Statistica Sinica Preprint No: SS-2024-0051

Title	Spatial-Sign Based Maxsum Test for High Dimensional Location Parameters
Manuscript ID	SS-2024-0051
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202024.0051
Complete List of Authors	Jixuan Liu, Long Feng, Ping Zhao and Zhaojun Wang
Corresponding Authors	Long Feng
E-mails	flnankai@nankai.edu.cn
Notice: Accepted author version.	

Statistica Sinica

Statistica Sinica: Newly accepted Paper

Statistica Sinica

SPATIAL-SIGN BASED MAXSUM TEST FOR HIGH DIMENSIONAL LOCATION PARAMETERS

Jixuan Liu, Long Feng, Ping Zhao and Zhaojun Wang

School of Statistics and Data Science, KLMDASR, LEBPS, and LPMC, Nankai University

Abstract: In this study, we explore a robust testing procedure for the high-dimensional location parameters testing problem. Initially, we introduce a spatial-sign based max-type test statistic, which exhibits excellent performance for sparse alternatives. Subsequently, we demonstrate the asymptotic independence between this max-type test statistic and the spatial-sign based sum-type test statistic (Feng and Sun, 2016). Building on this, we propose a spatial-sign based max-sum type testing procedure, which shows remarkable performance under varying signal sparsity. Our simulation studies underscore the superior performance of the procedures we propose.

Key words and phrases: Asymptotic independence, High dimensional data, Scalar-invariant, Spatial-sign.

Statistica Sinica

Statistica Sinica: Newly accepted Paper

1. Introduction

The testing of location parameters is a crucial and extensively researched area in multivariate statistics with a fixed dimension. The conventional Hotelling's T^2 test is commonly applied, but it fails in high-dimensional scenarios where the variable's dimension p exceeds the sample sizes n. Consequently, numerous efforts have been made to develop a high-dimensional mean test procedure. One straightforward approach is to substitute the Mahalanobis distance with the Euclidean distance. For the two-sample location problem, Bai and Saranadasa (1996) employed the L_2-norm of the difference between two sample means. Chen and Qin (2010) eliminated some redundant terms in Bai and Saranadasa (1996)'s test statistics and made no assumptions about the relationship between the dimension and sample sizes. Srivastava (2009), Park and Ayyala (2013), and Feng et al. (2015) suggested some scalar-invariant test statistics that replace the sample covariance matrix in Hotelling's T^2 test statistics with its diagonal matrix. All these methods are built on the assumption of normal distribution or diverging factor models, which perform poorly for heavy-tailed distributions. For instance, the well-known multivariate t-distribution does not meet the above assumption. Therefore, numerous studies have also considered robust high-dimensional test procedures.

Statistica Sinica: Newly accepted Paper

In traditional multivariate analysis, numerous methods have been developed to extend classic univariate rank and signed rank techniques to a multivariate context. A significant method is based on spatial signs and ranks, utilizing the so-called Oja median (Oja, 2010). Wang et al. (2015) proposed a high-dimensional spatial-sign test that replaces the scatter matrix with the identity matrix for a one-sample location problem. Similarly, Feng and Sun (2016) proposed a high-dimensional spatial sign test that replaces the scatter matrix with its diagonal matrix, which has a scalar-invariant property. Furthermore, Feng et al. (2016) considered the high-dimensional two-sample location problem based on the spatial-sign method. Feng, Liu, and Ma (2021) devised an inverse norm sign test that considers not only the direction of the observations but also the modulus of the observation. Huang et al. (2023) extended the inverse norm sign test for a high-dimensional two-sample location problem. Feng et al. (2020) demonstrated that the spatial-rank method also performs well for a high-dimensional two-sample problem. All these methods are constructed using the L_2-norm of the spatial median, which performs well under dense alternatives, meaning many variables have non-zero means. However, it is well-known that these sum-type test procedures perform poorly for sparse alternatives, where only a few variables have non-zero means.

Statistica Sinica: Newly accepted Paper

In high-dimensional settings, numerous max-type test procedures have been introduced to detect sparse alternatives. Cai et al. (2013) proposed a test statistic for the high-dimensional two-sample mean problem, which is based on the maximum difference between the means of two samples' variables under the Gaussian or sub-Gaussian assumption. For heavy-tailed distributions, Cheng et al. (2023) established a Gaussian approximation for the sample spatial median over the class of hyperrectangles and constructed a max-type test procedure using a multiplier bootstrap algorithm. However, their proposed test statistic is not scalar-invariant, and they did not provide the limit null distribution of their test statistic. The multiple bootstrap algorithm is also time-consuming. In this paper, we first introduce a novel Bahadur representation of the scaled sample spatial median and then construct a new max-type test statistic. We demonstrate that the limit null distribution of the proposed test statistic is still a Type I Gumbel distribution. We also establish the consistency of the proposed max-type test procedure. Simulation studies further illustrate its superiority over existing methods under sparse alternatives and heavy-tailed distributions.

In practical scenarios, it's often unknown whether the alternative is dense or sparse. This has led to numerous studies proposing an adaptive strategy that combines the sum-type test and max-type test. For high-

Statistica Sinica: Newly accepted Paper

dimensional mean problems, Xu et al. (2016) integrated different L_r-norms of the sample means. He et al. (2021) introduced a family of U-statistics as an unbiased estimator of the L_r-norm of the mean vectors, covariance matrices and regression coefficients, demonstrating that U-statistics of different finite orders are asymptotically independent, normally distributed, and also independent from the maximum-type test statistic. Feng et al. (2022) relaxed the covariance matrix assumption to establish independence between the sum-type test statistic and the max-type test statistic. There are also many other studies showing the asymptotic independence between the sum-type test statistics and the max-type test statistic for other high-dimensional problems. For instance, Wu et al. (2019) and Wu et al. (2020) examined the coefficient test in high-dimensional generalized linear models. Feng et al. (2022) looked at the cross-sectional independence test in high-dimensional panel data models. Yu et al. (2022) focused on testing the high-dimensional covariance matrix. Feng et al. (2022) considered the high-dimensional white noise test, while Wang and Feng (2023) looked at high-dimensional change point inference. Ma et al. (2023) and Yu et al. (2023) considered testing the alpha of high-dimensional time-varying and linear factor pricing models, respectively.

However, all these methods assume a normal or other light-tailed dis-

Statistica Sinica: Newly accepted Paper

tributions. There's a gap in the literature when it comes to considering the asymptotic independence between the sum-type test statistic and the max-type test statistic under heavy-tailed distributions. In this paper, we first establish the asymptotic independence between Feng and Sun (2016)'s sum-type test statistic and a newly proposed spatial sign-based max-type test statistic for high dimensional one sample location parameter problem. We then propose a Cauchy combination test procedure (Liu and Xie, 2020) to handle general alternatives. Both simulation studies and theoretical results demonstrate the advantages of our newly proposed methods.

This paper are organized as follow. Section 2 introduce Bahadur representation of the scaled spatial median and establish the max-type test statistic. In section 3, we prove the asymptotic independence between the sum-type test statistic and the new proposed max-type test statistic and construct the Cauchy combination test procedure. Section 4 show some simulation studies. Some discussion are stated in Section 5. Two real data applications and all the technical details are in the Supplementary Material.

Notations: For d-dimensional $\boldsymbol{x}$, we use the notation $\|\boldsymbol{x}\|$ and $\|\boldsymbol{x}\|_\infty$ to denote its Euclidean norm and maximum-norm respectively. Denote $a_n \lesssim b_n$ if there exists constant C, $a_n \leq Cb_n$ and $a_n \asymp b_n$ if both $a_n \lesssim b_n$ and $b_n \lesssim a_n$ hold. Let $\psi_\alpha(x) = \exp(x^\alpha) - 1$ be a function defined on $[0, \infty)$

Statistica Sinica: Newly accepted Paper

for $\alpha > 0$. Then the Orlicz norm $\|\cdot\|_{\psi_\alpha}$ of a $\boldsymbol{X}$ is defined as $\|\boldsymbol{X}\|_{\psi_\alpha} = \inf\{t > 0, \mathbb{E}\{\psi_\alpha(|\boldsymbol{X}|/t)\} \leqslant 1\}$. Let $\mathrm{tr}(\cdot)$ be a trace for matrix, $\lambda_{min}(\cdot)$ and $\lambda_{max}(\cdot)$ be the minimum and maximum eigenvalue for symmetric martix. $\mathbf{I}_p$ represents a p-dimensional identity matrix, and $\mathrm{diag}\{v_1, v_2, \cdots, v_p\}$ represents the diagonal matrix with entries $\boldsymbol{v} = (v_1, v_2, \cdots, v_p)$. For $a, b \in \mathbb{R}$, we write $a \wedge b = \min\{a, b\}$.

2. Max-type test

Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be a sequence of independent and identically distributed (i.i.d.) p-dimensional random vectors from a population X with cumulative distribution function F_X in $\mathbb{R}^p$. We consider the following model:

$$\boldsymbol{X}_i = \boldsymbol{\theta} + v_i \boldsymbol{\Gamma} \boldsymbol{W}_i,$$

where $\boldsymbol{\theta}$ is location parameter, $\boldsymbol{W}_i$ is a p-dimensional random vector with independent components, $E(\boldsymbol{W}_i) = 0$, $\boldsymbol{\Sigma} = \boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top$, v_i is a nonnegative univariate random variable and is independent with the spatial sign of $\boldsymbol{W}_i$. The distribution of $\boldsymbol{X}$ depends on $\boldsymbol{\Gamma}$ through the shape matrix. Model (2) encompasses a wide range of frequently utilized multivariate models and distribution families, such as the independent components model (Nordhausen et al., 2009; Ilmonen and Paindaveine, 2011; Yao et al., 2015) and the family of elliptical distributions (Hallin and Paindaveine, 2006; Oja,

Statistica Sinica: Newly accepted Paper

2010; Fang, 2018).

