Hindawi
Journal of Mathematics
Volume 2023, Article ID 9350523, 11 pages
https://doi.org/10.1155/2023/9350523

Research Article

One-Way High-Dimensional ANOVA

Tansheng Chen [1] and Lukun Zheng [2]

[1] *Department of Communication and Engineering, School of Information Science and Technology, Southwest Jiaotong University, Chengdu 611756, China*
[2] *Department of Mathematics, Western Kentucky University, Bowling Green, KY 42101, USA*

Correspondence should be addressed to Tansheng Chen; ctansheng@my.swjtu.edu.cn

Received 13 December 2022; Revised 14 February 2023; Accepted 15 February 2023; Published 28 March 2023

Academic Editor: Jiancheng Jiang

ANOVA is one of the most important tools in comparing the treatment means among different groups in repeated measurements. The classical F test is routinely used to test if the treatment means are the same across different groups. However, it is inefficient when the number of groups or dimension gets large. We propose a smoothing truncation test to deal with this problem. It is shown theoretically and empirically that the proposed test works regardless of the dimension. The limiting null and alternative distributions of our test statistic are established for fixed and diverging number of treatments. Simulations demonstrate superior performance of the proposed test over the F test in different settings.

1. Introduction

In bioscience, given a treatments, a central interesting problem is to compare the treatment mean differences. To deal with this problem, one usually employs the traditional univariate ANOVA to analyse independent random samples $Y_{i1}, \ldots, Y_{in}$ from the i th treatment, $i = 1, 2, \ldots, a$. A critical assumption is that the sample $\{Y_{ij}\}_{i=1}^{n}$ is from an $N(\mu_i, \sigma^2)$ population. Then, Y_{ij}, $i = 1, \ldots, a$, $j = 1, \ldots, n$ is a sequence of independent random variables satisfying

$$Y_{ij} = \mu_i + \varepsilon_{ij}, \tag{1}$$

where ε_{ij} follows the $N(0, \sigma^2)$ distribution. Let $\tau_i = \mu_i - \mu$, where $\mu = a^{-1}\sum_{i=1}^{a}\mu_i$. Then, $\sum_{i=1}^{a}\tau_i = 0$, and τ_i is referred to as the effect of the i th treatment. Furthermore, model (1) can be rewritten as

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}. \tag{2}$$

Let $\overline{Y}_{i\cdot} = n^{-1}\sum_{j=1}^{n}Y_{ij}$ and $\overline{Y}_{\cdot\cdot} = N^{-1}\sum_{i=1}^{a}\sum_{j=1}^{n}Y_{ij}$, where $N =$ an. Therefore, for model (2), one of the important problems is to test if the treatment means are different, which is amount to testing

$$H_0\colon \boldsymbol{\tau} = (\tau_1, \ldots, \tau_a)' = \overrightarrow{0} \text{ versus } H_a\colon \boldsymbol{\tau} \neq \overrightarrow{0}. \tag{3}$$

The classical F test is routinely employed in practice and takes the form

$$F_a = \frac{\mathrm{MS_{tr}}}{\mathrm{MSE}}, \tag{4}$$

where $MS_{tr} = n/a - 1\sum_{i=1}^{a}(\overline{Y}_{i\cdot} - \overline{Y}_{\cdot\cdot})^2$ is the mean sum of squares due to treatments, and $\mathrm{MSE} = 1/N - a\sum_{i=1}^{a}\sum_{j=1}^{n}(Y_{ij} - \overline{Y}_{i\cdot})^2$ is the mean sum of squares due to errors. In the current research, we relax the normality assumption by assuming that ε_{ij} are independent and identically distributed noises of mean zero and variance σ^2.

The properties of the F test have been well studied in the conventional low-dimensional setting. It enjoys desirable properties when the dimension a is fixed, see, for example, Casella and Berger [1]. The F test is also robust to the normality assumption, if a is fixed and $n \longrightarrow \infty$ [2]. Akritas and Papadatos [3] (Theorem 2.1) proved that if $n \longrightarrow \infty$ as $a \longrightarrow \infty$, then under H_0, $\sqrt{a}(F_a - 1) \xrightarrow{d} N(0, 2)$. This shows that the F test is asymptotically accurate as $a \longrightarrow \infty$ when the normality assumption does not hold.

For a range of applications including anomaly detection, medical imaging, and genomics, the means of two levels are typically identical or are quite similar in the sense that they possibly differ in only a small number of levels or groups. In other words, under the alternative H_1, the treatment effects are sparse. This is equivalent to the sparse alternative, see, e.g., Cao and Worsley [4] and Taylor and Worsley [5]. In these sparse settings, the F test is not powerful, and the power of the F test in general is fast decreasing as the number of levels increases. This motivates us to propose a smoothing truncation test which smoothly downweights the contributions of those data with treatment means close to zero. This is desired since those data with small treatment means are just noise. Our test is different from the adaptive Neyman test for goodness-of-fit in Fan [6] which works for two-sample means without repeated measurements. Our test is also different from the multisample ANOVA tests for high-dimensional means by Chen et al. [7] where the number of samples is fixed, since testing problem (3) for model (1) can be regarded as a test for n-sample a-dimensional means, with diverging n and a.

We establish asymptotic distributions of the proposed test under the null and the alternative. Simulations demonstrate superior performance of the proposed test over the classical F test. Our test performs well in general and is particularly much more powerful against sparse alternatives than the F test in high-dimensional settings.

Our approach can be extended to heteroscedastic and unbalance cases considered in Akritas and Papadatos [3]. It is also obviously applicable to general multifactor models in Wang (2004) [8]. Since the one-way F statistic coincides with the lack-of-fit statistic for testing if a regression function is constant against a general alternative at the repeated measurement settings, thus our methodology can be applied to this problem, with the current repeated measurements replaced by the residuals under the general alternative. Interested readers are recommended to refer to Härdle and Mammen [9] and Hart [10], among others.

The remainder of the paper is organized as follows. In Section 2, we introduce the smoothing truncation test. In Section 3, we establish the asymptotic distributions of the test under the null and the alternatives. In Section 4, we conduct simulations to compare finite sample performances of different tests.

2. Smoothing Truncation Test

Let $\tilde{z}_{in} = \sqrt{n(\overline{Y}_{i\cdot} - \overline{Y}_{\cdot\cdot})}/\sqrt{\text{MSE}}$, which normalizes the estimator, $\overline{Y}_{i\cdot} - \overline{Y}_{\cdot\cdot}$, of the i th treatment effect. Then, the F statistic in (3) can be rewritten as

$$F_n = \frac{1}{a-1}\sum_{i=1}^{a}\tilde{z}_{\text{in}}^2, \quad (5)$$

where each treatment receives the same weight in the average. For fixed a, F_n is asymptotically χ^2-distributed with degrees of freedom a, and only those treatments of nonzero means contribute to the power of the test. Hence, the F test can be improved if different weights are used in its definition. To this end, we downweigh the contributions of those data with small treatment effect by smoothly truncating the contribution of each $\tilde{z}_{\text{in}}$:

$$T_n = \sum_{i=1}^{a} w_i\tilde{z}_{\text{in}}^2, \quad (6)$$

where $w_i = K(0) - K(\tilde{z}_{in}^2/h_n)$ with K being a kernel function which can be taken as the standard normal density function. The smoothing parameter h_n controls the size of weight. Intuitively, $\tilde{z}_{\text{in}}^2$ is large if the i th treatment effect is nonzero; otherwise, it is small noise. Therefore, weight w_i gets smaller as the treatment effect gets closer to zero, and the truncation test should be more powerful than the F test. Like the F test, large values of T_n suggest rejection of the null, so it is a right-tailed test. Other ways can also be developed to downweigh the contributions of small $\tilde{z}_{\text{in}}^2$ and will be explored in the future.

