Journal of Mathematical Research with Applications
Jan., 2025, Vol. 45, No. 1, pp. 105–124
DOI:10.3770/j.issn:2095-2651.2025.01.009
Http://jmre.dlut.edu.cn

Limit Laws for the Maximum Interpoint Distance under a 1-Dependent Assumption

Haibin ZHANG, Yong ZHANG, Xue DING*
School of Mathematics, Jilin University, Jilin 130012*, P. R. China*

Abstract Let $\mathcal{M}_{n,p} = (X_{i,k})_{n\times p}$ be an $n \times p$ random matrix whose p columns $\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(p)}$ are an n-dimensional i.i.d. random sample of size p from 1-dependent Gaussian populations. Instead of investigating the special case where p and n are comparable, we consider a much more general case in which $\log n = o(p^{1/3})$. We prove that the maximum interpoint distance $M_n = \max\{|\boldsymbol{X}_i - \boldsymbol{X}_j|; 1 \le i < j \le n\}$ converges to an extreme-value distribution, where $\boldsymbol{X}_i$ and $\boldsymbol{X}_j$ denote the i-th and j-th row of $\mathcal{M}_{n,p}$, respectively. The proofs are completed by using the Chen-Stein Poisson approximation method and the moderation deviation principle.

Keywords maximum interpoint distance; extreme-value distribution; Chen-Stein Poisson approximation; moderation deviation; 1-dependent

MR(2020) Subject Classification 60F05; 62H10

1. Introduction

Let $\mathcal{M}_{n,p} = (X_{i,k})_{n\times p}$ be an $n \times p$ random matrix, where the n rows are $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ and the p columns are $\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(p)}$. In this paper, we are concerned with the limiting distribution of the maximum interpoint distance

$$M_n = \max\{|\boldsymbol{X}_i - \boldsymbol{X}_j|; 1 \le i < j \le n\}, \tag{1.1}$$

where $|\cdot|$ is the Euclidean norm on $\mathbb{R}^p$. We obviously find that the largest interpoint distance is obtained by two vectors $\boldsymbol{X}_i$ and $\boldsymbol{X}_j$ which have almost largest lengths and are almost opposite in direction. Then, the statistic M_n has some applications in detecting outliers [1]. In recent years, several authors have shown that the limiting distribution of M_n is mainly one of the three different extreme-value distributions which are known as Gumbel, Weibull and Fréchet distributions. And the distribution functions of the three distributions are

$$\text{Gumbel}: \exp(-e^{-x}), \quad -\infty < x < \infty;$$
$$\text{Weibull}: \exp(-|x|^d), \quad -\infty < x < \infty,\ d > 0;$$

Received March 29, 2024; Accepted September 2, 2024
Supported by the National Natural Science Foundation of China (Grant Nos. 11771178; 12171198), the Science and Technology Development Program of Jilin Province (Grant No. 20210101467JC), the Technology Program of Jilin Educational Department During the "14th Five-Year" Plan Period (Grant No. JJKH20241239KJ) and the Fundamental Research Funds for the Central Universities.
* Corresponding author
E-mail address: hbzhang21@mails.jlu.edu.cn (Haibin ZHANG); zyong2661@jlu.edu.cn (Yong ZHANG); dingxue83@jlu.edu.cn (Xue DING)

$$\text{Fréchet}: \exp(-x^{-d}), \quad 0 < x < \infty, \ d > 0.$$

In the univariate case $p = 1$, the asymptotic distribution of M_n is well known [2]. Jammalamadaka and Janson [3] supposed that $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ are i.i.d. symmetric real-valued random variables and found the limiting distribution of M_n as follows.

Lemma 1.1 ([3]) *Suppose that there are sequences a_n and b_n of positive numbers such that for any fixed real x, $b_n = o(a_n)$ and $P(|\boldsymbol{X}_i| > a_n + xb_n) = \frac{1+o(1)}{n}e^{-x}$ hold as $n \to \infty$. Then*

$$\frac{M_n - 2a_n}{b_n} + 2\log 2 \to dV_+ + V_-,$$

where V_+ and V_- are two independent random variables with the Gumbel distribution $F_V(x) = e^{-e^{-x}}$.