In this paper, we focus on the following one sample testing problem

$$H_0 : \boldsymbol{\theta} = \mathbf{0} \text{ versus } \quad H_1 : \boldsymbol{\theta} \neq \mathbf{0}. \tag{2.1}$$

The spatial sign function is defined as $U(\mathbf{x}) = \|\mathbf{x}\|^{-1}\mathbf{x}I(\mathbf{x} \neq \mathbf{0})$. In traditional fixed p circumstance, the following so-called "inner centering and inner standardization" sign-based procedure is usually used (cf., Chapter 6 of Oja (2010)), with statistic $Q_n^2 = np\overline{\boldsymbol{U}}^\top \overline{\boldsymbol{U}}$, where$\overline{\boldsymbol{U}} = \frac{1}{n}\sum_{i=1}^n \hat{\boldsymbol{U}}_i, \hat{\boldsymbol{U}}_i = U\left(\mathbf{S}^{-1/2}\boldsymbol{X}_i\right), \mathbf{S}^{-1/2}$ are Tyler's scatter matrix (cf., Section 6.1.3 of Oja (2010)). Q_n^2 is affine-invariant and can be regarded as a nonparametric counterpart of Hotelling's T^2 test statistic by using the spatial-signs instead of the original observations $\boldsymbol{X}_i$'s. However, when $p > n, Q_n^2$ is not defined as the matrix $\mathbf{S}^{-1/2}$ is not available in high-dimensional settings.

In high-dimensional settings, Wang et al. (2015) proposed a method where Tyler's scatter matrix is replaced by the identity matrix. This led to the test statistic $T_{WPL} = \sum_{i<j} U^\top(\boldsymbol{X}_i)U(\boldsymbol{X}_j)$. Building on this, Feng and Sun (2016) extended the method and introduced a scalar-invariant spatial-sign based test procedure, which will be detailed in section 3. Both methods utilize sum-type test statistics, which perform well under dense alternatives where many elements of $\boldsymbol{\theta}$ are nonzeros. However, their power decreases under sparse alternatives where only a few elements of $\boldsymbol{\theta}$ are nonzeros.

Statistica Sinica: Newly accepted Paper

It is well-known that max-type tests have good performance under sparse alternatives (Cai et al., 2013). Therefore, Cheng et al. (2023) first provided the Bahadur representation of the classic spatial median $\widetilde{\boldsymbol{\theta}}$, defined as $\widetilde{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \sum_{i=1}^{n} ||\mathbf{X}_i - \boldsymbol{\theta}||$. They then proposed a max-type test procedure based on Gaussian approximation. While this approach is robust and effective in high-dimensional settings, it loses scalar information of different variables and is not scalar-invariant. In real-world scenarios, different components may have entirely different physical or biological readings, and their scales would not be identical. Moreover, due to the unequal scale of $\widetilde{\boldsymbol{\theta}}$, it is not possible to derive the limited null distribution of $||\widetilde{\boldsymbol{\theta}}||_\infty$ even under weak correlation assumption. In this paper, we first provide the Bahadur representation and Gaussian approximation of the location estimator proposed in Feng et al. (2016). We then propose a new max-type test statistic and establish its limit null distribution under some mild conditions.

2.1 Bahadur representation and Gaussian approximation

Motivated by Feng et al. (2016), we suggest to find a pair of diagonal matrix $\mathbf{D}$ and vector $\boldsymbol{\theta}$ for each sample that simultaneously satisfy

$$\frac{1}{n}\sum_{i=1}^{n} U\left(\boldsymbol{\epsilon}_i\right) = 0 \text{ and } \frac{p}{n}\operatorname{diag}\left\{\sum_{i=1}^{n} U\left(\boldsymbol{\epsilon}_i\right) U\left(\boldsymbol{\epsilon}_i\right)^\top\right\} = \mathbf{I}_p, \tag{2.2}$$

Statistica Sinica: Newly accepted Paper

where $\boldsymbol{\epsilon}_i = \mathbf{D}^{-1/2}(\boldsymbol{X}_i - \boldsymbol{\theta})$. $(\mathbf{D},\ \boldsymbol{\theta})$ can be viewed as a simplified version of HettmanspergerRandles (HR) estimator , which is proposed in Hettmansperger and Randles (2002), without considering the off-diagonal elements of $\mathbf{S}$. We can adapt the recursive algorithm of Feng et al. (2016) to solve Equation 2.2. That is, repeat the following three steps until convergence:

(i) $\boldsymbol{\epsilon}_i \leftarrow \mathbf{D}^{-1/2}(\mathbf{X}_i - \boldsymbol{\theta}), \quad j = 1, \cdots, n;$

(ii) $\boldsymbol{\theta} \leftarrow \boldsymbol{\theta} + \frac{\mathbf{D}^{1/2}\sum_{j=1}^{n} U(\boldsymbol{\epsilon}_i)}{\sum_{j=1}^{n}\|\epsilon_i\|^{-1}};$

(iii) $\mathbf{D} \leftarrow p\mathbf{D}^{1/2}\,\mathrm{diag}\left\{n^{-1}\sum_{j=1}^{n} U(\boldsymbol{\epsilon}_i)\,U(\boldsymbol{\epsilon}_i)^\top\right\}\mathbf{D}^{1/2}.$

The resulting estimators of location and diagonal matrix are denoted as $\hat{\boldsymbol{\theta}}$ and $\hat{\mathbf{D}}$. The sample mean and sample variances can be utilized as initial estimators. Regrettably, no evidence has been found to confirm the convergence of the aforementioned algorithm, even in low-dimensional scenarios, despite its consistent practical effectiveness. The existence or uniqueness of the HR estimator mentioned above also lacks proof. This topic certainly warrants further investigation.

In this section, we investigate some theoretical properties based on maximum-norm about $\hat{\boldsymbol{\theta}}$. Similar to the proof of Lemma 1 and Theorem 1 in Cheng et al. (2023), we give the Bahadur representation of $\hat{\mathbf{D}}^{-1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ and the Gaussian approximation result for $\hat{\mathbf{D}}^{-1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ over hyperrectangles. Based on Gaussian approximation, we can easily derive the limiting

Statistica Sinica: Newly accepted Paper

distribution of $\hat{\boldsymbol{\theta}}$ based on the maximum-norm.

For $i = 1, 2, \cdots, n$, we denote $\boldsymbol{U}_i = U(\mathbf{D}^{-1/2}(\boldsymbol{X}_i - \boldsymbol{\theta}))$ and $R_i = \|\mathbf{D}^{-1/2}(\boldsymbol{X}_i - \boldsymbol{\theta})\|$ as the scale-invariant spatial-sign and radius of $\boldsymbol{X}_i - \boldsymbol{\theta}$, respectively. The moments of R_i is defined as $\zeta_k = \mathbb{E}\left(R_i^{-k}\right)$ for $k = 1, 2, 3, 4$. Denote $\boldsymbol{W}_i = (W_{i,1}, \ldots, W_{i,p})^\top$, the assumption is as follows.

Assumption 1. $W_{i,1}, \ldots, W_{i,p}$ are i.i.d. symmetric random variables with $\mathbb{E}(W_{i,j}) = 0, \mathbb{E}\left(W_{i,j}^2\right) = 1$, and $\|W_{i,j}\|_{\psi_\alpha} \leqslant c_0$ with some constant $c_0 > 0$ and $1 \leqslant \alpha \leqslant 2$.

Assumption 2. The moments $\zeta_k = \mathbb{E}\left(R_i^{-k}\right)$ for $k = 1, 2, 3, 4$ exist for large enough p. In addition, there exist two positive constants $\underline{b}$ and $\bar{B}$ such that $\underline{b} \leqslant \limsup_p \mathbb{E}\left(R_i/\sqrt{p}\right)^{-k} \leqslant \bar{B}$ for $k = 1, 2, 3, 4$.

Assumption 3. The shape matrix $\mathbf{R} = \mathbf{D}^{-1/2}\boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top\mathbf{D}^{-1/2} = (\sigma_{j\ell})_{p\times p}$ satisfies $\mathrm{tr}(\mathbf{R}) = p$ and $\max_{j=1,\cdots,p}\sum_{\ell=1}^p |\sigma_{j\ell}| \leqslant a_0(p)$ where $a_0(p)$ is a constant depending only on dimension p. In addition, $\liminf_{p\to\infty}\min_{j=1,2,\cdots,p} d_j > \underline{d}$ for some constant $\underline{d} > 0$, where $\mathbf{D} = \mathrm{diag}\{d_1^2, d_2^2, \cdots, d_p^2\}$.

Remark 1 Assumption 1 is the same as Condition C.1 in Cheng et al. (2023), which ensure that $\boldsymbol{\theta}$ in model (2) is the population spatial median and $W_{i,j}$ has a sub-exponential distribution. If $\boldsymbol{W}_i \sim N(\mathbf{0}, \mathbf{I}_p)$, $\boldsymbol{X}_i$ follows a elliptical symmetric distribution. Assumption 2 extend the Assumption 1 in

Statistica Sinica: Newly accepted Paper

Zou et al. (2014), which indicates that $\zeta_k \asymp p^{-k/2}$. It is a mild assumption and introduced to avoid $\boldsymbol{X}_i$ from concentrating too much near $\boldsymbol{\theta}$. For three commonly used distribution, multivariate normal, student-t and mixtures of multivariate normal distributions, Assumptions 1-2 are satisfied. For example, for standard multivariate normal distribution, $\mathbb{E}R_i^{-4}$ equals to $1/(p-2)(p-4))$ which restricts the dimension $p > 4$. See also discussions in Zou et al. (2014); Cheng et al. (2023) on similar assumptions. Assumption 3 means the correlation between those variables could not be too large, which is similar to the matrix class in Bickel and Levina (2008).