3. Asymptotic Distributions

To study the distributions of T_n under the null and alternative hypotheses, we first introduce some notations. For a vector $\tau = (\tau_1, \ldots, \tau_a)' \in R^a$, define the L_2-norm by $\|\tau\|_2 = (\sum_{i=1}^{a}|\tau_i|^2)^{1/2}$. Let $\mathbf{I_a}$ be the $a \times a$ identity matrix, and let $\hat{z}_{\text{in}} = \sqrt{n(\overline{Y}_{i\cdot} - \overline{Y}_{\cdot\cdot})/\sigma}$, then $\tilde{z}_{in} = \eta\hat{z}_{in}$, where $\eta = \sigma/\sqrt{\text{MSE}}$.

3.1. Smoothing Truncation Test with Fixed Number of Treatments. The following condition on the kernel function is needed for establishing the limiting distributions of T_n.

Condition 1. Assume $K(\cdot)$ is uniformly continuous and satisfies $\sup_{x\in R}|xK(x)| = M < \infty$ for some $M > 0$.

Condition 1 is satisfied for a wide range of kernel functions, for example, the standard normal density function. The boundness of the first moment of the kernel was used in Jiang [11], and the uniformly continuous assumption is satisfied by common choice of kernel functions, such as the standard normal density kernel and the Epanechnikov kernel.

Theorem 1. *Let $T_n^* = \sum_{i=1}^{a}\tilde{z}_{in}^2(K(0) - K(\tilde{z}_{in}^2/h_n))$. Assume that $h_n \longrightarrow 0$ as $n \longrightarrow \infty$. Then, $T_n - T_n^* \longrightarrow 0$ inprobability, as $n \longrightarrow \infty$.*

Proof. Let $g(x) = x^2(K(0) - K(x^2/h_n))$. Then, $g(x)$ is a uniformly continuous function. When a is finite, it is obvious that $\eta = 1 + o_p(1)$. Note that for each given i, $\hat{z}_{in} = O_p(1)$. Then, $\sup_i|\tilde{z}_{\text{in}} - \hat{z}_{\text{in}}| = o_p(1)$. It follows from the continuous mapping theorem that $T_n - T_n^* \longrightarrow 0$ in probability.

For studying the power, we consider a sequence of local Pitman alternatives $H_{\text{an}}^{(1)}$: $\tau = n^{-1/2}\tau_n$, where τ_n is a sequence of vectors in R^a such that $\tau_n \longrightarrow \tau^*$ with $0 < \|\tau^*\|_2 < \infty$. ☐

Theorem 2. *Assume that Condition 1 holds and $\sigma^2 = var(Y_{ij}) < \infty$. Then, as $n \longrightarrow \infty$,*

(i) *Under* H_0, $T_n^*/K(0) \xrightarrow{d} \chi^2(a-1)$, *where and thereafter* "$\xrightarrow{d}$" *represents converging in distribution*
(ii) *Under* $H_{an}^{(1)}$, $T_n^*/K(0) \xrightarrow{d} \chi^2(a-1, \|\tau^*\|_2^2/\sigma^2)$

Proof. We observe that

$$\begin{aligned}\frac{T_n^*}{K(0)} &= \frac{1}{K(0)}\left[\sum_{i=1}^{a} \widehat{z}_{\text{in}}^2\left(K(0) - K\left(\frac{\widehat{z}_{\text{in}}^2}{h_n}\right)\right)\right] \\ &= \sum_{i=1}^{a}\left(\widehat{z}_{\text{in}}^2\right)^2 - \frac{1}{K(0)}\sum_{i=1}^{a}\widehat{z}_{\text{in}}^2 K\left(\frac{\widehat{z}_{\text{in}}^2}{h_n}\right).\end{aligned} \tag{7}$$

Note that $\widehat{z}_{\text{in}} = \sqrt{n(\overline{Y}_{i\cdot} - \overline{Y}_{\cdot\cdot})}/\sigma$. Let $\varepsilon_{ij} = (Y_{ij} - \mu)/\sigma$ and $v_{n,i} = n^{-1/2}\sum_{j=1}^{n}\varepsilon_{\text{ij}}$. Then, v_{ni}, s are independent random variables with mean $\sqrt{n}\tau_i/\sigma$ and variance one, and

$$\widehat{z}_{in} = n^{-1/2}\sum_{j=1}^{n}\varepsilon_{ij} - \frac{\sqrt{n}(\overline{Y}_{\cdot\cdot} - \mu)}{\sigma} + \frac{\sqrt{n}\tau_i}{\sigma} = v_{n,i} - \overline{v}_n, \tag{8}$$

where $\overline{v}_n = a^{-1}\sum_{i=1}^{a} v_{n,i} = 1/a\mathbf{1}_a^{\top}\mathbf{v}_n$, with $\mathbf{v}_n = (v_{n,1}, \ldots, v_{n,a})^{\top}$ and $\mathbf{1}_a$ being an a-dimensional column vector of all components equal to one. For fixed a, by the Cramér–Wold device, $\mathbf{v}_n - \sqrt{n}\tau/\sigma$ is asymptotically normal distribution with mean zero and variance-covariance matrix $\mathbf{I}_a$. It is straightforward to verify that

$$\sum_{i=1}^{a}\widehat{z}_{\text{in}}^2 = \sum_{i=1}^{a}\left(v_{n,i} - \overline{v}_n\right)^2 = \mathbf{v}_n^{\top}\sum \mathbf{v}_n. \tag{9}$$

Then, there exists an orthogonal matrix $\mathbf{Q}$ and a diagonal matrix $\Lambda = \text{diag}\{\mathbf{1}_{a-1}, 0\}$ such that $\Sigma = \mathbf{Q}^{\top}\mathbf{\Lambda}\mathbf{Q}$. Let $\mathbf{Q}^{\top} = (\mathbf{q}_1, \ldots, \mathbf{q}_a)$, where $\mathbf{q}_i$ is the i-th row of $\mathbf{Q}$, and let $\mathbf{v}_n^* = \mathbf{Q}\mathbf{v}_n$ with $v_{n,i}^* = \mathbf{q}_i^{\top}\mathbf{v}_n$ being the i th entry of $\mathbf{v}_n^*$. Then, $\mathbf{v}_n^* - \mathbf{Q}\sqrt{n}\tau/\sigma = \mathbf{Q}(\mathbf{v}_n - \sqrt{n}\,\tau/\sigma)$ is asymptotically normal with mean zero and variance-covariance matrix $\mathbf{I_a}$. Recall from (2) that

$$\begin{aligned}\sum_{i=1}^{a}\widehat{z}_{\text{in}}^2 &= \mathbf{v}_n^{*\top}\mathbf{\Lambda}\mathbf{v}_n^* \\ &= \sum_{i=1}^{a-1} v_{n,i}^{*2}.\end{aligned} \tag{10}$$