For the multidimensional situation, it is much more difficult to study the limiting behavior of M_n. Matthews and Rukhin [1] showed that the limiting distribution of M_n is a Gumbel distribution under the special assumption of spherically symmetric normality. Henze and Klein [4] pointed out and corrected some errors in Matthews and Rukhin [1] and obtained the limit law of M_n for a spherically symmetric multivariate sample of the Kotz distribution. Henze and Lao [5] stated that the limiting distribution of M_n is none of the three types of classical extreme-value distributions if $\boldsymbol{X}_i$ obeys a power-tailed spherically decomposable distribution. Appel et al. [6] found that the limiting distribution of M_n is a convolution of two independent Weibull distributions if $\boldsymbol{X}_i$ has a uniform distribution in a planar set with unique major axis and sub-$\sqrt{x}$ decay of its boundary at the endpoints. Appel and Russo [7] considered the case of i.i.d. points distributed uniformly on the surface of a unit hypersphere in $\mathbb{R}^p$ and provided the limiting distribution of the maximum pairwise distance. Mayer and Molchanov [8] proved a limit theorem for the maximum interpoint distance for a sample of n i.i.d. points in the unit p-dimensional ball. Lao [9] derived the limit distribution of M_n when $\boldsymbol{X}_i$ obeys some distributions in the unit square, the uniform distribution in the unit hypercube, and the uniform distribution in a regular convex polygon. Jammalamadaka and Janson [3] obtained a Gumbel limit distribution for M_n when $\boldsymbol{X}_i$ is a $\mathbb{R}^p$-valued random vectors with a spherically symmetric distribution. Schrempp [10] studied the asymptotic behavior of the maximum interpoint distance of random points in a p-dimensional ellipsoid with a unique major axis. Furthermore, Schrempp [11] considered the same problem when the p-dimensional population has a unique diameter and a smooth boundary at the poles. Demichel et al. [12] studied the asymptotic behavior of the diameter or maximum interpoint distance when the random vector $\boldsymbol{X}_i$ in $\mathbb{R}^p$ has an elliptical distribution. Tang et al. [13] proved that M_n^2 under a suitable normalization has an asymptotic Gumbel distribution if the random vector $\boldsymbol{X}_i$ has independent sub-exponential components.

Lemma 1.2 ([13]) *Suppose that $\mathrm{Var}(X_{i,k}) = 1$ and $EX_{i,k}^4 = \kappa < 5$ for $1 \le k \le p$. If $p/n \to \chi \in (0, \infty)$, then*

$$4\sqrt{\log n}\Big(\frac{M_n^2 - 2p}{\sqrt{2(\kappa+1)p}} - 2\sqrt{\log n} + \frac{\log\log n}{4\sqrt{\log n}}\Big) \xrightarrow{d} \xi,$$

where ξ is a random variable whose distribution function is $F_\xi(x) = e^{-Ke^{-x/2}}$ with $K = \frac{1}{4\sqrt{2\pi}}$.

Previous studies preceding Lemma 1.2 primarily focused on fixed dimension p and large sample size n. Lemma 1.2 considered the high-dimensional case for the first time, but its moment conditions and the requirements for p and n are too restrictive. Furthermore, Heiny and Kleemann [14] relaxed the assumption of Lemma 1.2 to two high-dimensional settings: $n = O(p^\tau)$ for some $\tau > 0$ and $\log n = o(p^r)$ for some $0 < r \le 1/2$. And the corresponding moment conditions are also relaxed to $E[|X_{1,1}^{8\tau+4}(\log|X_{1,1}|)^{2\tau+1}] < \infty$ and $E[\exp(t|X_{1,1}|^{4r/(1+r)})] < \infty$ for some $t > 0$. In addition, we refer to Li [15], Modarres [16], Song and Modarres [17,18] for several other investigations on interpoint distance.

The above articles mainly assumed that the components $\{X_{i,k}; 1 \le i \le n, 1 \le k \le p\}$ of the random matrix $\mathcal{M}_{n,p}$ are independent. In practice, the independent assumption is too restrictive. Yan and Feng [19] assumed that $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ are a random sample from a p-dimensional population with dependent sub-gaussian components.

Although assuming the independence of the sample $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ may seem reasonable, verifying the independence of samples can sometimes be quite difficult [20]. Therefore, in this paper, we present the following assumption. We will consider a weak dependent case where two adjacent elements of each column vector are dependent and have a common correlation coefficient ρ_n.

Condition 1.3 *Let $\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(p)}$ be independent identically distributed random vectors with distribution $N_n(\boldsymbol{\mu}, \mathbf{R})$, where $\boldsymbol{\mu} \in \mathbb{R}^n$ is arbitrary and $\mathbf{R}$ has the tridiagonal structure, that is,*

$$\mathbf{R} = \begin{pmatrix} 1 & \rho_n & 0 & \cdots & 0 \\ \rho_n & 1 & \rho_n & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \rho_n \\ 0 & \cdots & 0 & \rho_n & 1 \end{pmatrix}, \tag{1.2}$$

where $\sup_{n\ge1}|\rho_n| < 1/2$. Let $\mathcal{M}_{n,p} = (\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(p)}) = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)' = (X_{i,k})_{n\times p}$.

Remark 1.4 By Horn and Johnson [21, Theorem 6.1.10], we find that $\mathbf{R}$ is positive definite if $|\rho_n| < 1/2$.

Condition 1.5 *Let $\{p_n; n \ge 1\}$ be a sequence of positive integers such that $p_n \to \infty$ with* $\log n = o(p^{1/3})$.

The main probabilistic tools in this paper are (i) the moderation deviation of the partial sum of the independent but not necessarily identically distributed random variables [22, Proposition 4.5], and (ii) the Chen-Stein Poisson approximation method which is a special case of Arratia et al. [23, Theorem 1]. In addition, some researcheres studied the central limit theorem and other weak theorems in a similar way [24–30].

