xi, z_{(\kappa)} > \sqrt{\xi}, -z_{(1)} > \sqrt{\xi}\right] - \operatorname{pr}\left[\mathbf{z}_{1,\kappa} < \nu + \frac{(2+\nu)\tilde{\xi}}{\sqrt{\xi}}\right]. \end{aligned} \tag{35}$$

By the fact that $\operatorname{pr}\left(\max_{1\leq\ell\leq\kappa} |\zeta_\ell| \geq \xi\right) \to 0$, $\operatorname{pr}\left(z_{(\kappa)} \leq \sqrt{\xi}\right) \to 0$, and $\operatorname{pr}\left(-z_{(1)} \leq \sqrt{\xi}\right) \to 0$ as $n \to \infty$, it follows that

$$\operatorname{pr}\left[\max_{1\leq\ell\leq\kappa} |\zeta_\ell| < \xi,\ z_{(\kappa)} > \sqrt{\xi},\ -z_{(1)} > \sqrt{\xi}\right] \to 1 \quad \text{as } n \to \infty. \tag{36}$$

Similar to the proof of (31), we obtain that as $n \to \infty$,

$$\operatorname{pr}\left[\mathbf{z}_{1,\kappa} \leq \nu + \frac{(2+\nu)\tilde{\xi}}{\sqrt{\xi}}\right] - \operatorname{pr}\left[\mathbf{z}_{1,\kappa} \leq \nu\right] \to 0. \tag{37}$$

Since

$$\begin{aligned} & \operatorname{pr}\left[\mathbf{z}_{1,\kappa} \leq \nu\right] \\ = & \operatorname{pr}\left[\mathbf{z}_{1,\kappa} \leq \nu, z_{(\kappa)} > 0, z_{(1)} < 0\right] + \operatorname{pr}\left[\mathbf{z}_{1,\kappa} \leq \nu, z_{(\kappa)} \leq 0 \text{ or } z_{(1)} \geq 0\right], \\ & \operatorname{pr}\left[\mathbf{z}_{1,\kappa} \leq \nu, z_{(\kappa)} \leq 0 \text{ or } z_{(1)} \geq 0\right] \to 0, \end{aligned}$$

and

$$\begin{aligned} & 1 - \operatorname{pr}\left[\mathbf{z}_{1,\kappa} \leq \nu, z_{(\kappa)} > 0, z_{(1)} < 0\right] \\ & = 1 - \int_{-\infty}^{0} \kappa\phi(x)\{[\Phi((1-\nu)x) - \Phi(x)]^{\kappa-1} - [\Phi(-x) - \Phi(x)]^{\kappa-1}\}dx \\ & + \int_0^{\infty} \kappa\phi(y)\{[\Phi(y) - \Phi(-y)]^{\kappa-1} - [\Phi(y) - \Phi((1-\nu)y)]^{\kappa-1}\}dy, \end{aligned}$$

by (35)-(37), it follows that (34) holds true, which, jointly with (33), concludes the theorem.

A.5. Proof of Theorem 4.4.

We first assume that δ_qs are all less than $-\varrho < 0$ by Assumptions A1 and A4. We compare $-x_{(1)}/\hat{\sigma}$ and $x_{(\kappa)}/\hat{\sigma}$. Since $x_\ell = \theta_\ell + \eta_\ell$ with $E(x_\ell) = \theta_\ell$ and $\max_{1\le\ell\le\kappa}(-\theta_\ell) = \max_{1\le\ell\le\kappa}(-\delta_\ell) > \varrho$,

$$\begin{aligned} -x_{(1)}/\hat{\sigma} &= -\min_{1\le\ell\le\kappa}(\theta_\ell/\hat{\sigma} + \eta_\ell/\hat{\sigma}) \\ &= \max_{1\le\ell\le\kappa}(-\theta_\ell/\hat{\sigma} - \eta_\ell/\hat{\sigma}) \ge \max_{1\le\ell\le\kappa}(-\theta_\ell)/\hat{\sigma} - \max_{1\le\ell\le\kappa}|\eta_\ell|/\hat{\sigma}, \end{aligned} \tag{38}$$

and

$$x_{(\kappa)}/\hat{\sigma} = \max_{1\le\ell\le\kappa}(\theta_\ell/\hat{\sigma} + \eta_\ell/\hat{\sigma}) \le \max_{1\le\ell\le\kappa}|\eta_\ell|/\hat{\sigma}. \tag{39}$$