The following lemma shows a Bahadur representation of $\hat{\boldsymbol{\theta}}$, which is the basis of Gaussian approximation result in Theorem 1.

Lemma 1. *(Bahadur representation) Assume Assumptions 1-3 with $a_0(p) \asymp p^{1-\delta}$ for some positive constant $\delta \leq 1/2$ hold. If $\log p = o(n^{1/3})$ and $\log n = o(p^{1/3\wedge\delta})$, then*

$$n^{1/2}\hat{\mathbf{D}}^{-1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) = n^{-1/2}\zeta_1^{-1}\sum_{i=1}^{n}\boldsymbol{U}_i + C_n,$$

where $\|C_n\|_\infty = O_p\{n^{-1/4}\log^{1/2}(np)+p^{-(1/6\wedge\delta/2)}\log^{1/2}(np)+n^{-1/2}(\log p)^{1/2}\log^{1/2}(np)\}$.

Remark 2 Feng et al. (2016) derived the Bahadur representation of the estimator $\hat{\boldsymbol{\theta}}$, where the remainder term $||C_n||$ is $o_p(\zeta_1^{-1})$, assuming a symmetric elliptical distribution. In this context, we provide the rate of the

Statistica Sinica: Newly accepted Paper

remainder term C_n subject to a maximum-norm constraint. It's important to note that in this Lemma, we scale the spatial-median estimator $\hat{\boldsymbol{\theta}}$ by $\hat{\mathbf{D}}^{-1/2}$. This is a departure from much of the existing literature on the Bahadur representation of the spatial median, which does not exhibit scalar invariance. Such works include Zou et al. (2014), Cheng et al. (2019), Li and Xu (2022), and Cheng et al. (2023).

Let $\mathcal{A}^{\mathrm{re}} = \left\{\prod_{j=1}^{p}\left[a_j, b_j\right] : -\infty \leqslant a_j \leqslant b_j \leqslant \infty, j=1,\ldots,p\right\}$ be the class of rectangles in $\mathbb{R}^p$. Based on the Bahadur representation of $\hat{\boldsymbol{\theta}}$, we acquire the following Gaussian approximation of $\hat{\mathbf{D}}^{-1/2}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})$ in rectangle $\mathcal{A}^{\mathrm{re}}$.

Lemma 2. *(Gaussian approximation) Assume Assumptions 1-3 with* $a_0(p) \asymp p^{1-\delta}$ *for some positive constant* $\delta \leqslant 1/2$ *hold. If* $\log p = o\left(n^{1/5}\right)$ *and* $\log n = o\left(p^{1/3\wedge\delta}\right)$, *then*

$$\rho_n\left(\mathcal{A}^{\mathrm{re}}\right) = \sup_{A\in\mathcal{A}^{\mathrm{re}}}\left|\mathbb{P}\left\{n^{1/2}\hat{\mathbf{D}}^{-1/2}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)\in A\right\} - \mathbb{P}(\boldsymbol{G}\in A)\right| \to 0,$$

as $n\to\infty$, *where* $\boldsymbol{G}\sim N\left(0,\zeta_1^{-2}\boldsymbol{\Sigma}_u\right)$ *with* $\boldsymbol{\Sigma}_u = \mathbb{E}\left(\boldsymbol{U}_1\boldsymbol{U}_1^{\top}\right)$.

The Gaussian approximation for $\hat{\boldsymbol{\theta}}$ indicates that the probabilities $\mathbb{P}\left\{n^{1/2}\hat{\mathbf{D}}^{-1/2}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)\in A\right\}$ can be approximated by that of a centered Gaussian random vector with covariance matrix $\zeta_1^{-2}\boldsymbol{\Sigma}_u$ for hyperrectangles $A\in\mathcal{A}^{\mathrm{re}}$. Since the region $\mathcal{A}^{t} = \left\{\prod_{j=1}^{p}\left[a_j, b_j\right] : -\infty = a_j \leqslant b_j = t, j=1,\ldots,p\right\}$ used in the following corollary is contained in the set $\mathcal{A}^{\mathrm{re}}$, it is clear that the Corollary 1 follows.

Statistica Sinica: Newly accepted Paper

Corollary 1. Under the assumptions of Lemma 2, as $n \to \infty$, we have

$$\rho_n = \sup_{t \in \mathbb{R}} \left| \mathbb{P}\left(n^{1/2} \|\hat{\mathbf{D}}^{-1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\|_\infty \leqslant t \right) - \mathbb{P}\left(\|\boldsymbol{G}\|_\infty \leqslant t \right) \right| \to 0,$$

where $\boldsymbol{G} \sim N\left(0, \zeta_1^{-2} \boldsymbol{\Sigma}_u\right)$.

Taking into account the relationships between $\boldsymbol{\Sigma}_u$ and $p^{-1}\mathbf{R}$, we propose a more straightforward Gaussian approximation.

Lemma 3. *(Variance approximation) Suppose* $\boldsymbol{G} \sim N\left(0, \zeta_1^{-2} \boldsymbol{\Sigma}_u\right)$ *and* $\boldsymbol{Z} \sim N\left(0, \zeta_1^{-2} p^{-1} \mathbf{R}\right)$*, under the assumptions of Lemma 2, as* $(n, p) \to \infty$*, we have*

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P}\left(\|\boldsymbol{Z}\|_\infty \leqslant t \right) - \mathbb{P}\left(\|\boldsymbol{G}\|_\infty \leqslant t \right) \right| \to 0.$$

By integrating Corollary 1 and Lemma 3, we can readily derive the principal theorem of Gaussian approximation.

Theorem 1. *Assume Assumptions 1-3 with* $a_0(p) = p^{1-\delta}$ *for some positive constant* $\delta \leqslant 1/2$ *hold. If* $\log p = o\left(n^{1/5}\right)$ *and* $\log n = o\left(p^{1/3 \wedge \delta}\right)$*, then*

$$\widetilde{\rho}_n = \sup_{t \in \mathbb{R}} \left| \mathbb{P}\left(n^{1/2} \|\hat{\mathbf{D}}^{-1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\|_\infty \leqslant t \right) - \mathbb{P}\left(\|\boldsymbol{Z}\|_\infty \leqslant t \right) \right| \to 0,$$

where $\boldsymbol{Z} \sim N\left(0, \zeta_1^{-2} p^{-1} \mathbf{R}\right)$.

2.2 Max-type test procedure

In order to guarantee that the maximum value of a sequence of normal variables adheres to a Gumbel limiting distribution, we introduce Assumption

Statistica Sinica: Newly accepted Paper

4, which is employed to specify the necessary correlation among variables.

Assumption 4. For some $\varrho \in (0,1)$, assume $|\sigma_{ij}| \le \varrho$ for all $1 \le i < j \le p$ and $p \ge 2$. Suppose $\{\delta_p; p \ge 1\}$ and $\{\kappa_p; p \ge 1\}$ are positive constants with $\delta_p = o(1/\log p)$ and $\kappa = \kappa_p \to 0$ as $p \to \infty$. For $1 \le i \le p$, define $B_{p,i} = \{1 \le j \le p; |\sigma_{ij}| \ge \delta_p\}$ and $C_p = \{1 \le i \le p; |B_{p,i}| \ge p^{\kappa}\}$. We assume that $|C_p|/p \to 0$ as $p \to \infty$.

Remark 3 Assumption 4 aligns with Assumption (2.2) in Feng et al. (2022). This assumption stipulates that for each variable, the count of other variables that exhibit a strong correlation with it cannot be excessively large. To the best of our understanding, this is the least restrictive assumption in the literature that allows for the limiting null distribution of the maximum of correlated normal random variables to follow a Gumbel distribution. Both Assumption 3 and 4 pertain to the correlation matrix $\mathbf{R}$. We examine two specific cases that satisfy both of these conditions. The first case is the classic AR(1) structure, denoted as $\mathbf{R} = (\rho^{|i-j|})_{1 \le i,j \le p}, \rho \in (-1,1)$. In this scenario, $\sum_{l=1}^{p} |\sigma_{jl}| \to \frac{1}{1+\rho}$, which allows δ to be one in Assumption 3. For Assumption 4, we set $\delta_p = (\log p)^{-2}$, leading to $B_{p,i} = \{j : |i-j| \le -2\log\log p/\log|\rho|\}$. As a result, $|B_{p,i}| \le -4\log\log p/\log|\rho| < p^{\kappa}$ with $\kappa = 5\log\log p/\log p$, which implies $|C_p| = 0$ and Assumption 4 is satisfied. The second case

Statistica Sinica: Newly accepted Paper

involves a banded correlation matrix, where $\sigma_{ij}=0$ if $|i-j|>\ell$. Here, $\sum_{j=1}^{p}|\sigma_{ij}|=O(\ell)$ and $|B_{p,i}|\le \ell$ for $\delta_p=(\log p)^{-2}$. Therefore, Assumptions 3 and 4 will hold if $\ell=o(p^{\kappa})$ for any $\kappa\to 0$.