The rest of the paper is organized as follows. We begin in Section 2 presenting some notations and technical tools. Section 2 studies the limiting law of the maximum interpoint distance under a 1-dependent normal population assumption. In Section 5, we list some specific examples and applications. And the proof is given in Section 4.

2. Preliminaries

Before showing the asymptotic behavior of the maximum interpoint distance, we give some notations which will be used in this paper. Throughout this paper, convergence in distribution and convergence in probability are denoted by $\xrightarrow{d}$ and $\xrightarrow{P}$, respectively, and $\xi \stackrel{d}{=} \eta$ implies that ξ and η have the same distribution. For two nonrandom sequences a_n and b_n, $b_n = o(a_n)$ means $\lim_{n\to\infty} b_n/a_n = 0$, and $b_n = O(a_n)$ means $\limsup_{n\to\infty} |b_n/a_n| < \infty$. We denote by C a positive constant independent of n and p, and its value may be different from line to line. In addition, the symbol $\operatorname{sgn}(\rho_n)$ means

$$\operatorname{sgn}(\rho_n) = \begin{cases} 1\,, & \rho_n > 0, \\ 0\,, & \rho_n = 0, \\ -1, & \rho_n < 0. \end{cases}$$

The first ingredient is the Chen–Stein Poisson approximation method.

Lemma 2.1 ([23]) *Let $\{\eta_\alpha; \alpha \in I\}$ be random variables on an index set I and $\{B_\alpha; \alpha \in I\}$ be a set of subsets of I, that is, for each $\alpha \in I$, $B_\alpha \subset I$. For any $t \in \mathbb{R}$, set $\lambda_p = \sum_{\alpha\in I} P(\eta_\alpha > t)$. Then we have*

$$|P(\max_{\alpha\in I} \eta_\alpha \le t) - e^{-\lambda_p}| \le (1 \wedge \lambda_p^{-1})(b_1 + b_2 + b_3),$$

where

$$b_1 = \sum_{\alpha\in I} \sum_{\beta\in B_\alpha} P\left(\eta_\alpha > t\right) P\left(\eta_\beta > t\right),$$

$$b_2 = \sum_{\alpha\in I} \sum_{\alpha\ne\beta\in B_\alpha} P\left(\eta_\alpha > t, \eta_\beta > t\right),$$

$$b_3 = \sum_{\alpha\in I} E|P\left\{\eta_\alpha > t \mid \sigma\left(\eta_\beta; \beta \notin B_\alpha\right)\right\} - P\left(\eta_\alpha > t\right)|,$$

and $\sigma(\eta_\beta; \beta \notin B_\alpha)$ is the σ-algebra generated by $\{\eta_\beta; \beta \notin B_\alpha\}$. In particular, if η_α is independent of $\{\eta_\beta; \beta \notin B_\alpha\}$ for each α, then b_3 vanishes.

The next one is the moderation deviation of the partial sum of the independent but not necessarily identically distributed random variables.

Lemma 2.2 ([22]) *Let $\{\eta_i; 1 \le i \le n\}$ be independent random variables with $E\eta_i = 0$ and $Ee^{h_n|\eta_i|} < \infty$ for some $h_n > 0$ and $1 \le i \le n$. Assume that $\sum_{i=1}^n E\eta_i^2 = 1$. Then*

$$\frac{P(\sum_{i=1}^n \eta_i \ge x)}{1 - \Phi(x)} = 1 + C_n(1 + x^3)\gamma e^{4x^3\gamma}$$

for all $0 \le x \le h_n$ and $\gamma = \sum_{i=1}^n E(|\eta_i|^3 e^{x|\eta_i|})$, where $\Phi(x)$ is the standard normal distribution function, $\sup_{n\ge1}|C_n| \le C$ and C is an absolute constant.

The following lemma can be obtained by using the method same as Fan and Jiang [25, Lemma 3.6]. And we can use it to estimate $E(|\eta|^3 e^{x|\eta|})$.

Lemma 2.3 *Let T, U, V, W be i.i.d. $N(0,1)$-distributed random variables. Let $a_1, a_2, \ldots, a_{11}$ be real numbers. Set $\eta = a_1T^2 + a_2U^2 + a_3V^2 + a_4W^2 + a_5TU + a_6TV + a_7TW + a_8UV + a_9UW + a_{10}VW + a_{11}$. Then*

$$E(|\eta|^3 e^{x|\eta|}) \leq C \cdot e^{x|a_{11}|} \cdot \sum_{i=1}^{10} |a_i|^3,$$

as $0 < x \leq \frac{1}{8\sum_{i=1}^{10}|a_i|}$, where C is a constant not depending on $a_1, a_2, \ldots, a_{11}$.

By applying a similar argument to Fan and Jiang [25, Lemma 3.8], we can derive the following lemma. The detailed proof process will not be described in this paper.

Lemma 2.4 *Let L_n be a random variable for each $n \geq 1$, $h \in \{0,1\}$ and $g > 0$ be a constant satisfying*

$$\lim_{n\to\infty} P(L_n \leq \sqrt{g^2 \log n - \log\log n + h \cdot \log 8 + x}) = F(x)$$

for any $x \in \mathbb{R}$, where $F(x)$ is a continuous distribution function on $\mathbb{R}$. Then

$$L_n = g\sqrt{\log n} - \frac{\log\log n}{2g\sqrt{\log n}} + \frac{h \cdot \log 8}{2g\sqrt{\log n}} + \frac{1}{2g\sqrt{\log n}} U_n,$$

where U_n converges weakly to a probability measure with distribution function $F(x)$.