Since $2f(0)\sqrt{m}\hat{\sigma} \to 1$ in probability, by (26) and Assumption 2, we have

$$1/\hat{\sigma} \gg \max_{1\le\ell\le\kappa}|x_\ell - E(x_\ell)|/\hat{\sigma} \quad \text{in probability,} \tag{40}$$

where $A_\kappa \gg B_\kappa$ means A/B goes to infinity when $\kappa \to \infty$. Thus, by (38)-(40), we obtain that

$$-x_{(1)}/\hat{\sigma} \gg x_{(\kappa)}/\hat{\sigma} \quad \text{in probability,}$$

which implies that in probability,

$$\begin{aligned} T_\kappa &= \frac{x_{(\kappa)} - x_{(1)}}{\min\{x_{(\kappa)}, -x_{(1)}\}} = 1 + \frac{\max\{x_{(\kappa)}, -x_{(1)}\}}{\min\{x_{(\kappa)}, -x_{(1)}\}} \\ &= 1 + \frac{\max\{x_{(\kappa)}/\hat{\sigma}, -x_{(1)}/\hat{\sigma}\}}{\min\{x_{(\kappa)}/\hat{\sigma}, -x_{(1)}/\hat{\sigma}\}} = 1 + \frac{-x_{(1)}/\hat{\sigma}}{x_{(\kappa)}/\hat{\sigma}} \to \infty. \end{aligned}$$

We now assume that δ_qs are all larger than $\varrho > 0$ by Assumptions A1 and A4. Similar to the derivations of (38)-(39), it can be shown that

$$-x_{(1)}/\hat{\sigma} = -\min_{1\le\ell\le\kappa}(\theta_\ell/\hat{\sigma} + \eta_\ell/\hat{\sigma}) = \max_{1\le\ell\le\kappa}(-\theta_\ell/\hat{\sigma} - \eta_\ell/\hat{\sigma}) \le \max_{1\le\ell\le\kappa}|\eta_\ell|/\hat{\sigma},$$

and

$$x_{(\kappa)}/\hat{\sigma} = \max_{1\le\ell\le\kappa}(\theta_\ell/\hat{\sigma} + \eta_\ell/\hat{\sigma}) \ge \max_{1\le\ell\le\kappa}\theta_\ell/\hat{\sigma} - \max_{1\le\ell\le\kappa}|\eta_\ell|/\hat{\sigma},$$

which, jointly with (40), yield that $x_{(\kappa)}/\hat{\sigma} \gg -x_{(1)}/\hat{\sigma}$ in probability. Thus,

$$T_\kappa = 1 + \frac{x_{(\kappa)}/\hat{\sigma}}{-x_{(1)}/\hat{\sigma}} \to \infty \quad \text{in probability.}$$

Acknowledgments

We would like to thank Professor Nancy Reid and Professor Anthony Davison for their helpful comments on the first draft.

References

[1] Antile, A., Kersting, G., and Zucchini, W. (1982). Testing asymmetry. *Journal of the American Statistical Association* **77**, 639–46. MR0675891

[2] Asadi, P., Davison, A. C. and Engelke, S. (2015). Extremes on river networks. *The Annals of Applied Statistics* **9**, 2023–50. MR3456363

[3] Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society: Series B* **36**, 192–236. MR0373208

[4] Bhattacharya, P. K., Gastwirth, J. L., and Wright, A. L. (1982). Two modified Wilcoxon tests for symmetry about an unknown location parameter. *Biometrika* **69**, 377–82. MR0671975

[5] Cabilio, P. and Masaro, J. (1996). A simple test of symmetry about an unknown median. *Canadian Journal of Statistics* **24**, 349–61.