Suppose Assumption 1-4 hold, by the Theorem 2 in Feng et al. (2022), we can see that $p\zeta_1^2\max_{1\le i\le p}Z_i^2-2\log p+\log\log p$ converges to a Gumbel distribution with cdf $F(x)=\exp\left\{-\frac{1}{\sqrt{\pi}}e^{-x/2}\right\}$ as $p\to\infty$. In combining with the Theorem 1 we can conclude that,

$$\mathbb{P}\left(n^{1/2}\left\|\hat{\mathbf{D}}^{-1/2}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})\right\|_{\infty}^{2}p\zeta_1^2-2\log p+\log\log p\le x\right)\to\exp\left\{-\frac{1}{\sqrt{\pi}}e^{-x/2}\right\}. \tag{2.3}$$

Next we replace $E\left(R^{-1}\right)$ with its estimators. We denote $\hat{R}_i=\|\hat{\mathbf{D}}^{-1/2}(\mathbf{X}_i-\hat{\boldsymbol{\theta}})\|$. Then the estimator is defined as $\hat{\zeta}_1:=\frac{1}{n}\sum_{i=1}^{n}\hat{R}_i^{-1}$, and the proof of consistency is shown in Lemma 3 in Supplementary Material. Because the convergence rate of maximum is very slow, we propose a adjust max-type test statistic which based on the scalar-invariant spatial median $\hat{\boldsymbol{\theta}}$,

$$T_{MAX}=n\left\|\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}}\right\|_{\infty}^{2}\hat{\zeta}_1^2p\cdot\left(1-n^{-1/2}\right).$$

Theorem 2. *Suppose the assumptions in Theorem 1 and Assumption 4 hold. Under the null hypothesis, as $(n,p)\to\infty$, we have*

$$P\left(T_{MAX}-2\log p+\log\log p\le x\right)\to\exp\left\{-\frac{1}{\sqrt{\pi}}e^{-x/2}\right\}.$$

According to Theorem 2, a level- α test will then be performed through

Statistica Sinica: Newly accepted Paper

rejecting H_0 when $T_{MAX} - 2\log p + \log\log p$ exceeds the $(1-\alpha)$ quantile $q_{1-\alpha} = -\log\pi - 2\log\log(1-\alpha)^{-1}$ of the Gumbel distribution $F(x)$. The following theorem shows the consistency of the proposed max-type test.

Theorem 3. *Suppose the Assumptions in Theorem 2 hold. For any given* $\alpha \in (0,1)$, *if* $\|\boldsymbol{\theta}\|_\infty \geq \widetilde{C} n^{-1/2}\{\log p - 2\log\log(1-\alpha)^{-1}\}^{1/2}$ *for some large enough constant* $\widetilde{C}$, *as* $n \to \infty$, *we have*

$$\mathbb{P}(T_{MAX} - 2\log p + \log\log p > q_{1-\alpha} \mid H_1) \to 1.$$

Given a fixed significant level α, the test T_{MAX} attains consistency if $\|\boldsymbol{\theta}\|_\infty \geq \widetilde{C}\sqrt{\log p/n}$, provided that $\widetilde{C}$ is sufficiently large. This is the minimax rate optimal for testing against sparse alternatives, as stated in Theorem 3 of Cai et al. (2013). If $\widetilde{C}$ is adequately small, then it becomes impossible for any α-level test to reject the null hypothesis with a probability approaching one. Therefore, Theorem 3 also demonstrates the optimality of our proposed test T_{MAX}.

To show the high dimensional asymptotic relative efficiency, we consider a special alternative hypothesis:

$$H_1 : \boldsymbol{\theta} = (\theta_1, 0, \cdots, 0)^\top, \theta_1 > 0,$$

which means there are only one variable has nonzero mean. Let $x_\alpha =$

Statistica Sinica: Newly accepted Paper

$2\log p - \log\log p + q_{1-\alpha}$. In this case,

$$\mathbb{P}\left(\hat{d}_1^{-2}\hat{\theta}_1^2 n\hat{\zeta}_1^2 p \geq x_\alpha\right) \leq \mathbb{P}\left(T_{MAX} \geq x_\alpha\right) \leq \mathbb{P}\left(\hat{d}_1^{-2}\hat{\theta}_1^2 n\hat{\zeta}_1^2 p \geq x_\alpha\right) + \mathbb{P}\left(\max_{2\leq i\leq p} \hat{d}_i^{-2}\hat{\theta}_i^2 n\hat{\zeta}_1^2 p \geq x_\alpha\right).$$

Under this special alternative hypothesis, we can easily have

$$\mathbb{P}\left(\max_{2\leq i\leq p} \hat{d}_i^{-2}\hat{\theta}_i^2 n\hat{\zeta}_1^2 p \geq x_\alpha\right) \to \alpha, \text{ and } \mathbb{P}\left(\hat{d}_1^{-2}\hat{\theta}_1^2 n\hat{\zeta}_1^2 p \geq x_\alpha\right) \to \Phi\left(-\sqrt{x_\alpha} + (np)^{1/2} d_1^{-1}\theta_1\zeta_1\right).$$

So the power function of our proposed T_{MAX} test is

$$\beta_{MAX}(\boldsymbol{\theta}) \in (\Phi\left(-\sqrt{x_\alpha} + (np)^{1/2} d_1^{-1}\theta_1\zeta_1\right), \Phi\left(-\sqrt{x_\alpha} + (np)^{1/2} d_1^{-1}\theta_1\zeta_1\right) + \alpha).$$

Similarly, the power function of Cai et al. (2013)'s test is

$$\beta_{CLX}(\boldsymbol{\theta}) \in (\Phi\left(-\sqrt{x_\alpha} + n^{1/2}\varsigma_1^{-1}\theta_1\right), \Phi\left(-\sqrt{x_\alpha} + n^{1/2}\varsigma_1^{-1}\theta_1\right) + \alpha),$$

where ς_i^2 is the variance of $X_{ki}, i = 1, \cdots, p$. Thus, the asymptotic relative efficiency of T_{MAX} with respective to Cai et al. (2013)'s test could be approximated as $ARE(T_{MAX}, T_{CLX}) = \{E(R_i^{-1})\}^2 E(R_i^2)$, which is the same as the asymptotic relative efficiency of Feng and Sun (2016)'s test with respective to Srivastava (2009)'s test. If $\boldsymbol{X}_i$ are generated from standard multivariate t-distribution with ν degrees of freedom ($\nu > 2$),

$$ARE(T_{MAX}, T_{CLX}) = \frac{2}{\nu - 2}\left(\frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)}\right)^2.$$

For different $\nu = 3, 4, 5, 6$, the above ARE are $2.54, 1.76, 1.51, 1.38$, respectively. Under the multivariate normal distribution ($\nu = \infty$), our T_{MAX} test is the same powerful as Cai et al. (2013)'s test. However, our T_{MAX} test is much more powerful under the heavy-tailed distributions.

Statistica Sinica: Newly accepted Paper

3. Maxsum test

By Feng et al. (2016), we have the following theorem and assumptions:

Assumption 5. Variables $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n\}$ in the n-th row are independently and identically distributed (i.i.d.) from p-variate elliptical distribution with density functions $\det(\boldsymbol{\Sigma})^{-1/2}\, g\left(\left\|\boldsymbol{\Sigma}^{-1/2}(\mathbf{x}-\boldsymbol{\theta})\right\|\right)$ where $\boldsymbol{\theta}$'s are the symmetry centers and $\boldsymbol{\Sigma}$'s are the positive definite symmetric $p \times p$ scatter matrices.

Assumption 6. $\operatorname{tr}(\mathbf{R}^4) = o(\operatorname{tr}^2(\mathbf{R}^2))$.

Assumption 7. (i) $\operatorname{tr}(\mathbf{R}^2) - p = o(n^{-1}p^2)$, (ii) $n^{-2}p^2/\operatorname{tr}(\mathbf{R}^2) = O(1)$ and $\log p = o(n)$.

Remark 4 Assumption 6 is a common condition for sum-type test statistic in high dimensions, see Chen and Qin (2010); Feng et al. (2016); Feng and Sun (2016), which requires that the eigenvalues of $\mathbf{R}$ not diverge excessively. If all the eigenvalues of $\mathbf{R}$ are bounded, $\operatorname{tr}(\mathbf{R}^2) = O(p), \operatorname{tr}(\mathbf{R}^4) = O(p)$. So the Assumption 6 holds trivially. In this case, Assumption 7 becomes $p = O(n^2)$ and $p/n \to \infty$. Actually, it is not necessary for the eigenvalues to be bounded. For b unbounded eigenvalues with respect dimension p, the sufficient condition for Assumption 6 and 7 are $\lambda_{(p)}/\lambda_{(1)} = o((p-b)^{1/2}b^{-1/4})$ and $\lambda_i = 1 + o(p^{1/2}n^{-1/2})$. When the assumptions about the spectrum of $\mathbf{R}$

Statistica Sinica: Newly accepted Paper

do not hold, there would be a bias therm in sum-type statistic that difficult to calculate and deserves to be investigated further, see the Supplemental Material of Feng et al. (2016) for more details.

The following Lemma restate the Theorem 1 in Feng and Sun (2016), which gives the asymptotic null distribution of T_{SUM} under the symmetric elliptical distribution assumption.

Lemma 4. *Under Assumptions 5-7. and H_0, as $(p,n)\to\infty$, $T_{SUM}/\sigma_n \xrightarrow{d} N(0,1)$, where $\sigma_n^2 = \frac{2}{n(n-1)p^2}\operatorname{tr}(\mathbf{R}^2)$.*

To broaden the application, we re-derive the limiting null distribution of T_{SUM} under a more generalized model (2).