3. Main results

In this section, we will show the limit distribution of the maximum interpoint distance M_n. Let ξ be a Gumbel distributed random variable with distribution function

$$F_\xi(x) = \exp(-\frac{1}{4\sqrt{2\pi}} e^{-x/2}), \quad x \in \mathbb{R}. \tag{3.1}$$

Set

$$\mu_1 = 2\sqrt{\log n} - \frac{\log\ \log n}{4\sqrt{\log n}} \text{ and } \mu_2 = \sqrt{2\log n} - \frac{\log\log n}{2\sqrt{2\log n}} + \frac{\log 8}{2\sqrt{2\log n}}.$$

Theorem 3.1 *Set $M_{n1} = \max\{|\boldsymbol{X}_i - \boldsymbol{X}_j|; 1 \leq i < j \leq n, i < j-1\}$. Under Conditions 1.3 and 1.5, the following holds as $n \to \infty$:*

$$4\sqrt{\log n}(\frac{M_{n1}^2 - 2p}{\sqrt{p(8 - 6\rho_n^2)}} - \mu_1) \xrightarrow{d} \xi.$$

Theorem 3.2 *Set $M_{n2} = \max\{|\boldsymbol{X}_i - \boldsymbol{X}_j|; 1 \leq i < j \leq n, i = j-1\}$. Under Conditions 1.3 and 1.5, the following holds as $n \to \infty$:*

$$2\sqrt{2\log n}(\frac{M_{n2}^2 - 2p(1-\rho_n)}{\sqrt{p(2\rho_n^2 - 12\rho_n + 8)}} - \mu_2) \xrightarrow{d} \xi.$$

Theorem 3.3 *Under Conditions 1.3 and 1.5, suppose that $\frac{\rho_n\sqrt{p}}{\sqrt{\log n}} \to \lambda$ as $n \to \infty$. Then, the following holds as $n \to \infty$:*

(i) *If* $\lambda \in [-\infty, 2-2\sqrt{2})$, *then*

$$2\sqrt{2\log n}\Big(\frac{M_n^2-2p(1-\rho_n)}{\sqrt{p(2\rho_n^2-12\rho_n+8)}}-\mu_2\Big)\xrightarrow{d}\xi.$$

(ii) *If* $\lambda \in [2-2\sqrt{2}, \infty]$, *then*

$$4\sqrt{\log n}\Big(\frac{M_n^2-2p}{\sqrt{p(8-6\rho_n^2)}}-\mu_1\Big)\xrightarrow{d}\xi.$$

Remark 3.4 Similar to Cai and Jiang [24, Theorem 4], we first study M_{n1} under the 1-dependent normal population assumption. Then, we define a new statistic M_{n2} due to the complexity of the dependent case. Finally, the limiting distribution of $M_n = M_{n1}\vee M_{n2}$ can be obtained by the limiting distributions of M_{n2} when $\lambda \in [-\infty, 2-2\sqrt{2})$, and the limiting distribution of $M_n = M_{n1}\vee M_{n2}$ can be obtained by the limiting distributions of M_{n1} when $\lambda \in [2-2\sqrt{2}, \infty]$.

Theorems 3.1–3.3 imply immediately the following laws of large numbers.

Corollary 3.5 *Under Conditions 1.3 and 1.5, the following holds as* $n\to\infty$*:*

$$\frac{M_{n1}^2-2p}{\sqrt{p(8-6\rho_n^2)\log n}}\xrightarrow{P}2 \text{ and } \frac{M_{n2}^2-2p(1-\rho_n)}{\sqrt{p(2\rho_n^2-12\rho_n+8)\log n}}\xrightarrow{P}\sqrt{2}.$$

Corollary 3.6 *Under Conditions 1.3 and 1.5, suppose that* $\frac{\rho_n\sqrt{p}}{\sqrt{\log n}}\to\lambda$ *as* $n\to\infty$*. Then, the following holds as* $n\to\infty$*:*

(i) *If* $\lambda \in [-\infty, 2-2\sqrt{2})$, *then*

$$\frac{M_n^2-2p(1-\rho_n)}{\sqrt{p(2\rho_n^2-12\rho_n+8)\log n}}\xrightarrow{P}\sqrt{2}.$$

(ii) *If* $\lambda \in [2-2\sqrt{2}, \infty]$, *then*

$$\frac{M_n^2-2p}{\sqrt{p(8-6\rho_n^2)\log n}}\xrightarrow{P}2.$$

The following is an obvious consequence of Theorem 3.3.