It follows that

(i) Under H_0, $\mathbf{v}_n^*$ is asymptotically normal with mean zero and variance-covariance matrix $\mathbf{I_a}$, so that $\sum_{i=1}^{a}(v_{n,i} - \overline{v}_n)^2 \xrightarrow{d} \chi^2(a-1)$. Then $T_n^*/K(0) \xrightarrow{d} \chi^2(a-1)$, since the 2nd term on the righthand of (1) is $o(1)$ for $\widehat{z}_{\text{in}}^2 K(\widehat{z}_{\text{in}}^2/h_n) \le Mh_n \longrightarrow 0$, as $n \longrightarrow \infty$.
(ii) Under $H_{\text{an}}^{(1)}$, $\sqrt{n}\tau/\sigma \longrightarrow \tau^*/\sigma$, and $\mathbf{v}_n^*$ is asymptotically normal with mean τ_i^*/σ and variance-covariance matrix $\mathbf{I_a}$, so that $\sum_{i=1}^{a}(v_{n,i} - \overline{v}_n)^2 \xrightarrow{d} \chi^2(a-1, \|\tau^*\|_2^2/\sigma^2)$. Hence, $T_n^*/K(0) \xrightarrow{d} \chi^2(a-1, \|\tau^*\|_2^2/\sigma^2)$.

Combining Theorems 1 and 2 gives us the following asymptotic distribution of T_n. □

Theorem 3. *Assume that conditions in Theorem 2 hold. Then, as* $n \longrightarrow \infty$,

(i) $T_n/K(0) \xrightarrow{d} \chi^2(a-1)$
(ii) $T_n/K(0) \xrightarrow{d} \chi^2(a-1, \|\tau^*\|^2/\sigma^2)$

The above theorem demonstrates that the smoothing truncation can detect local alternatives close to the null at rate of $\sqrt{n}$, *which is the optimal rate that all regular parametric tests can achieve.*

3.2. Smoothing Truncation Test with Diverging Number of Treatments. Let $\Lambda = n\sigma^{-2}\|\tau\|_2^2$. To obtain the limiting null and alternative distributions, we need additional conditions.

Condition 2. $n \longrightarrow \infty$, *as* $a \longrightarrow \infty$; $h_n\sqrt{a} = o(1)$, *as* $n \longrightarrow \infty$

Condition 3. Suppose the Cramér condition holds for Y_{ij}, i.e.,

$$E\left|Y_{ij}\right|^m \le m!M^{m-2}\frac{\sigma_i^2}{2}, \tag{11}$$

for all i and j, where M is a positive constant, $m \ge 2$, and $\sigma_i^2 = \text{var}(Y_{\text{ij}}) < \infty$.

The first part of Condition 2 means that we consider high-dimensional settings with a diverging number of populations. It is a setting considered in Akritas and Papadatos [3]. Condition 2 restricts the smoothing parameters h_n. This is a wild condition. As $a \longrightarrow \infty$, it only requires $h_n = o(a^{-1/2})$. Condition 3 is trivially fulfilled if Y_{ij}, s are bounded; for Gaussian variables, it obviously holds.

By the definition of F_n, we have

$$F_n = \eta(a-1)^{-1}\sum_{i=1}^{a}\widehat{z}_{\text{in}}^2. \tag{12}$$

The following result shows that the difference between $\widetilde{z}_{\text{in}}$ and $\widehat{z}_{\text{in}}$ is $o_p(1)$ uniformly in $i = 1, \ldots, a$ for extremely large a.

Lemma 1. *Assume that* $a = O(\exp(n^{\delta}))$ *for* $0 \le \delta < 1/2$ *and* $sup_{1\le i\le a}|\tau_i| = O(n)$. *Then,* $sup_{1\le i\le a}|\widetilde{z}_{in} - \widehat{z}_{in}| = o_p(1)$, *as* $n \longrightarrow \infty$.

Proof

(i) We show that $\eta = \sigma/\sqrt{\text{MSE}} = 1 + O_p(\sqrt{u_n})$. It is easy to show the identity:

$$\frac{1}{n-1}\sum_{j=1}^{n}\left(Y_{ij}-\overline{Y}_{i\cdot}\right)^2 = \frac{1}{n-1}\sum_{j=1}^{n}\left(Y_{ij}-\mu_i\right)^2 - \frac{n}{n-1}\left(\overline{Y}_{i\cdot}-\mu_i\right)^2 = L_{n1,i} - L_{n2,i}. \tag{13}$$

Then, by the definition of MSE and the triangle inequality,

$$\left|\mathrm{MSE}-\sigma^2\right| \le a^{-1}\sum_{i=1}^{a}\left|L_{n1,i}-\sigma^2\right| + a^{-1}\sum_{i=1}^{a}\left|L_{n2,i}\right| \equiv Q_{n1} + Q_{n2}. \tag{14}$$

Hence, for any $C > 0$, we have

$$\begin{aligned} P\left(\left|Q_{n1}\right| > u_n C\right) &= P\left(\sum_{i=1}^{a}\left|L_{n1,i}-\sigma^2\right| > u_n a C\right) \\ &\le P\left(\sup_{1\le i\le a}\left|L_{n1,i}-\sigma^2\right| > u_n C\right) \\ &\le \sum_{i=1}^{a} P\left(\left|L_{n1,i}-\sigma^2\right| > u_n C\right). \end{aligned} \tag{15}$$

Let $\varepsilon_{ij}^{*} = (Y_{ij}-\mu_i)/\sigma$. Then, $\left\{\varepsilon_{ij}^{*}\right\}_{j=1}^{n}$ is a sequence of iid random variables with mean zero and variance 1 and satisfying Condition 3. Using the Bernstein inequality, we obtain that

$$\begin{aligned} P\left(\left|\frac{n-1}{n}L_{n1,i}-\sigma^2\right| > u_n C\right) &= P\left(\left|\sum_{j=1}^{n}\left(\varepsilon_{ij}^{*2}-1\right)\right| > \mathrm{nu}_n\frac{C}{\sigma^2}\right) \\ &\le 2\ \exp\left\{-\frac{n^2u_n^2C^2}{\left[2\sigma^4\left(n\mathrm{var}\left(\varepsilon_{i1}^{*2}\right)+nu_n\mathrm{CM}/\sigma^2\right)\right]}\right\}. \end{aligned} \tag{16}$$