Theorem 4. *Under Assumptions 1-3,6-7 and H_0, as $(p,n)\to\infty$, $T_{SUM}/\sigma_n \xrightarrow{d} N(0,1)$.*

Similar to Feng and Sun (2016), we propose the following estimator to estimate the trace term in σ_n^2

$$\widehat{\operatorname{tr}(\mathbf{R}^2)} = \frac{p^2}{n(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\left(U\left(\hat{\mathbf{D}}_{ij}^{-1/2}\left(\boldsymbol{X}_i-\hat{\boldsymbol{\theta}}_{ij}\right)\right)^T U\left(\hat{\mathbf{D}}_{ij}^{-1/2}\left(\boldsymbol{X}_j-\hat{\boldsymbol{\theta}}_{ij}\right)\right)\right)^2,$$

where $\left(\hat{\boldsymbol{\theta}}_{ij},\hat{\mathbf{D}}_{ij}\right)$ are the corresponding spatial median and diagonal matrix estimators using leave-two-out sample $\{\boldsymbol{X}_k\}_{k\neq i,j}^n$. Similar to the proof of Proposition 2 in Feng et al. (2016), we can easily obtain that $\widehat{\operatorname{tr}(\mathbf{R}^2)}/\operatorname{tr}(\mathbf{R}^2) \xrightarrow{p}$

Statistica Sinica: Newly accepted Paper

1 as $(p, n) \to \infty$ under model (2). Consequently, a ratio-consistent estimator of σ_n^2 under H_0 is $\hat{\sigma}_n^2 = \frac{2}{n(n-1)p^2}\widehat{\operatorname{tr}(\mathbf{R}^2)}$. We reject the null hypothesis with α level of significance if $T_{SUM}/\hat{\sigma}_n > z_\alpha$, where z_α is the upper α quantile of $N(0,1)$.

And we also re-derive the asymptotic distribution of T_{SUM} under the following alternative hypothesis:

$$H_1 : \boldsymbol{\theta}^\top \mathbf{D}^{-1} \boldsymbol{\theta} = O\left(\zeta_1^{-2}\sigma\right) \text{ and } \boldsymbol{\theta}^\top \mathbf{R} \boldsymbol{\theta} = o\left(\zeta_1^{-2} np\sigma^2\right) \tag{3.1}$$

Theorem 5. *Under Assumptions 1-3,6-7 and the alternative hypothesis (3.1), as* $(p, n) \to \infty$,

$$\frac{T_{SUM} - \zeta_1^2 \boldsymbol{\theta}^\top \mathbf{D}^{-1} \boldsymbol{\theta}}{\sigma_n} \xrightarrow{d} N(0,1).$$

By Theorem 4-5, the power function of T_{SUM} can be approximated as

$$\beta_{SUM}(\boldsymbol{\theta}) = \Phi\left(-z_\alpha + \frac{\zeta_1^2 np \boldsymbol{\theta}^T \mathbf{D}^{-1} \boldsymbol{\theta}}{\sqrt{2\operatorname{tr}(\mathbf{R}^2)}}\right).$$

Hence, T_{SUM} is expected to perform well under the dense alternative hypothesis. For a more detailed discussion on the asymptotic relative efficiency of T_{SUM} compared to other tests, refer to Feng and Sun (2016). The power comparison between T_{SUM} and T_{MAX} will be addressed in the following subsection.

Statistica Sinica: Newly accepted Paper

3.1 Maxsum test

In this subsection, we demonstrate that our proposed test statistic T_{MAX} is asymptotically independent of the statistic T_{SUM} presented in Feng and Sun (2016). This allows us to carry out a Cauchy p-value combination of the two asymptotically independent p-values, resulting in a new test. This test is tailored to accommodate both sparse and dense alternatives.

Assumption 8. There exist $C > 0$ and $\varrho \in (0, 1)$ so that $\max_{1 \le i < j \le p} |\sigma_{ij}| \le \varrho$ and $\max_{1 \le i \le p} \sum_{j=1}^{p} \sigma_{ij}^2 \le (\log p)^C$ for all $p \ge 3$; $p^{-1/2}(\log p)^C \ll \lambda_{\min}(\mathbf{R}) \le \lambda_{\max}(\mathbf{R}) \ll \sqrt{p}(\log p)^{-1}$ and $\lambda_{\max}(\mathbf{R})/\lambda_{\min}(\mathbf{R}) = O(p^{\tau})$ for some $\tau \in (0, 1/4)$.

Remark 5 Assumption 8 is the same as the condition (2.3) in Feng et al. (2022). As shown in Feng et al. (2022), Assumption 8 is more restrictive than Assumption 3, 4 and 6. Under Assumption 8, we have $p^{1/2}(\log p)^C \lesssim \mathrm{tr}(\mathbf{R}^2) \lesssim p^{3/2}\log^{-1} p$. So Assumption 7 will hold if $n = o(p^{3/2}\log p)$ and $p^{3/4}(\log p)^{-C/2} = O(n)$. Intuitively speaking, if all the eigenvalues of $\mathbf{R}$ are bounded and $p/n \to c \in (0, \infty)$, all the assumptions 3, 4, 6, 7 and 8 hold.

Theorem 6. *Under Assumptions 1-4, 7-8 and H_0, T_{SUM} and T_{MAX} are asymptotically independent, i.e., as $(n, p) \to \infty$,*

$$\mathbb{P}\left(T_{SUM}/\sigma_n \le x, T_{MAX} - 2\log p + \log\log p \le y\right) \to \Phi(x)F(y).$$

Statistica Sinica: Newly accepted Paper

According to Theorem 6, we suggest combining the corresponding p-values by using Cauchy Combination Method (Liu and Xie, 2020), to wit,

$$p_{CC} = 1 - G[0.5\tan\{(0.5 - p_{MAX})\pi\} + 0.5\tan\{(0.5 - p_{SUM})\pi\}],$$

$$p_{MAX} = 1 - F(T_{MAX} - 2\log p + \log\log p),\ \ p_{SUM} = 1 - \Phi(T_{SUM}/\hat{\sigma}_n),$$

where $G(\cdot)$ is the CDF of the standard Cauchy distribution. If the final p-value is less than a pre-specified significant level $\alpha \in (0,1)$, we reject H_0.

Next, we consider the relationship between T_{SUM} and T_{MAX} under local alternative hypotheses:

$$H_1 : \|\boldsymbol{\theta}\| = O(\zeta_1^{-2}\sigma), \|\mathbf{R}^{1/2}\boldsymbol{\theta}\| = o\left(\zeta_1^{-2}np\sigma^2\right) \text{ and } |\mathcal{A}| = o\left(\frac{\lambda_{\min}(\mathbf{R})[\operatorname{tr}(\mathbf{R}^2)]^{1/2}}{(\log p)^C}\right), \tag{3.2}$$

where $\mathcal{A} = \{i \mid \theta_i \neq 0, 1 \leq i \leq p\}$, $\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_p)^\top$. The following theorem establish the asymptotic independence between T_{SUM} and T_{MAX} under this special alternative hypothesis.

Theorem 7. *Under Assumptions 1-4, 6-8 and the alternative hypothesis (3.2), T_{SUM} and T_{MAX} are asymptotically independent, i.e., as $(n,p) \to \infty$*

$$\mathbb{P}\left(T_{SUM}/\sigma_n \leq x, T_{MAX} - 2\log p + \log\log p \leq y\right) \to$$
$$\mathbb{P}\left(T_{SUM}/\sigma_n \leq x\right)\mathbb{P}\left(T_{MAX} - 2\log p + \log\log p \leq y\right).$$

According to Long et al. (2023), the Cauchy combination-based test has more power than the test based on the minimum of p_{MAX} and p_{SUM}, which

Statistica Sinica: Newly accepted Paper

is also known as the minimal p-value combination. This is represented as $\beta_{M\wedge S,\alpha} = P(\min\{\mathrm{p}_{MAX}, \mathrm{p}_{SUM}\} \le 1 - \sqrt{1-\alpha})$. It is clear that:

$$\begin{aligned}\beta_{M\wedge S,\alpha} &\ge P(\min\{\mathrm{p}_{MAX}, \mathrm{p}_{SUM}\} \le \alpha/2) \\ &= \beta_{MAX,\alpha/2} + \beta_{SUM,\alpha/2} - P(\mathrm{p}_{MAX} \le \alpha/2, \mathrm{p}_{SUM} \le \alpha/2) \\ &\ge \max\{\beta_{MAX,\alpha/2}, \beta_{SUM,\alpha/2}\}. \end{aligned} \tag{3.3}$$

On the other hand, under the local alternative hypothesis (3.2), we have,

$$\beta_{M\wedge S,\alpha} \ge \beta_{MAX,\alpha/2} + \beta_{SUM,\alpha/2} - \beta_{MAX,\alpha/2}\beta_{SUM,\alpha/2} + o(1), \tag{3.4}$$

which is due to the asymptotic independence implied by Theorem 7.

For a small α, the difference between $\beta_{MAX,\alpha}$ and $\beta_{MAX,\alpha/2}$ is small, and the same applies to $\beta_{SUM,\alpha}$. Therefore, by equations (3.3) and (3.4), the power of the adaptive test is at least as large as, or even significantly larger than, that of either the max-type or sum-type test. For a detailed comparison of the performance of each test type under varying conditions of sparsity and signal strength, please refer to Table 1 in Ma et al. (2024).

4. Simulation

In this section, we incorporated various methods into our study:

- the proposed test T_{MAX}, referred as SS-MAX;

Statistica Sinica: Newly accepted Paper

- sum-type test proposed by Feng and Sun (2016), referred as SS-SUM;
- the proposed test T_{CC}, referred as SS-CC;
- max-type method proposed by Cai et al. (2013) , referred as MAX;
- sum-type method proposed by Srivastava (2009), referred as SUM;
- combination test proposed by Feng et al. (2022), referred as COM.

The following scenarios are firstly considered.