Corollary 3.7 *Let* $\rho\in(-1/2,1/2)$ *be fixed. Under Conditions 1.3 and 1.5, the following holds as* $n\to\infty$*:*

$$4\sqrt{\log n}\Big(\frac{M_n^2-2p}{\sqrt{p(8-6\rho^2)}}-2\sqrt{\log n}+\frac{\log\log n}{4\sqrt{\log n}}\Big)\xrightarrow{d}\xi.$$

4. Examples

In this section, we will show several examples based on the main results discussed above.

Example 4.1 Let $\{\xi_i; i=0,1,\ldots,n\}$ be independent and identically distributed standard normal random variables. Set

$$\boldsymbol{X}^{(k)}=(\mathrm{x}_1,\ldots,\mathrm{x}_n)'=\sqrt{\theta}(\xi_0,\ldots,\xi_{n-1})'+\sqrt{1-\theta}(\xi_1,\ldots,\xi_n)'$$

for $\theta \in [0,1]$. Obviously, $\mathrm{x}_i \sim N(0,1)$ and

$$\mathrm{Cov}(\mathrm{x}_i, \mathrm{x}_j) = \begin{cases} 0, & |i-j| > 1, \\ \sqrt{\theta(1-\theta)}, & |i-j| = 1 \end{cases}$$

for $1 \le i, j \le n$. Therefore, $\boldsymbol{X}^{(k)}$ obeys a 1-dependent normal distribution. The decomposition structure mentioned above is crucial in the proofs of theorems.

Example 4.2 Assume that $\boldsymbol{X}^{(k)}$ follows a standard multivariate normal distribution in $\mathbb{R}^n$. Under Condition 1.5, Theorem 3.3 yields

$$4\sqrt{\log n}\Big(\frac{M_n^2 - 2p}{2\sqrt{2p}} - 2\sqrt{\log n} + \frac{\log\, \log n}{4\sqrt{\log n}}\Big) \xrightarrow{d} \xi,$$

as shown by Tang et al. [13].

Example 4.3 Test the covariance structure. Evaluating the covariance structure of a distribution is a crucial issue in statistical inference for high-dimensional data. Let $\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(p)}$ be a random sample from the population $N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. As the application of Theorem 3.3, we wish to test whether $\boldsymbol{X}^{(k)}$ comes from a 1-dependent normal population, or equivalently in terms of the correlation matrix, we wish to test whether $\boldsymbol{\Sigma}$ has the tridiagonal structure like $\mathbf{R}$ as in (1.2). The specific hypotheses are

$$H_0 : \boldsymbol{\Sigma} = \mathbf{R} \text{ v.s. } H_1 : \boldsymbol{\Sigma} \neq \mathbf{R}.$$

M_n in (1.1) can be used as a test statistic. Set $q_\alpha = -\log(32\pi) - 2\log\log(1-\alpha/2)^{-1}$ for $0 < \alpha < 1$. It is obvious to find that q_α is the $(1-\alpha/2)$-quantile of the distribution $F_\xi(x)$ in (3.1). Then, we can get the following rejection regions $\mathcal{K}_1$ and $\mathcal{K}_2$ when $\lambda \in [-\infty, 2-2\sqrt{2})$ and $\lambda \in [2-2\sqrt{2}, \infty]$, respectively.

$$\mathcal{K}_1 = \{\Big|\frac{M_n^2 - 2p(1-\rho_n)}{\sqrt{p(2\rho_n^2 - 12\rho_n + 8)}} - \sqrt{2\log n} + \frac{\log\log n}{2\sqrt{2\log n}} - \frac{\log 8}{2\sqrt{2\log n}}\Big| \ge \frac{q_\alpha}{2\sqrt{2\log n}}\},$$
$$\mathcal{K}_2 = \{\Big|\frac{M_n^2 - 2p}{\sqrt{p(8-6\rho_n^2)}} - 2\sqrt{\log n} + \frac{\log\, \log n}{4\sqrt{\log n}}\Big| \ge \frac{q_\alpha}{4\sqrt{\log n}}\}.$$

In particular, one can test the independence of the Gaussian random variables when $\mathbf{R} = \mathbf{I}_n$. And the corresponding rejection region is

$$\mathcal{K}_3 = \{\Big|\frac{M_n^2 - 2p}{2\sqrt{2p}} - 2\sqrt{\log n} + \frac{\log\, \log n}{4\sqrt{\log n}}\Big| \ge \frac{q_\alpha}{4\sqrt{\log n}}\}.$$

Example 4.4 Test for the presence of outlier points. Outlier detection refers to the process of identifying and analyzing observations in a dataset that are significantly different from other observations. And outlier detection techniques are commonly used in various fields, including finance, healthcare, and environmental science, to identify these anomalous observations and either remove them from the dataset or further investigate them. The largest interpoint distance test is a reliable method for detecting the presence of a potential outlier in the data. Let $\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(p)}$ be a random sample from an n-variate normal population with the covariance matrix $\Sigma_{n\times n} = \mathbf{R}$ as in (1.2). We wish to test the hypothesis

$$H_0: \text{ There is no outlier point among } \boldsymbol{X}_1, \ldots, \boldsymbol{X}_n.$$

Similar to the argument in Example 4.3, we choose M_n as the test statistic. Under Condition 1.5, two rejection regions of the asymptotic size-α test are given by $\mathcal{K}_1$ and $\mathcal{K}_2$ when $\lambda \in [-\infty, 2-2\sqrt{2})$ and $\lambda \in [2-2\sqrt{2}, \infty]$, respectively.