Combining (5) and (6) yields that

$$P\left(\left|Q_{n1}\right| > u_n C\right) \le 2a\ \exp\left\{-\frac{n^2u_n^2C^2}{\left[2\left(n\mathrm{var}\left(\varepsilon_{i1}^{*2}\right)+\mathrm{nu}_n\mathrm{CM}\sigma^2\right)\right]}\right\} \longrightarrow 0, \tag{17}$$

if $a = o(\exp\{\min(\mathrm{nu}_n^2, \mathrm{nu}_n)\})$. Thus, $Q_{n1} = O_p(u_n)$. Let $Q_{n2}^{*} = a^{-1}\sum_{i=1}^{a}(\overline{Y}_i-\mu_i)^2/\sigma^2$. Then, $Q_{n2} = n\sigma^2/n-1Q_{n2}^{*}$ and $Q_{n2}^{*} = a^{-1}\sum_{i=1}^{a}(n^{-1}\sum_{j=1}^{n}\varepsilon_{ij})^{*2}$. It follows that, for a positive sequence u_n and positive constant C,

$$\begin{aligned} P\left(\left|Q_{n2}^{*})\right| > u_n C\right) &\le P\left\{\sum_{i=1}^{a}\left(n^{-1}\sum_{j=1}^{n}\varepsilon_{ij}^{*}\right)^2 > \mathrm{au}_n C\right\} \\ &\le \sum_{i=1}^{a}P\left\{\left(n^{-1}\sum_{j=1}^{n}\varepsilon_{ij}^{*}\right)^2 > u_n C\right\} \\ &\le \sum_{i=1}^{a}P\left\{\left|\sum_{j=1}^{n}\varepsilon_{ij}^{*}\right| > n\sqrt{u_n C}\right\}. \end{aligned} \tag{18}$$

By the Bernstein inequality, we have

$$P(|Q_{n2}^{*})| > u_n C) \le 2a\ \exp\left\{-\frac{n^2u_nC}{2\left[n\mathrm{var}\left(\varepsilon_{i1}^{*}\right)+n\sqrt{u_n\mathrm{CM}}\right]}\right\} \longrightarrow 0, \tag{19}$$

if $a = o(\exp\{\min(\mathrm{nu}_n, n\sqrt{u_n})\})$. Therefore, $Q_{n2}^* = O_p(u_n)$. Then, by (4), $\mathrm{MSE} = \sigma^2 + O_p(u_n)$, and thus, $\eta - 1 = O_p(\sqrt{u_n})$.

(ii) Since $\tilde{z}_{\mathrm{in}} = \eta \hat{z}_{\mathrm{in}}$, we have

$$\tilde{z}_{\mathrm{in}} - \hat{z}_{\mathrm{in}} = (\eta - 1)\hat{z}_{\mathrm{in}} = O_p(\sqrt{u_n})\hat{z}_{\mathrm{in}}. \quad (20)$$

Rewrite $\hat{z}_{in} = n^{-1/2}\sum_{j=1}^n \varepsilon_{ij}^* - n^{-1/2}(\overline{Y}_{..} - \mu_i)/\sigma \equiv \xi_{n1,i} - \xi_{n2,i}$, where $\varepsilon_{ij}^* = (Y_{ij} - \mu_i)/\sigma$. Note that $\{\varepsilon_{ij}^*\}_{j=1}^n$ is a sequence of iid random variables with mean zero and variance $\sigma_i^{*2} = 1$ and satisfying Condition 3. It follows from the Bernstein exponential inequality that

$$\begin{aligned} P\left(\sup_{1\le i\le a}|\xi_{n1,i}| > u_n^* C\right) &\le \sum_{i=1}^a P\left(\left|\sum_{j=1}^n \varepsilon_{ij}^*\right| > \sqrt{\mathrm{nu}_n^*}\,C\right) \\ &\le 2a\exp\left\{-\frac{nu_n^{*2}C^2}{\left[2\left(n\sigma_i^{*2} + \sqrt{\mathrm{nu}_n^*}\,\mathrm{CM}\right)\right]}\right\} \longrightarrow 0, \end{aligned} \quad (21)$$

if $a = o(\exp\{\min(u_n^{*2}, \sqrt{n}u_n^*)\})$. Hence, $\sup_{1\le i\le a}|\xi_{n1,i}| = O_p(u_n^*)$. Note that

$$\begin{aligned} \xi_{n2,i} &= n^{-1/2}a^{-1}\sum_{i=1}^a n^{-1}\sum_{j=1}^n \frac{(Y_{ij} - \mu_i)}{\sigma} + n^{-1/2}\frac{(\mu - \mu_i)}{\sigma} \\ &= n^{-1/2}a^{-1}\sum_{i=1}^a n^{-1}\sum_{j=1}^n \varepsilon_{ij}^* - \frac{n^{-1/2}\tau_i}{\sigma} \\ &\equiv \nu_n - \frac{n^{-1/2}\tau_i}{\sigma}, \end{aligned} \quad (22)$$

where $\nu_n = O_p(u_n^*)$ and $n^{-1/2}\sup_i|\tau_i|/\sigma = O(u_n^*)$ if $\sup_{1\le i\le a}|\tau_i| = O(\sqrt{\mathrm{nu}_n^*})$. In fact,

$$\begin{aligned} P\left(|\nu_n| > u_n^* C\right) &\le P\left(\sum_{i=1}^a\sum_{j=1}^n |\varepsilon_{ij}^*| > n^{3/2}\mathrm{au}_n^* C\right) \\ &\le P\left(\text{there are at least one } i \text{ such that} \sum_{j=1}^n |\varepsilon_{ij}^*| > n^{3/2}u_n C\right) \\ &\le \sum_{i=1}^a P\left(\sum_{j=1}^n |\varepsilon_{ij}^*| > n^{3/2}u_n^* C\right). \end{aligned} \quad (23)$$

Applying the Bernstein exponential inequality again, we get that

$$P\left(|\nu_n| > u_n^* C\right) \le 2a\exp\left\{-\frac{n^3 u_n^{*2}C^2}{\left[2\left(n\sigma_i^{*2} + n^{3/2}u_n^*\mathrm{CM}\right)\right]}\right\} \longrightarrow 0, \quad (24)$$

if $a = o(\exp\{\min(n^2 u_n^{*2}, n^{3/2}u_n^*)\})$. Then, $\sup_{1\le i\le a}|\xi_{n2,i}| = O_p(u_n^*)$, and thus, $\sup_{1\le i\le a}|\hat{z}_{in}| = O_p(u_n^*)$, which combined with (12) leads to $\sup_{1\le i\le a}|\tilde{z}_{\mathrm{in}} - \hat{z}_{\mathrm{in}}| = O_p(\sqrt{u_n}u_n^*) = o_p(1)$, if we take $u_n = n^{-1/4}$ and $u_n^* = n^{1/2}$, under the condition of $a = o(\exp(n^{1/2}))$. □

Lemma 2. *Assume that Conditions 1–3 hold. Under H_0, we have*

$$\frac{\sum_{i=1}^a \hat{z}_{\mathrm{in}}^2 - (a-1)}{\sqrt{a-1}} = \sqrt{2}Z + o_p(1) \xrightarrow{d} N(0,2), \quad (25)$$

where Z is a standard normal random variable.