(I) Multivariate normal distribution. $\mathbf{X}_i \sim N(\boldsymbol{\theta}, \boldsymbol{\Sigma})$.

(II) Multivariate t-distribution $t_{p,4}$. $\mathbf{X}_i$'s are generated from $t_{p,4}$ with location parameter $\boldsymbol{\theta}$ and scatter matrix $\boldsymbol{\Sigma}$.

(III) Multivariate mixture normal distribution $\text{MN}_{p_n,\gamma,9}$. $\mathbf{X}_i$'s are generated from $\gamma f_{p_n}(\boldsymbol{\theta}, \boldsymbol{\Sigma}) + (1-\gamma) f_{p_n}(\boldsymbol{\theta}, 9\boldsymbol{\Sigma})$, denoted by $\text{MN}_{p_n,\gamma,9}$, where $f_{p_n}(\cdot;\cdot)$ is the density function of p_n-variate multivariate normal distribution. γ is chosen to be 0.8.

Here we consider the scatter matrix $\boldsymbol{\Sigma} = (0.5^{|i-j|})_{1\le i,j\le p}$. Two sample sizes $n = 50, 100$ and three dimensions $p = 200, 400, 600$ are considered. All the findings in this section are derived from 1000 repetitions. Table 1 presents the empirical sizes of the six tests mentioned above. It was observed that

Statistica Sinica: Newly accepted Paper

the spatial-sign based tests–SS-MAX, SS-SUM, and SS-CC–are able to effectively manage the empirical sizes in a majority of scenarios. Under the normality assumption, the Type I error of the MAX method increases as the ratio p/n grows. This may be due to the component $n\bar{x}_i/\hat{\sigma}_{ii}$ of the max statistic following a $t(n)$ distribution, which deviates from the normal distribution. When p is fixed and n increases, leading to a closer approximation to the normal distribution, the MAX method exhibits improved control over the Type I error. When dealing with distributions that are not multivariate normal, the SUM test tends to have empirical sizes that fall below the nominal level. Similarly, the COM test also exhibits smaller sizes when operating under non-normal distributions. In contrast, the SS-MAX method is more robust to variations in data distributions.

To compare the power performance of each test, we consider $\boldsymbol{\theta} = (\kappa, \kappa, \kappa, 0, \cdots, 0)$ where the first s components of $\boldsymbol{\theta}$ are all equal to $\kappa = \sqrt{0.5/s}$. Figure 4 illustrates the power curves for each test. In the case of the multivariate normal distribution, SS-SUM and SUM exhibit similar performance, aligning with the findings of Feng and Sun (2016). The spatial-sign based max-type test procedure, SS-MAX, is slightly less powerful than its mean-based counterpart, MAX. The two combination type test procedures demonstrate comparable performance in this scenario. However,

Statistica Sinica: Newly accepted Paper

Table 1: Empirical size comparison of various tests with a nominal level 5%.

	$n=50$			$n=100$		
p	200	400	600	200	400	600
	Multivariate Normal Distribution					
SS-MAX	0.051	0.061	0.049	0.025	0.04	0.032
SS-SUM	0.061	0.056	0.041	0.06	0.059	0.068
SS-CC	0.071	0.065	0.048	0.057	0.056	0.056
MAX	0.095	0.125	0.116	0.052	0.081	0.072
SUM	0.076	0.086	0.054	0.069	0.064	0.081
COM	0.095	0.108	0.089	0.063	0.072	0.058
	Multivariate t_3 Distribution					
SS-MAX	0.063	0.062	0.063	0.061	0.063	0.058
SS-SUM	0.067	0.053	0.064	0.062	0.064	0.053
SS-CC	0.058	0.061	0.068	0.058	0.052	0.059
MAX	0.044	0.052	0.047	0.033	0.036	0.04
SUM	0.005	0.001	0.001	0.002	0.001	0
COM	0.021	0.03	0.027	0.019	0.014	0.019
	Multivariate Mixture Normal Distribution					
SS-MAX	0.056	0.061	0.07	0.037	0.037	0.044
SS-SUM	0.066	0.05	0.051	0.054	0.058	0.061
SS-CC	0.067	0.058	0.064	0.052	0.046	0.059
MAX	0.037	0.042	0.056	0.031	0.038	0.03
SUM	0.004	0	0	0.007	0.002	0
COM	0.021	0.019	0.028	0.013	0.02	0.01

when dealing with non-normal distributions, the spatial-sign based test procedures surpass the mean-based ones. Moreover, the newly proposed test, SS-CC, outperforms the others in most scenarios. In extremely sparse scenarios ($s < 5$), SS-CC's performance is akin to SS-MAX. In highly dense scenarios ($s > 10$), SS-CC performs similarly to SS-SUM. However, when the signal is neither very sparse nor very dense, SS-CC proves to be the most effective among all test procedures. This underscores the superiority of our proposed max-sum procedures, not only in handling signal sparsity but also in dealing with heavy-tailed distributions.

Statistica Sinica: Newly accepted Paper

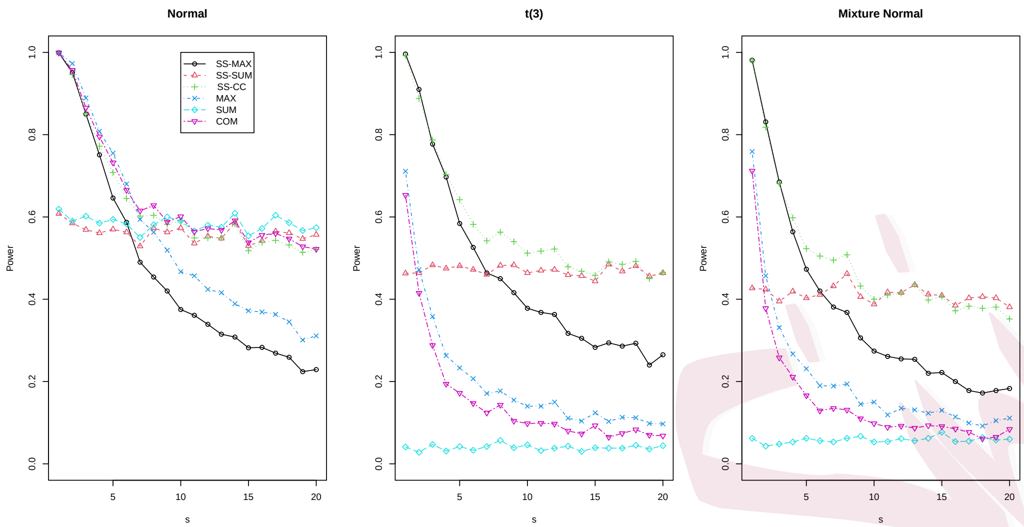

Figure 1: Power of tests with different sparsity levels over $(n, p) = (100, 200)$.

Next, we consider the power comparison of those tests under different signal strength. Here we consider three sparsity level $s = 2, 20, 50$ and the signal parameter $\kappa = \sqrt{\delta/s}$. Figures 2-4 present the power curves for various testing methods under Scenarios I to III. As the signal strength increases, the power of all tests also increases. Despite the presence of heavy-tailed distributions, spatial-sign based testing methods continue to surpass those based on means. Among all the tests, the proposed CC test consistently delivers the best performance.

As shown in Feng and Sun (2016), for the sum-type test procedure, SS-SUM is more powerful than the non scalar-invariant test (Wang et al., 2015). Here we also compare our proposed test T_{MAX} with Cheng et al. (2023)'s

Statistica Sinica: Newly accepted Paper

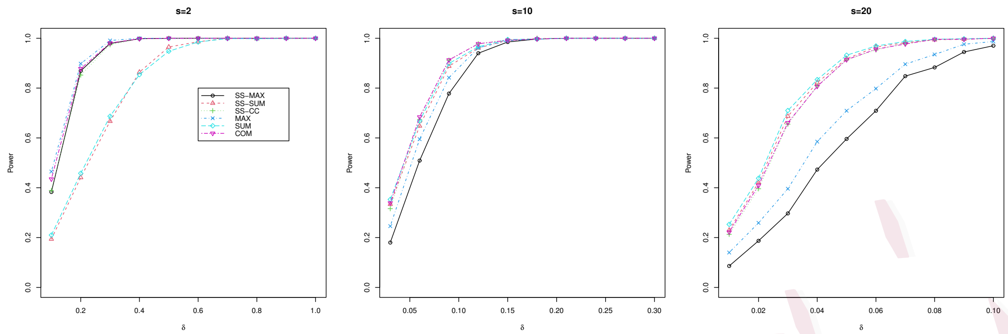

Figure 2: Power of tests with different signal strength for multivariate normal distribution over $(n, p) = (100, 200)$.