5. Proofs of main results

Now we are in a position to prove Theorems 3.1–3.3.

For any $\rho_n \in (-1/2, 1/2)$, there exists a $\theta_n \in [0, 1/2)$ such that $\sqrt{\theta_n(1-\theta_n)} = |\rho_n|$. So we substitute $\sqrt{\theta_n(1-\theta_n)}$ for $|\rho_n|$ in the proofs of theorems. Assume that the random variables

$$\{\xi_k, \xi_{i,k}; k = 1, 2, \ldots, i = 0, 1, 2, \ldots\} \text{ are i.i.d. as } N(0,1). \tag{5.1}$$

Given $\theta_n \in [0, 1/2)$ for each $n \geq 1$, set

$$a_n = \theta_n,\ b_n = 1-\theta_n,\ c_n = \sqrt{\theta_n(1-\theta_n)}. \tag{5.2}$$

For $x \in \mathbb{R}$ and integer $n \geq 1$, set

$$s_n = \sqrt{4\log n - \log\log n + x} \tag{5.3}$$

and

$$s_n' = \sqrt{2\log n - \log\log n + \log 8 + x}. \tag{5.4}$$

Since M_{n1}, M_{n2} and M_n are invariant under translation of the vectors $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$, we assume that $\boldsymbol{\mu} = \mathbf{0}$ without loss of generality. Review (5.1). Write

$$x_{i,k} = \operatorname{sgn}(\rho_n)\cdot\sqrt{\theta_n}\ \xi_{i-1,k} + \sqrt{1-\theta_n}\xi_{i,k},\quad 1 \leq i \leq n,\ 1 \leq k \leq p. \tag{5.5}$$

We find that the p columns of the matrix $(x_{i,k})_{n\times p}$ are p i.i.d. random vectors and $x_{l,1} \sim N(0,1)$ for each $1 \leq l \leq n$. In addition, for $1 \leq i, j \leq n$, we have

$$\operatorname{Cov}(x_{i,1}, x_{j,1}) = \begin{cases} 0, & |i-j| > 1, \\ \operatorname{sgn}(\rho_n)\sqrt{\theta_n(1-\theta_n)}, & |i-j| = 1, \end{cases}$$

where $\operatorname{sgn}(\rho_n)\sqrt{\theta_n(1-\theta_n)} = \rho_n$. That is, each column of the matrix $(x_{i,k})_{n\times p}$ follows $N_n(\mathbf{0}, \mathbf{R})$. Therefore, $\mathcal{M}_{n,p}$ and $(x_{i,k})_{n\times p}$ have the same distribution. Recall (1.1), (5.2) and (5.5). Write

$$\begin{aligned}
|\boldsymbol{X}_i - \boldsymbol{X}_j|^2 &= \sum_{k=1}^{p}(x_{i,k}^2 + x_{j,k}^2 - 2x_{i,k}x_{j,k}) \\
&= \sum_{k=1}^{p}[a_n(\xi_{i-1,k}^2 + \xi_{j-1,k}^2 - 2\xi_{i-1,k}\xi_{j-1,k}) + b_n(\xi_{i,k}^2 + \xi_{j,k}^2 - 2\xi_{i,k}\xi_{j,k})] + \\
&\quad \operatorname{sgn}(\rho_n)\cdot c_n \sum_{k=1}^{p}(\xi_{i-1,k}\xi_{i,k} + \xi_{j-1,k}\xi_{j,k} - 2\xi_{i-1,k}\xi_{j,k} - 2\xi_{i,k}\xi_{j-1,k}) \\
&:= \sum_{k=1}^{p}\eta_{ijk}.
\end{aligned} \tag{5.6}$$

5.1. Proof of Theorem 3.1

In this section, we will prove Theorem 3.1 by Lemma 2.1 and provide a detailed proof.

Proof of Theorem 3.1 Set $I' = \{(i,j); 1 \le i < j \le n, i < j-1\}$. For $\alpha = (i,j) \in I'$, define

$$X_\alpha = \frac{1}{\sqrt{p}\sigma_{n1}} \sum_{k=1}^{p} (\eta_{ijk} - 2)$$

and

$$B'_\alpha = \{(k,l) \in I';\ \{k,l\} \cap \{i-1, i, i+1, j-1, j, j+1\} \ne \emptyset \text{ but } (k,l) \ne \alpha\},$$

where $\sigma_{n1}^2 := 6\theta_n^2 - 6\theta_n + 8$. Then we find that random variable X_α is independent of $\{X_\beta; \beta \notin B'_\alpha\}$. By Lemma 2.1, we have

$$\Big| P\Big(\max_{\alpha \in I'} X_\alpha \le s_n\Big) - e^{-\lambda_{p1}} \Big| \le u_1 + u_2,$$