Proof. Under H_0, we have $Y_{\mathrm{ij}} = \mu + \varepsilon_{\mathrm{ij}}$. By Theorem 2.1(b) of Akritas and Paradatos [3], we have $\sqrt{a}(F_n - 1) \xrightarrow{d} N(0,2)$, as $a \longrightarrow \infty$, assuming that $\mathrm{EY}_{\mathrm{ij}}^4 < \infty$. That is, $\sqrt{a}(F_n - 1) = \sqrt{2}Z + o_p(1)$, where Z is a standard normal random variable. Then, by (3) and part (i) of the proof for Lemma 1,

$$\begin{aligned}\sum_{i=1}^{a}\hat{z}_{in}^2 &= (a-1)\eta^{-1}F_n\\ &= (a-1)\left(1+o_p(1)\right)\left\{1+\sqrt{\frac{2}{aZ}}+o_p\left(\frac{1}{\sqrt{a}}\right)\right\}\\ &= a-1+\sqrt{2(a-1)Z}+o_p(\sqrt{a}).\end{aligned} \tag{26}$$

Hence, $\sum_{i=1}^{a}\hat{z}_{in}^2-(a-1)/\sqrt{2(a-1)} = Z+o_p(1)\xrightarrow{d} N(0,1)$. □

Theorem 4. *Assume that Conditions 1–3 are satisfied. Then, under the null hypothesis H_0: $\tau_i=0$ for $i=1,\ldots,a$,*

$$\frac{T_n^*-K(0)(a-1)}{K(0)\sqrt{a-1}}\xrightarrow{d} N(0,2), \tag{27}$$

as $n\longrightarrow\infty$.

Proof. Observe that

$$\begin{aligned}\frac{T_n^*}{K(0)} &= \frac{1}{K(0)}\left[\sum_{i=1}^{a}\hat{z}_{\mathrm{in}}^2\left(K(0)-K\left(\frac{\hat{z}_{\mathrm{in}}^2}{h_n}\right)\right)\right]\\ &= \sum_{i=1}^{a}\left(\hat{z}_{\mathrm{in}}^2\right)^2-\frac{1}{K(0)}\sum_{i=1}^{a}\hat{z}_{\mathrm{in}}^2K\left(\frac{\hat{z}_{\mathrm{in}}^2}{h_n}\right).\end{aligned} \tag{28}$$

It follows that

$$\begin{aligned}\frac{T_n^*-K(0)(a-1)}{K(0)\sqrt{2a-2}} &= \frac{\sum_{i=1}^{a}\left(\hat{z}_{\mathrm{in}}^2\right)^2-(a-1)}{\sqrt{2a-2}}-\frac{1}{K(0)\sqrt{2a-2}}\sum_{i=1}^{a}\hat{z}_{\mathrm{in}}^2K\left(\frac{\hat{z}_{\mathrm{in}}^2}{h_n}\right)\\ &\equiv K_{n1}+K_{n2}.\end{aligned} \tag{29}$$

Since

$$\begin{aligned}E\left(\sum_{i=1}^{a}\hat{z}_{\mathrm{in}}^2K\left(\frac{\hat{z}_{\mathrm{in}}^2}{h_n}\right)\right) &= \sum_{i=1}^{a}E\left(\hat{z}_{\mathrm{in}}^2K\left(\frac{\hat{z}_{\mathrm{in}}^2}{h_n}\right)\right)\\ &= \sum_{i=1}^{a}E\left(h_n\frac{\hat{z}_{\mathrm{in}}^2}{h_n}K\left(\frac{\hat{z}_{\mathrm{in}}^2}{h_n}\right)\right),\\ &\le \sum_{i=1}^{a}Mh_n = \mathrm{Mah}_n,\end{aligned} \tag{30}$$

by the Markov inequality,

$$P\left(\frac{1}{K(0)\sqrt{2a-2}}\sum_{i=1}^{a}\hat{z}_{\mathrm{in}}^2K\left(\frac{\hat{z}_{\mathrm{in}}^2}{h_n}\right)>\varepsilon\right)\le\frac{\mathrm{Mah}_n}{K(0)\sqrt{2a-2}}=\frac{1}{\sqrt{2}K(0)}\frac{\mathrm{Mh}_n\sqrt{a}}{\sqrt{1-1/a}}\longrightarrow 0, \tag{31}$$

if $h_n\sqrt{a}=o(1)$. Then, $K_{n2}=o_p(1)$. By Lemma 2, we have $K_{n1}\xrightarrow{d} N(0,1)$, as $a\longrightarrow\infty$. These, combined (10) and Slusky's theorem, yield the result of theorem.

To study the power of the test, we consider the same local alternatives as Akritas and Paradatos [3], which specifies

$$H_{\mathrm{an}}^{(2)}:\ \tau_i=a^{3/4}n^{-1/2}\int_{(i-1)/a}^{i/a}g(t)\mathrm{d}t, \tag{32}$$

where $g(t)$ is a continuous function on [0, 1] such that $\int_0^1 g(t)\mathrm{d}t=0$. With such local alternatives, we have

$$\begin{aligned}\lambda &= \sigma^{-2}a^{3/2}\sum_{i=1}^{a}\left\{\int_{(i-1)/a}^{i/a}g(t)\mathrm{d}t\right\}^2\\ &= \sigma^{-2}a^{-1/2}\sum_{i=1}^{a}g^2(t_i)=\sigma^{-2}a^{1/2}\theta^2(1+o(1)),\end{aligned} \tag{33}$$

where $t_i\in[i/a,(i-1)/a]$ and $\theta^2=\int_0^1 g^2(t)\mathrm{d}t$. Obviously, $H_{\mathrm{an}}^{(2)}$ converges to H_0 at the rate of $a^{-1/4}n^{-1/2}$. □

Lemma 3. *Assume that Conditions 1–3 hold. Under $H_{an}^{(2)}$, we have*

$$\frac{\sum_{i=1}^{a}\widehat{z}_{\text{in}}^{2}-(a-1)}{\sqrt{a-1}}=\sqrt{2}Z+\sigma^{-2}\theta^{2}+o_{p}(1)\overset{d}{\longrightarrow}N\left(\sigma^{-2}\theta^{2},2\right). \tag{34}$$

Proof. Under H_a, we have $Y_{\text{ij}}=\mu+\tau_i+\varepsilon_{\text{ij}}$. Then,

$$\sqrt{n}\left(\overline{Y}_{i\cdot}-\overline{Y}_{\cdot\cdot}\right)=\sqrt{n}\left(\overline{Y}_{i\cdot}-\tau_i\right)-\sqrt{n}\left(\overline{Y}_{\cdot\cdot}-\tau_i\right). \tag{35}$$

Let $Y_{\text{ij}}^{*}=Y_{\text{ij}}-\tau_i$. Then,

$$Y_{\text{ij}}^{*}=\mu+\varepsilon_{\text{ij}}, \tag{36}$$

$\overline{Y}_{i\cdot}^{*}=\overline{Y}_{i\cdot}-\tau_i$, and $\overline{Y}_{\cdot\cdot}^{*}=\overline{Y}_{\cdot\cdot}$, since $\sum_{i=1}^{a}\tau_i=0$. Thus, by (27),

$$\sqrt{n}\left(\overline{Y}_{i\cdot}-\overline{Y}_{\cdot\cdot}\right)=\sqrt{n}\left(\overline{Y}_{i\cdot}^{*}-\overline{Y}_{\cdot\cdot}^{*}\right)+\sqrt{n}\tau_i. \tag{37}$$