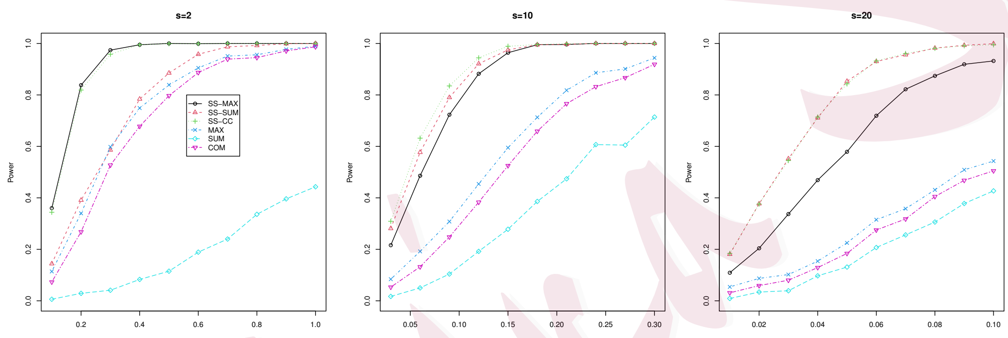

Figure 3: Power of tests with different signal strength for multivariate t_3 distribution over $(n, p) = (100, 200)$.

test (abbreviated as CPZ hereafter) to show the importance of scalar-invariant for max-type test procedure. We consider two scatter matrix case for $\mathbf{\Sigma}$: (i) $\mathbf{\Sigma} = (0.5^{|i-j|})_{1\le i,j\le p}$; (ii) $\mathbf{\Sigma} = \mathbf{D}^{1/2}\mathbf{R}\mathbf{D}^{1/2}, \mathbf{R} = (0.5^{|i-j|})_{1\le i,j\le p}, \mathbf{D} = \mathrm{diag}\{d_1, \cdots, d_p\}$ where $d_i = 1, i \le p/2, d_i = 3, i > p/2$. The other settings are all the same as above. Table 2 presents the empirical sizes of the SS-MAX and CPZ tests. Both tests are capable of controlling the empirical

Statistica Sinica: Newly accepted Paper

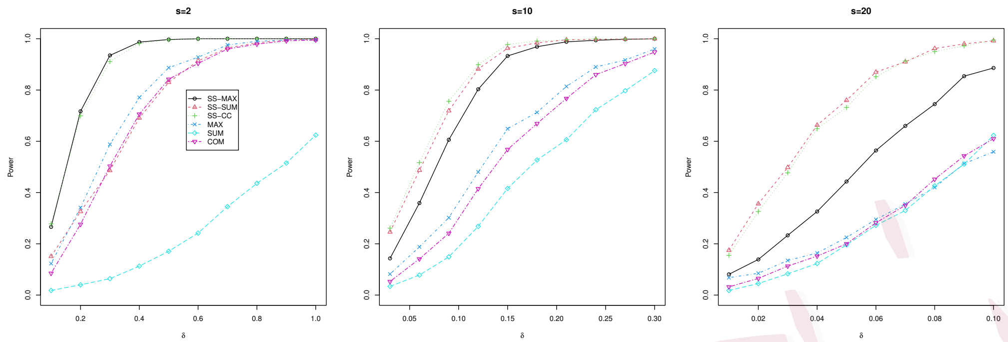

Figure 4: Power of tests with different signal strength for multivariate mixture normal distribution over $(n, p) = (100, 200)$.

sizes in the majority of cases. Moreover, we conduct a power comparison of these two max-type tests under identical settings as previously mentioned, but with two distinct scatter matrix cases. Figures 5 and 6 depict the power curves of SS-MAX and CPZ under scatter matrix cases (i) and (ii), respectively. We observe that SS-MAX performs comparably to CPZ when all elements of the diagonal matrix of $\mathbf{\Sigma}$ are equal. However, SS-MAX exhibits greater power than CPZ when the elements of the diagonal matrix of the scatter matrix are unequal, underscoring the necessity of scalar-invariance.

5. Conclusion

In this paper, we address a one-sample testing problem in high-dimensional settings for heavy-tailed distributions. We begin by providing a Bahadur

Statistica Sinica: Newly accepted Paper

Table 2: Empirical size comparison of SS-MAX and CPZ tests with a nominal level 5%.

	$n=50$			$n=100$			$n=50$			$n=100$		
p	200	400	600	200	400	600	200	400	600	200	400	600
	Scatter Matrix Case (i)						Scatter Matrix Case (ii)					
	Multivariate Normal Distribution											
SS-MAX	0.046	0.04	0.054	0.03	0.036	0.058	0.038	0.054	0.032	0.044	0.044	0.032
CPZ	0.076	0.054	0.076	0.056	0.08	0.088	0.07	0.052	0.074	0.058	0.056	0.05
	Multivariate t_3 Distribution											
SS-MAX	0.06	0.09	0.096	0.06	0.054	0.054	0.064	0.08	0.068	0.058	0.068	0.064
CPZ	0.06	0.064	0.06	0.052	0.07	0.038	0.07	0.05	0.064	0.064	0.086	0.066
	Multivariate Mixture Normal Distribution											
SS-MAX	0.072	0.06	0.072	0.038	0.052	0.038	0.062	0.056	0.048	0.038	0.036	0.05
CPZ	0.086	0.08	0.074	0.078	0.066	0.046	0.068	0.05	0.05	0.084	0.052	0.052

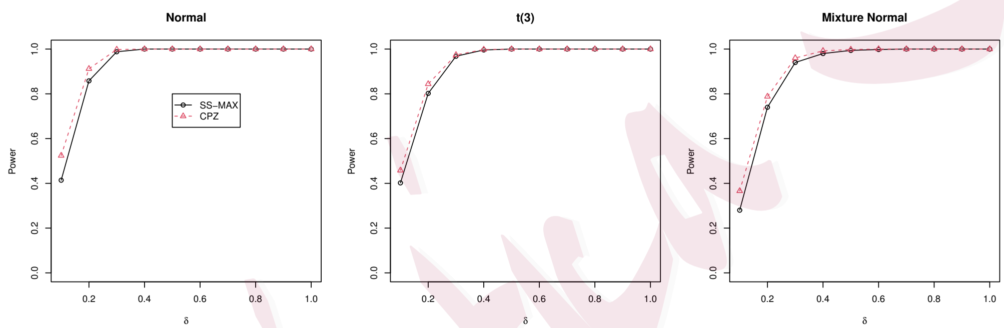

Figure 5: Power of max-type tests with different signal strength for matrix case (i) over $(n, p) = (100, 200)$.

representation and Gaussian approximation of the spatial median estimator, as discussed in Feng et al. (2016). Following this, we introduce a spatial-sign based max-type test procedure for sparse alternatives and establish the limit null distribution and consistency of the proposed max-type test statistic. Next, we reformulate the sum-type test statistic, originally proposed by Feng and Sun (2016), under a general model. This sum-type test

Statistica Sinica: Newly accepted Paper

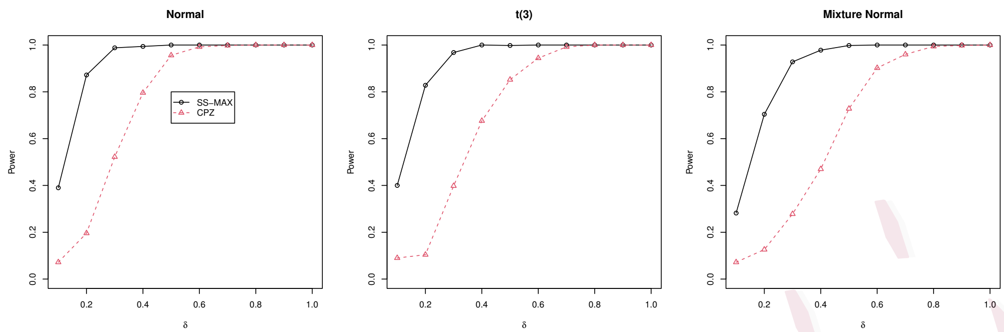

Figure 6: Power of max-type tests with different signal strength for matrix case (ii) over $(n, p) = (100, 200)$.

exhibits superior performance under dense alternatives. Finally, we demonstrate the asymptotic independence between the aforementioned max-type test statistic and the sum-type test statistic, given some mild conditions. We then propose a Cauchy combination test procedure, which performs exceptionally well under both sparse and dense alternatives. Both simulation studies and real data applications underscore the superiority of the proposed maxsum-type test procedure.

We propose several directions for future research. Firstly, the sum-type test statistic in Feng and Sun (2016) only takes into account the direction of the sample, neglecting the information of the sample's radius. Feng et al. (2021) introduced a more powerful inverse norm sign test. It would be intriguing to derive a max-type test statistic that also considers the radius of the sample. Furthermore, it remains an open question whether this new max-type test statistic maintains asymptotic independence with the sum-

Statistica Sinica: Newly accepted Paper

type test statistic proposed by Feng et al. (2021).

Secondly, the newly proposed methods can be extended to address other high-dimensional testing problems. These include the high-dimensional two-sample location problem (Chen and Qin, 2010; Feng et al., 2016), high-dimensional covariance matrix tests (Chen et al., 2010; Li and Chen, 2012; Cutting et al., 2017; Cheng et al., 2019), testing the martingale difference hypothesis in high dimension (Chang et al., 2023) and high-dimensional white noise test (Paindaveine and Verdebout, 2016; Chang et al., 2017; Feng et al., 2022; Zhao et al., 2023). Additionally, the alpha test in the high-dimensional linear factor pricing model is a significant problem that has been explored in practical applications.

Thirdly, our paper's theoretical results are predicated on the assumption of identical and independent distribution. However, there may occasionally be auto-correlations among the sample sizes. Recent literature, such as Zhang and Cheng (2018) and Chang et al. (2024), has considered the Gaussian approximation of the sample mean under a dependent assumption. Therefore, it would be intriguing to establish the Bahadur representation and Gaussian approximation of the spatial median in the context of dependent observations. Building on these findings, we can also suggest the implementation of max-type and maxsum-type testing methods

Statistica Sinica: Newly accepted Paper

for addressing high-dimensional location problem in the context of dependent observations (Ayyala et al., 2017; Ma et al., 2024).