where

$$\lambda_{p1} = \sum_{\alpha \in I'} P(X_\alpha > s_n) = \frac{(n-1)(n-2)}{2} P\Big(\frac{1}{\sqrt{p}\sigma_{n1}} \sum_{k=1}^{p} (\eta_{13k} - 2) > s_n \Big),$$

$$\begin{aligned} u_1 &= \sum_{\alpha \in I'} \sum_{\beta \in B'_\alpha} P(X_\alpha > s_n) P(X_\beta > s_n) \\ &\le \frac{(n-1)(n-2)}{2} \cdot (6n) \cdot P\Big(\frac{1}{\sqrt{p}\sigma_{n1}} \sum_{k=1}^{p} (\eta_{13k} - 2) > s_n \Big)^2 \end{aligned}$$

and

$$\begin{aligned} u_2 &= \sum_{\alpha \in I'} \sum_{\beta \in B'_\alpha} P(X_\alpha > s_n, X_\beta > s_n) \\ &\le \frac{(n-1)(n-2)}{2} \cdot (6n) \cdot P\Big(\frac{1}{\sqrt{p}\sigma_{n1}} \sum_{k=1}^{p} (\eta_{13k} - 2) > s_n, \frac{1}{\sqrt{p}\sigma_{n1}} \sum_{k=1}^{p} (\eta_{15k} - 2) > s_n \Big). \end{aligned}$$

Note that there are other possible combinations when estimating u_2. In this paper, we provide a detailed proof for only one set; the other combinations can be derived using the same method.

Write

$$\begin{aligned} \eta_{13k} =& a_n(\xi_{0,k}^2 + \xi_{2,k}^2 - 2\xi_{0,k}\xi_{2,k}) + b_n(\xi_{1,k}^2 + \xi_{3,k}^2 - 2\xi_{1,k}\xi_{3,k}) + \\ & \text{sgn}(\rho_n) \cdot c_n(\xi_{0,k}\xi_{1,k} + \xi_{2,k}\xi_{3,k} - 2\xi_{0,k}\xi_{3,k} - 2\xi_{1,k}\xi_{2,k}) \end{aligned}$$

for each k. Obviously,

$$E(\eta_{13k}) = 2 \text{ and } \text{Var}(\eta_{13k}) = 6\theta_n^2 - 6\theta_n + 8 := \sigma_{n1}^2. \tag{5.7}$$

Then, $\sigma_{n1}^2 \in (13/2, 8]$ due to $\theta_n \in [0, 1/2)$. Denote

$$a = \frac{a_n}{\sqrt{p}\sigma_{n1}},\ b = \frac{b_n}{\sqrt{p}\sigma_{n1}},\ c = \frac{\text{sgn}(\rho_n) \cdot c_n}{\sqrt{p}\sigma_{n1}},\ d = -\frac{2}{\sqrt{p}\sigma_{n1}}.$$

Set

$$\eta_k = a(\xi_{0,k}^2 + \xi_{2,k}^2 - 2\xi_{0,k}\xi_{2,k}) + b(\xi_{1,k}^2 + \xi_{3,k}^2 - 2\xi_{1,k}\xi_{3,k}) +$$

$$c(\xi_{0,k}\xi_{1,k}+\xi_{2,k}\xi_{3,k}-2\xi_{0,k}\xi_{3,k}-2\xi_{1,k}\xi_{2,k})+d.$$

Then it follows from (5.7) that

$$E(\eta_k)=0 \text{ and } \sum_{k=1}^{p}\mathrm{Var}(\eta_k)=1. \tag{5.8}$$

Furthermore, we have

$$\max\{|a|,|b|,|c|,|d|\}\le\frac{1}{\sqrt{p}}. \tag{5.9}$$

Then, by using the Hölder inequality, the fact that $|\xi_{1,k}\xi_{2,k}|\le\xi_{1,k}^2+\xi_{2,k}^2$ and independence, we see

$$\begin{aligned}Ee^{h|\eta_k|}&\le E\exp[\frac{4h}{\sqrt{p}}(\xi_{0,k}^2+\xi_{1,k}^2+\xi_{2,k}^2+\xi_{3,k}^2)]\cdot e^{\frac{h}{\sqrt{p}}}\\&\le E\exp(\frac{4h}{\sqrt{p}}\xi_{0,k}^2)\cdot E\exp(\frac{4h}{\sqrt{p}}\xi_{1,k}^2)\cdot E\exp(\frac{4h}{\sqrt{p}}\xi_{2,k}^2)\cdot E\exp(\frac{4h}{\sqrt{p}}\xi_{3,k}^2)\cdot e^{\frac{h}{\sqrt{p}}}<\infty\end{aligned} \tag{5.10}$$

for all h,k,p satisfying $0<h<h_p:=\frac{\sqrt{p}}{16}$ and $1\le k\le p$. By Lemma 2.3, we have

$$\gamma:=\sum_{k=1}^{p}E(|\eta_k|^3e^{s_n|\eta_k|})\le\sum_{k=1}^{p}C\cdot(|a|^3+|b|^3+|c|^3)e^{s_n/\sqrt{p}}\le\sum_{k=1}^{p}\frac{C\cdot e^{s_n/\sqrt{p}}}{p^{3/2}}\le\frac{C\cdot e^{s_n/\sqrt{p}}}{\sqrt{p}}.$$