Let $\widehat{z}_{\text{in}}^{*}=\sqrt{n}\left(\overline{Y}_{i\cdot}^{*}-\overline{Y}_{\cdot\cdot}^{*}\right)/\sigma$. It follows from (3) that

$$\begin{aligned}F_n&=\frac{\sigma}{\sqrt{\text{MSE}}}\frac{1}{a-1}\sum_{i=1}^{a}\left\{\sqrt{n}\;\frac{\left(\overline{Y}_{i\cdot}^{*}-\overline{Y}_{\cdot\cdot}^{*}\right)}{\sigma}+\sqrt{n}\,\tau_i\sigma\right\}^{2}\frac{\sigma}{\sqrt{\text{MSE}}}\frac{1}{a-1}\sum_{i=1}^{a}\left\{\widehat{z}_{\text{in}}^{*2}+\frac{n\tau_i^{2}}{\sigma^{2}}+2\widehat{z}_{\text{in}}^{*}\sqrt{n}\;\frac{\tau_i}{\sigma}\right\}\\&\equiv L_{n1}+L_{n2}+L_{n3},\end{aligned} \tag{38}$$

where

$$\begin{aligned}L_{n1}&=\frac{\sigma}{\sqrt{\text{MSE}}}\frac{1}{a-1}\sum_{i=1}^{a}\widehat{z}_{\text{in}}^{*2},\\L_{n2}&=\frac{\sigma}{\sqrt{\text{MSE}}}\frac{1}{a-1}\sum_{i=1}^{a}n\frac{\tau_i^{2}}{\sigma^{2}},\\L_{n3}&=\frac{\sigma}{\sqrt{\text{MSE}}}\frac{2}{a-1}\sum_{i=1}^{a}\widehat{z}_{\text{in}}^{*}\sqrt{n}\;\frac{\tau_i}{\sigma}.\end{aligned} \tag{39}$$

Since $L_{n1}=(\sigma/\sqrt{\text{MSE}})(1/a-1)\sum_{i=1}^{a}\widehat{z}_{\text{ij}}^{*2}$ is the F statistic, denoted by F_n^{*} with a little bit of abuse of notation, for model (12) with all $\tau_i=0$, $F_n^{*}=(\sigma/\sqrt{\text{MSE}})(a-1)^{-1}\sum_{i=1}^{a}\widehat{z}_{\text{in}}^{*2}$, using Theorem 2.1 of Akritas and Paradatos [3], we have $\sqrt{a}(F_n^{*}-1)\overset{d}{\longrightarrow}N(0,2)$, or equivalently, $F_n^{*}=1+\sqrt{2/a}Z+o_p(a^{-1/2})$, so that

$$\begin{aligned}\sqrt{a}L_{n1}&=\sqrt{a}F_n^{*}\\&=\sqrt{a}+\sqrt{2}Z+o_p(1).\end{aligned} \tag{40}$$

Note that the 2nd term is

$$\begin{aligned}L_{n2}&=\frac{1}{a-1}\sum_{i=1}^{a}n\frac{\tau_i^{2}}{\sigma^{2}}\left(1+o_p(1)\right)=(a-1)^{-1}n\frac{\|\boldsymbol{\tau}\|_2^{2}}{\sigma^{2}}\left(1+o_p(1)\right)\\&=(a-1)^{-1}\lambda\left(1+o_p(1)\right)\\&=a^{-1/2}\sigma^{-2}\theta^{2}\left(1+o_p(1)\right),\end{aligned} \tag{41}$$

and the 3rd term is

$$\begin{aligned}L_{n3}&=\frac{2}{a-1}\sum_{i=1}^{a}\sqrt{n}\frac{\overline{Y}_{i\cdot}^{*}-\overline{Y}_{\cdot\cdot}^{*}}{\sigma}\sqrt{n}\;\frac{\tau_i}{\sigma}\\&=\frac{2n}{a-1}\sigma^{-1}\sum_{i=1}^{a}\tau_i\frac{\overline{Y}_{i\cdot}^{*}}{\sigma}.\end{aligned} \tag{42}$$

It is straightforward to show that $E(L_{n3})=0$ and $\text{var}(L_{n3})=(4n/(a-1)^{2})(\|\tau\|_2^{2}/\sigma^{2})=(4\lambda/(a-1)^{2})$. Since $\lambda=\sqrt{a}\sigma^{-2}\theta^{2}(1+o(1))$, $L_{n3}=O_p(a^{-1}\sqrt{\lambda})=o_p(a^{-1/2})$. Thus, $\sqrt{a}L_{n2}=\sigma^{-2}\theta^{2}(1+o_p(1))$, and $\sqrt{a}L_{n3}=o_p(1)$, which together with (10)–(14) yields that

$$\sqrt{a}(F_n-1)=\sqrt{2}Z+\sigma^{-2}\theta^{2}+o_p(1). \tag{43}$$

Then, by (3),

$$\sum_{i=1}^{a}\widehat{z}_{\text{in}}^{2}=\frac{\sqrt{\text{MSE}}}{\sigma}\times\left\{a-1+\sqrt{a-1}\left(\sqrt{2}\,Z+\sigma^{-2}\theta^{2}\right)+o_p\left(\sqrt{a}\right)\right\}. \tag{44}$$

Hence,

$$\frac{\sum_{i=1}^{a}\widehat{z}_{\text{in}}^{2}-(a-1)}{\sqrt{a-1}}=\sqrt{2}Z+\sigma^{-2}\theta^{2}+o_p(1)\longrightarrow N\left(\sigma^{-2}\theta^{2},2\right). \tag{45}$$

□

Theorem 5. *Under the alternative hypothesis $H_{an}^{(2)}$: $\tau_n\neq\overrightarrow{0}$,*

$$\frac{T_n^{*}-K(0)(a-1)}{K(0)\sqrt{a-1}}\overset{d}{\longrightarrow}N\left(\sigma^{-2}\theta^{2},2\right), \tag{46}$$

as $n\longrightarrow\infty$, provided that Conditions 1-3 are satisfied.