Acknowledgement

The authors thank the editor, the associate editor and two anonymous referees for helpful comments and discussions. This paper is partially supported by Shenzhen Wukong Investment Company, Tianjin Science Fund for Outstanding Young Scholar (23JCJQJC00150), the Fundamental Research Funds for the Central Universities under Grant No. ZB22000105 and 63233075, the China National Key R&D Program (Grant Nos. 2019YFC1908502, 2022YFA1003703, 2022YFA1003802, 2022YFA1003803) and the National Natural Science Foundation of China Grants (Nos. 12271271, 11925106, 12231011, 11931001 and 11971247). Long Feng and Ping Zhao are the co-corresponding authors.

Supplementary Materials

The supplement provide the proofs of the lemmas, theorems given in Section 2-3 and two real data applications. The proofs for lemmas and theorems are given in Appendix S1 and S2 respectively and the real data applications are in Appendix S4, in the online Supplementary Material.

Statistica Sinica: Newly accepted Paper

References

Ayyala, D. N., J. Park, and A. Roy (2017). Mean vector testing for high-dimensional dependent observations. *Journal of Multivariate Analysis 153*, 136–155.

Bai, Z. and H. Saranadasa (1996). Effect of high dimension: by an example of a two sample problem. *Statistica Sinica*, 311–329.

Bickel, P. J. and E. Levina (2008). Covariance regularization by thresholding. *The Annals of Statistics 36*(6), 2577–2604.

Cai, T., W. Liu, and Y. Xia (2013, 08). Two-sample test of high dimensional means under dependence. *Journal of the Royal Statistical Society Series B: Statistical Methodology 76*(2), 349–372.

Chang, J., X. Chen, and M. Wu (2024). Central limit theorems for high dimensional dependent data. *Bernoulli 30*(1), 712–742.

Chang, J., Q. Jiang, and X. Shao (2023). Testing the martingale difference hypothesis in high dimension. *Journal of Econometrics 235*(2), 972–1000.

Chang, J., Q. Yao, and W. Zhou (2017). Testing for high-dimensional white noise using maximum cross-correlations. *Biometrika 104*(1), 111–127.

Chen, S. X. and Y.-L. Qin (2010). A two-sample test for high-dimensional data with applications to gene-set testing. *The Annals of Statistics 38*(2), 808–835.

Chen, S. X., L.-X. Zhang, and P.-S. Zhong (2010). Tests for high-dimensional covariance ma-

Statistica Sinica

Statistica Sinica: Newly accepted Paper

trices. *Journal of the American Statistical Association 105*(490), 810–819.

Cheng, G., B. Liu, L. Peng, B. Zhang, and S. Zheng (2019). Testing the equality of two high-dimensional spatial sign covariance matrices. *Scandinavian Journal of Statistics 46*(1), 257–271.

Cheng, G., L. Peng, and C. Zou (2023). Statistical inference for ultrahigh dimensional location parameter based on spatial median. *arXiv preprint arXiv:2301.03126*.

Cutting, C., D. Paindaveine, and T. Verdebout (2017). Testing uniformity on high-dimensional spheres against monotone rotationally symmetric alternatives. *The Annals of Statistics 45*(3), 1024–1058.

Fang, K. W. (2018). *Symmetric multivariate and related distributions*. CRC Press.

Feng, L., T. Jiang, X. Li, and B. Liu (2022). Asymptotic independence of the sum and maximum of dependent random variables with applications to high-dimensional tests. *arXiv preprint arXiv:2205.01638*.

Feng, L., T. Jiang, B. Liu, and W. Xiong (2022). Max-sum tests for cross-sectional independence of high-dimensional panel data. *Annals of Statistics 50*(2), 1124–1143.

Feng, L., B. Liu, and Y. Ma (2021). An inverse norm sign test of location parameter for high-dimensional data. *Journal of Business & Economic Statistics 39*(3), 807–815.

Feng, L., B. Liu, and Y. Ma (2022). Testing for high-dimensional white noise. *arXiv preprint arXiv:2211.02964*.

Statistica Sinica: Newly accepted Paper

Feng, L. and F. Sun (2016). Spatial-sign based high-dimensional location test. *Electronic Journal of Statistics 10*, 2420–2434.

Feng, L., X. Zhang, and B. Liu (2020). A high-dimensional spatial rank test for two-sample location problems. *Computational Statistics & Data Analysis 144*, 106889.

Feng, L., C. Zou, and Z. Wang (2016). Multivariate-sign-based high-dimensional tests for the two-sample location problem. *Journal of the American Statistical Association 111*(514), 721–735.

Feng, L., C. Zou, Z. Wang, and L. Zhu (2015). Two-sample behrens-fisher problem for high-dimensional data. *Statistica Sinica 25*, 1297–1312.

Hallin, M. and D. Paindaveine (2006). Semiparametrically efficient rank-based inference for shape. i. optimal rank-based tests for sphericity. *Annals of Statistics 34*(6), 2707–2756.

He, Y., G. Xu, C. Wu, and W. Pan (2021). Asymptotically independent u-statistics in high-dimensional testing. *Annals of Statistics 49*(1), 151–181.

Hettmansperger, T. P. and R. H. Randles (2002). A practical affine equivariant multivariate median. *Biometrika 89*(4), 851–860.

Huang, X., B. Liu, Q. Zhou, and L. Feng (2023). A high-dimensional inverse norm sign test for two-sample location problems. *Canadian Journal of Statistics 51*(4), 1004–1033.

Ilmonen, P. and D. Paindaveine (2011). Semiparametrically efficient inference based on signed ranks in symmetric independent component models. *Annals of Statistics 39*(5), 2448–2476.

Statistica Sinica: Newly accepted Paper

Li, J. and S. X. Chen (2012). Two sample tests for high-dimensional covariance matrices. *Annals of Statistics 40*(2), 908–940.

Li, W. and Y. Xu (2022). Asymptotic properties of high-dimensional spatial median in elliptical distributions with application. *Journal of Multivariate Analysis 190*, 104975.

Liu, Y. and J. Xie (2020). Cauchy combination test: A powerful test with analytic p-value calculation under arbitrary dependency structures. *Journal of the American Statistical Association 115*(529), 393–402.

Long, M., Z. Li, W. Zhang, and Q. Li (2023). The cauchy combination test under arbitrary dependence structures. *The American Statistician 77*(2), 134–142.

Ma, H., L. Feng, and Z. Wang (2023). Adaptive testing for alphas in conditional factor models with high dimensional assets. *arXiv preprint arXiv:2307.09397*.

Ma, H., L. Feng, Z. Wang, and B. Jigang (2024). Testing alpha in high dimensional linear factor pricing models with dependent observations. *arXiv preprint arXiv:2401.14052*.

Nordhausen, K., H. Oja, and D. Paindaveine (2009). Signed-rank tests for location in the symmetric independent component model. *Journal of Multivariate Analysis 100*(5), 821–834.

Oja, H. (2010). *Multivariate nonparametric methods with R: an approach based on spatial signs and ranks*. Springer Science & Business Media.

Paindaveine, D. and T. Verdebout (2016). On high-dimensional sign tests. *Bernoulli 22*(3),

Statistica Sinica: Newly accepted Paper

1745–1769.

Park, J. and D. N. Ayyala (2013). A test for the mean vector in large dimension and small samples. *Journal of Statistical Planning and Inference 143*(5), 929–943.

Srivastava, M. S. (2009). A test for the mean vector with fewer observations than the dimension under non-normality. *Journal of Multivariate Analysis 100*(3), 518–532.

Wang, G. and L. Feng (2023). Computationally efficient and data-adaptive changepoint inference in high dimension. *Journal of the Royal Statistical Society Series B: Statistical Methodology 85*(3), 936–958.

Wang, L., B. Peng, and R. Li (2015). A high-dimensional nonparametric multivariate test for mean vector. *Journal of the American Statistical Association 110*(512), 1658–1669.

Wu, C., G. Xu, and W. Pan (2019). An adaptive test on high-dimensional parameters in generalized linear models. *Statistica Sinica 29*(4), 2163–2186.

Wu, C., G. Xu, X. Shen, and W. Pan (2020). A regularization-based adaptive test for high-dimensional generalized linear models. *The Journal of Machine Learning Research 21*(1), 5005–5071.

Xu, G., L. Lin, P. Wei, and W. Pan (2016). An adaptive two-sample test for high-dimensional means. *Biometrika 103*(3), 609–624.

Yao, J., S. Zheng, and Z. Bai (2015). Sample covariance matrices and high-dimensional data analysis. *Cambridge UP, New York*.

Statistica Sinica: Newly accepted Paper

Yu, X., D. Li, and L. Xue (2022). Fisher's combined probability test for high-dimensional covariance matrices. *Journal of the American Statistical Association Online publised*(1-36).

Yu, X., J. Yao, and L. Xue (2023). Power enhancement for testing multi-factor asset pricing models via fisher's method. *Journal of Econometrics In press*.

Zhang, X. and G. Cheng (2018). Gaussian approximation for high dimensional vector under physical dependence. *Bernoulli 24*(4A), 2640–2675.

Zhao, P., D. Chen, and Z. Wang (2023). Spatial-sign based high dimensional white noises test. *arXiv preprint arXiv:2303.10641*.

Zou, C., L. Peng, L. Feng, and Z. Wang (2014). Multivariate sign-based high-dimensional tests for sphericity. *Biometrika 101*(1), 229–236.

School of Statistics and Data Science, KLMDASR, LEBPS, and LPMC, Nankai University

E-mail: jxliu@mail.nankai.edu.cn

School of Statistics and Data Science, KLMDASR, LEBPS, and LPMC, Nankai University

E-mail: flnankai@nankai.edu.cn

School of Statistics and Data Science, KLMDASR, LEBPS, and LPMC, Nankai University

E-mail: zpnanakai@hotmail.com

School of Statistics and Data Science, KLMDASR, LEBPS, and LPMC, Nankai University

E-mail: zjwangnk@126.com