[6] Cai, T.T., Jeng, J. and Li, H. (2012). Robust detection and identification of sparse segments in ultrahigh dimensional data analysis. *Journal of the Royal Statistical Society: Series B* **74**, 773–97. MR2988906

[7] Davison, A.C. (2003). *Statistical Models*. Cambridge: Cambridge University Press. MR1998913

[8] Davison, A. C., Padoan, S. A. and Ribatet, M. (2012). Statistical Modeling of Spatial Extremes. *Statistical Science* **9**, 161–86. MR2963980

[9] Delgado, M. A., and Song, X. (2018). Nonparametric tests for conditional symmetry. *Journal of Econometrics*, **206**(2), 447–71. MR3850990

[10] Drikvandi, R., Modarres, R., and Jalilian, A.H. (2011). A bootstrap test for symmetry based on ranked set samples. *Computational statistics & data analysis*, **55**(4), 1807–14. MR2748681

[11] Ekström, M., and Jammalamadaka, S.R. (2007). An asymptotically distribution-free test of symmetry. *Journal of statistical planning and inference*, **137**(3), 799–810. MR2301716

[12] Ferreira, A. and de Haan, L. (2015). On the block maxima method in extreme value theory: PWM estimators. *The Annals of Statistics* **43**, 276–98. MR3285607

[13] Gastwirth, J. L. (1971). On the sign test for symmetry. *Journal of the American Statistical Association* **66**, 821–3.

[14] Grady, E. (2016). Gastrointestinal bleeding scintigraphy in the early 21st century. *Journal of Nuclear Medicine* **57**, 252–9.

[15] Gretton, A., Fukumizu, K., Teo, C.H., Song, L., Schölkopf, B. and Smola, A.J. (2008). A kernel statistical test of independence. *Advances in neural information processing systems*, 585–92.

[16] Gumbel, E. J. (1958). *Statistics of Extremes*. New York: Columbia Univ. Press. MR0096342

[17] Huber, P.J. (1981). *Robust Statistics*. New York: John Wiley and Sons. MR0606374

[18] Jairath, V., Hearnshaw, S., Brunskill, S.J., Doree, C., Hopewell, S., Hyde, C.and Travis, S. and Murphy, M.F.(2010). Red cell transfu-

sion for the management of upper gastrointestinal haemorrhage. *Cochrane Database Syst Rev.* **9**, CD006613.

[19] Li, T., and Yuan, M. (2019).On the optimality of Gaussian kernel based nonparametric tests against smooth alternatives. arXiv preprint arXiv:1909.03302.

[20] Lyubchich, V., Wang, X., Heyes, A., and Gel, Y.R. (2016). A distribution-free m-out-of-n bootstrap approach to testing symmetry about an unknown median. *Computational Statistics & Data Analysis*, **104**, 1–9. MR3540982

[21] McWilliams, T. P. (1990). A distribution-free test for symmetry based on a runs statistic. *Journal of the American Statistical Association* **85**, 1130–3. MR1134510

[22] Mira, A. (1999). Distribution-free test for symmetry based on Bonferroni's measure. *Journal of Applied Statistics*, **26**(8), 959–72. MR1731875

[23] Ou, S and Yi, D. (2009). *Robustness analysis and algorithm of expected shortfall based on extreme-value block minimum model*. International Conference on Business Intelligence and Financial Engineering.

[24] Pitas, I. (2000). *Digital Image Processing Algorithms and Applications*. New York: Wiley. MR1272249

[25] Psaradakis, Z. (2016). Using the bootstrap to test for symmetry under unknown dependence. *Journal of Business & Economic Statistics*, **34**(3), 406–15. MR3523784

[26] Psaradakis, Z., and Vavra, M. (2019). Bootstrap-assisted tests of symmetry for dependent data. *Journal of Statistical Computation and Simulation*, **89**(7), 1203–26. MR3928457

[27] Quackenbush, J. (2002). Microarray data normalization and transformation. *Nature Genetics* **32**, 496–501.

[28] Randles, R. H., Fligner, M. A., Policello, G. E., and Wolfe, D. A. (1980). An asymptotically distribution-free test for symmetry versus asymmetry. *Journal of the American Statistical Association* **75**, 168–72. MR0568588

[29] Shi, X., Wu, Y. and Rao, C. R. (2018). Consistent and powerful non-Euclidean graph-based change-point test with applications to segmenting random interfered video data. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 5914–9. MR3827579

[30] Xing, X., Shang, Z., Du, P., Ma, P., Zhong, W., and Liu, J.S. (2019) Minimax nonparametric two-sample test. arXiv preprint arXiv:1911.02171. MR4119162

[31] Zheng, T., and Gastwirth, J.L. (2010). On bootstrap tests of symmetry about an unknown median. *Journal of data science: JDS*, **8**(3), 413.