Proof. Since

$$\begin{aligned}\frac{T_n^*}{K(0)} &= \frac{1}{K(0)}\left[\sum_{i=1}^{a}\widehat{z}_{\text{in}}^2\left(K(0)-K\left(\frac{\widehat{z}_{\text{in}}^2}{h_n}\right)\right)\right]\\ &= \sum_{i=1}^{a}\left(\widehat{z}_{\text{in}}^2\right)^2 - \frac{1}{K(0)}\sum_{i=1}^{a}\widehat{z}_{\text{in}}^2 K\left(\frac{\widehat{z}_{\text{in}}^2}{h_n}\right),\end{aligned} \tag{47}$$

it can be rewritten that

$$\begin{aligned}\frac{T_n^* - K(0)(a-1)}{K(0)\sqrt{a-1}} &= \frac{\sum_{i=1}^{a}\left(\widehat{z}_{\text{in}}^2\right)^2-(a-1)}{\sqrt{a-1}} - \frac{1}{K(0)\sqrt{a-1}}\sum_{i=1}^{a}\widehat{z}_{\text{in}}^2 K\left(\frac{\widehat{z}_{\text{in}}^2}{h_n}\right)\\ &\equiv K_{n3}+K_{n4}.\end{aligned} \tag{48}$$

Note that

$$\begin{aligned}E\left(\sum_{i=1}^{a}\widehat{z}_{\text{in}}^2 K\left(\frac{\widehat{z}_{\text{in}}^2}{h_n}\right)\right) &= \sum_{i=1}^{a}E\left(\widehat{z}_{\text{in}}^2 K\left(\frac{\widehat{z}_{\text{in}}^2}{h_n}\right)\right)\\ &= \sum_{i=1}^{a}E\left(h_n\frac{\widehat{z}_{\text{in}}^2}{h_n}K\left(\frac{\widehat{z}_{\text{in}}^2}{h_n}\right)\right)\\ &\le \sum_{i=1}^{a}\text{Mh}_n = \text{Mah}_n.\end{aligned} \tag{49}$$

It follows that from Markov's inequality, for any $\varepsilon > 0$,

$$\begin{aligned}&P\left(\frac{1}{\sqrt{a-1}}\sum_{i=1}^{a}\widehat{z}_{\text{in}}^2 K\left(\frac{\widehat{z}_{\text{in}}^2}{h_n}\right) > \varepsilon\right)\\ &\le \frac{\text{Mah}_n}{\varepsilon\sqrt{a-1}} = \frac{M\sqrt{\text{ah}_n}}{\varepsilon\sqrt{1-1/a}} \longrightarrow 0,\end{aligned} \tag{50}$$

if $\sqrt{a}h_n \longrightarrow 0$. Thus, $K_{n4} = o_p(1)$. By Lemma 3, we have $K_{n3} \longrightarrow N(\sigma^{-2}\theta^2, 2)$. Then, by Slusky's theorem and (15), the result of theorem holds. □

Corollary 1. *Under the null hypothesis* H_0: $\tau_i = 0$, $i = 1, \ldots, a$,

$$\frac{T_n - K(0)(a-1)}{K(0)\sqrt{a-1}} \xrightarrow{d} N(0,2), \tag{51}$$

as $n \longrightarrow \infty$, *provided that Conditions 1–3 are satisfied.*

From Corollary 1, one gets the rejection region of the T_n test:

$$W = \left\{T_n\colon T_n > K(0)(a-1) + \sqrt{2}Z_{1-\alpha}K(0)\sqrt{a-1}\right\}, \tag{52}$$

where $Z_{1-\alpha}$ is the upper α-percentile of $N(0,1)$.

Corollary 2. *Under the alternative hypothesis* $H_{an}^{(2)}$: $\tau_n \ne \overrightarrow{0}$,

$$\frac{T_n - K(0)(a-1)}{K(0)\sqrt{a-1}} \xrightarrow{d} N\left(\sigma^{-2}\theta^2, 2\right), \tag{53}$$

as $n \longrightarrow \infty$, *provided that Conditions 1–3 are satisfied.*

It is obvious that the power of T_n for testing problem H_0 against $H_{an}^{(2)}$ is

$$\begin{aligned}P\left(T_n \in W | H_{\text{an}}^{(2)}\right) &= P\left\{\frac{\left[T_n - K(0)(a-1)/K(0)\sqrt{a-1} - \sigma^{-2}\theta^2\right]}{\sqrt{2}} > Z_{1-\alpha} - \frac{\sigma^{-2}\theta^2}{\sqrt{2}}\right\}\\ &= 1 - \Phi\left(Z_{1-\alpha} - \frac{\sigma^{-2}\theta^2}{\sqrt{2}}\right).\end{aligned} \tag{54}$$

Table 1: Comparison of the powers between the F test and the proposed test.

a	Critical value of F_n	Power of F_n	Critical value of T_n	Power of T_n
50	1.3565	0.8796	1.1601	0.9980
100	1.2463	0.7191	2.3446	0.9930
150	1.1992	0.6006	3.6925	0.9920
200	1.1715	0.5163	5.4545	0.9810
250	1.1528	0.4545	6.7949	0.9720
300	1.1390	0.4078	8.0685	0.9690
350	1.1284	0.3713	10.3332	0.9700
400	1.1199	0.3421	11.3591	0.9640
450	1.1129	0.3183	13.0851	0.9500
500	1.1070	0.2984	14.7558	0.9490

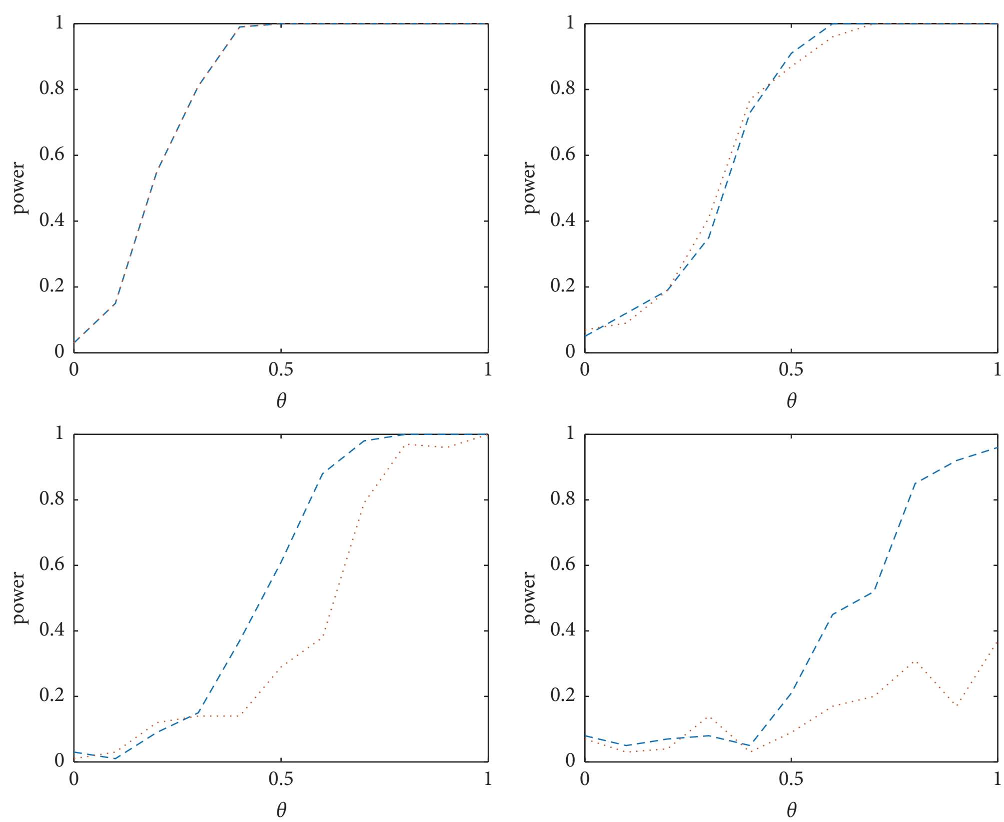

Figure 1: Powers of the tests. Dashed: T_n, dotted: F_n.

This means that the test still has power to distinguish $H_{\text{an}}^{(2)}$ from H_0.

Finally, it is worth pointing out that our theorems above are established under the equal sample size setting, but they can be extended to allow for different sample sizes. Since it involves in more dedicated proofs, we leave this as an open problem that can be explored in the future.

4. Numerical Results

In this section, we consider the numerical performance of the proposed test T_n and compare it with the F test.

4.1. Simulations. Without loss of generality, we take $\mu = 0$ and $\sigma = 1$ in model (1). Our test involves kernel function K and the bandwidth h_n. We take K as the standard normal density function and set $h_n = 100(\text{na})^{-0.2}$ which satisfies Condition 2. For the following two examples, we draw samples from the normal distribution for model (1). Specifically, for each level i, we draw a sample of size n from $N(\tau_i, 1)$ with τ_i being different in the following two examples. For each setting, we conduct 1000 simulations to calculate the critical values of tests under the null hypothesis. That is, for significance level $\alpha = 5\%$, we calculate the values of the test statistics in each simulation and then use the

$100(1-\alpha)$ th percentile of the realized values of test statistics in 1000 simulations. To evaluate the global power of test, we generate 1000 normally distributed samples from this alternative and evaluate the test statistics for each sample. The power of test is calculated as the proportion of the realized values of test statistic larger than the critical value.

Example 1. (global power). We consider different model sizes with $n = 100$ and $a = 50 \times k$ for $k = 1, 2, \ldots, 10$, and want to test the hypotheses

$$H_0\colon \tau_i = 0, \quad \forall i = 1, \ldots, a \text{ against } H_a\colon \tau_1 = 1,\ \tau_2 = \cdots = \tau_a = 0. \tag{55}$$

The simulation results are summarized in Table 1. It is shown that the power of the F test drops significantly as the number a of levels increases while the power of our proposed test drops just slightly. It is clear that our proposed test significantly outperforms the F test.

Example 2. (local power). With different model sizes of $n = 200$ and $a = 2,\ 20,\ 200,\ 2000$, we test

$$H_0\colon \boldsymbol{\tau} = \boldsymbol{\tau}_0 \text{ versus } H_{\text{an}}\colon \boldsymbol{\tau} = (1-\theta)\boldsymbol{\tau}_0 + \theta\boldsymbol{\tau}_1, \tag{56}$$

with $\theta = 0,\ 0.1, \ldots, 1$, $\tau_0 = (0, \ldots, 0)^T$, $\tau_1 = 0.5 \times (-1, \ldots, -1, 1, \ldots, 1)^T$, where τ_1 has the 1st half components of -0.5 and the 2nd half of 0.5. Figure 1 displays the powers of tests, which verifies desired results on the power: when $\theta = 0$, the null and the alternative coincide, so that the power of test should be the significance level α; as θ increases, the alternative gets further away from the null, and the power should become larger. It is seen that the proposed smoothing truncation test has same performance as the F test in low-dimension settings and is much better than the F test in high-dimensional settings. In particular, our test exhibits robust performance as the dimension changes, but the F test has difficulty to distinguish the alternatives from the null.

4.2. A Real Example. In this section, we apply the proposed test and the traditional F test to analyse a breast tumor dataset. This dataset contains 107 cDNA microarray experiments [12]. As indicated in Benito et al. [13], there were two distinct experiment biases in the data which might be from different handling procedures. Jiang et al. [14] corrected the systematic batch biases in the cDNA microarray data and published the batch-adjusted dataset on the website: https://www.stat.unc.edu/postscript/papers/marron/GeneArray/. The data consist of vectors representing relative expression of $a = 5961$ genes for each of which there are $n = 107$ total cases. To perform high-dimensional tests, we keep the samples unchanged for first two genes and centralized and standardized the sample for each of the remaining genes. Hence, in this transformed dataset, the two samples for the first two genes have different means from the others, which results in a high-dimensional sparse setting for hypothesis testing problem (3). Now, we employ the traditional F test and the proposed test for this problem. With $h = 10(\text{an})^{-0.2}$, we calculate the p values of F and T_n as 1 and 0.024. That is, at 5% level, the F test fails in distinguishing the population mean differences, but our test is successful for this testing problem. This is expected, since the F test loses its power from 5959 noise samples and ours wins due to its ability in reducing the contributions of noises.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] G. Casella and R. L. Berger, *Statistical Inference*, Thomson Learning, Boston, MA, USA, 2nd edition, 2002.

[2] S. F. Arnold, "Asymptotic validity of F-tests for the ordinary linear model and the multiple correlation model," *Journal of the American Statistical Association*, vol. 75, no. 372, pp. 890–895, 1980.

[3] M. G. Akritas and N. Papadatos, "Heteroscedastic one-way ANOVA and lack-of-fit tests," *Journal of the American Statistical Association*, vol. 99, pp. 368–382, 2004.

[4] J. Cao and K. J. Worsley, "The detection of local shape changes via the geometry of Hotelling's T^2 fields," *Annals of Statistics*, vol. 27, pp. 925–942, 1999.

[5] J. E. Taylor and K. J. Worsley, "Random fields of multivariate test statistics, with applications to shape analysis," *Annals of Statistics*, vol. 36, pp. 1–27, 2008.

[6] J. Fan, "Test of significance based on wavelet thresholding and neyman's truncation," *Journal of the American Statistical Association*, vol. 91, no. 434, pp. 674–688, 1996.

[7] S. X. Chen, J. Li, and P.-S. Zhong, "Two-sample and ANOVA tests for high dimensional means," *Annals of Statistics*, vol. 47, no. 3, pp. 1443–1474, 2019.

[8] H. Wang, *Rank Tests in Multifactor Heteroscedastic ANOVA and Repeated Measures Designs with Large Number of Levels*, Ph.D. thesis, The Pennsylvania State University, Pennsylvania, PA, USA, 2004.

[9] W. Härdle and E. Mammen, "Comparing nonparametric versus parametric regression fits," *Annals of Statistics*, vol. 21, no. 4, pp. 1926–1947, 1993.

[10] J. D. Hart, *Nonparametric Smoothing and Lack-Of-Fit Tests*, Springer-Verlag, Berlin, Germany, 1997.

[11] J. Jiang, "Multivariate functional-coefficient regression models for nonlinear vector time series data," *Biometrika*, vol. 101, no. 3, pp. 689–702, 2014.

[12] C. M. Perou, T. Sørlie, M. B. Eisen et al., “Molecular portraits of human breast tumours,” *Nature*, vol. 406, no. 6797, pp. 747–752, 2000.

[13] M. Benito, J. Parker, Q. Du et al., “Adjustment of systematic microarray data biases,” *Bioinformatics*, vol. 20, no. 1, pp. 105–114, 2004.

[14] J. Jiang, J. S. Marron, and X. Jiang, “Robust centroid based classification with minimum error rates for high dimension low sample size data,” *Journal of Statistical Planning and Inference*, vol. 139, no. 8, pp. 2571–2580, 2009.