We first estimate λ_{p1}. By Condition 1.5, $s_n^3\gamma=O(s_n^3p^{-1/2}e^{s_n/\sqrt{p}})\to0$ as $n\to\infty$. From (5.8), (5.10), Lemma 2.2 and the formula $P(N(0,1)\ge x)=\frac{1}{\sqrt{2\pi}x}e^{-x^2/2}(1+o(1))$ as $x\to\infty$, we conclude

$$\begin{aligned}\frac{n^2}{2}\cdot P\Big(\frac{1}{\sqrt{p}\sigma_{n1}}\sum_{k=1}^{p}(\eta_{13k}-2)>s_n\Big)&=\frac{n^2}{2}\cdot P\Big(\sum_{k=1}^{p}\eta_k>s_n\Big)\\&=\frac{n^2}{2}\cdot[1-\Phi(s_n)]\cdot[1+O(1)(1+s_n^3)\gamma e^{4s_n^3\gamma}]\\&=\frac{n^2}{2\sqrt{2\pi}s_n}e^{-\frac{s_n^2}{2}}[1+o(1)]\to\frac{1}{4\sqrt{2\pi}}e^{-\frac{x}{2}}\end{aligned} \tag{5.11}$$

as $n\to\infty$. Thus, $\lim_{n\to\infty}\lambda_{p1}=\frac{1}{4\sqrt{2\pi}}e^{-\frac{x}{2}}$ for any $x\in\mathbb{R}$.

Then, we can see from (5.11) that

$$P\Big(\frac{1}{\sqrt{p}\sigma_{n1}}\sum_{k=1}^{p}(\eta_{13k}-2)>s_n\Big)^2=O(\frac{1}{n^4}) \tag{5.12}$$

as $n\to\infty$, which means $\lim_{n\to\infty}u_1=0$.

Finally, for u_2, it is obvious to find that

$$\begin{aligned}u_2&\le\frac{(n-1)(n-2)}{2}\cdot(6n)\cdot P\Big(\frac{1}{\sqrt{p}\sigma_{n1}}\sum_{k=1}^{p}(\eta_{13k}-2)>s_n,\frac{1}{\sqrt{p}\sigma_{n1}}\sum_{k=1}^{p}(\eta_{15k}-2)>s_n\Big)\\&\le\frac{(n-1)(n-2)}{2}\cdot(6n)\cdot P\Big(\frac{1}{\sqrt{p}\sigma_{n1}}\sum_{k=1}^{p}(\eta_{13k}+\eta_{15k}-4)>2s_n\Big).\end{aligned}$$

Write

$$\eta_{15k}=a_n(\xi_{0,k}^2+\xi_{4,k}^2-2\xi_{0,k}\xi_{4,k})+b_n(\xi_{1,k}^2+\xi_{5,k}^2-2\xi_{1,k}\xi_{5,k})+$$

$$\mathrm{sgn}(\rho_n)\cdot c_n(\xi_{0,k}\xi_{1,k}+\xi_{4,k}\xi_{5,k}-2\xi_{0,k}\xi_{5,k}-2\xi_{1,k}\xi_{4,k}).$$

By some calculations, we can get that

$$E(\eta_{13k}+\eta_{15k}-4)=0 \text{ and } \mathrm{Var}(\eta_{13k}+\eta_{15k}-4)=18\theta_n^2-18\theta_n+20:=\sigma_{n2}^2.$$

Then, $\sigma_{n2}^2\in(31/2,20]$ and $\sigma_{n1}/\sigma_{n2}>\sqrt{\frac{2}{5}}$ due to $\theta_n\in[0,1/2)$. Thus,

$$E(\frac{\eta_{13k}+\eta_{15k}-4}{\sqrt{p}\sigma_{n2}})=0 \text{ and } \sum_{k=1}^{p}\mathrm{Var}(\frac{\eta_{13k}+\eta_{15k}-4}{\sqrt{p}\sigma_{n2}})=1. \tag{5.13}$$

Furthermore,

$$E\exp(\frac{h|\eta_{13k}+\eta_{15k}-4|}{\sqrt{p}\sigma_{n2}})<\infty \tag{5.14}$$

for all h,k,p satisfying $0<h<h_p:=\frac{\sqrt{p}}{16}$ and $1\le k\le p$. By the same argument as in (5.11), we obtain that

$$\begin{aligned}
&P\Big(\frac{1}{\sqrt{p}\sigma_{n1}}\sum_{k=1}^{p}(\eta_{13k}+\eta_{15k}-4)>2s_n\Big)\\
&=P\Big(\frac{1}{\sqrt{p}\sigma_{n2}}\sum_{k=1}^{p}(\eta_{13k}+\eta_{15k}-4)>\frac{2\sigma_{n1}}{\sigma_{n2}}s_n\Big)\\
&\le P\Big(\sum_{k=1}^{p}\frac{\eta_{13k}+\eta_{15k}-4}{\sqrt{p}\sigma_{n2}}>\frac{2\sqrt{10}s_n}{5}\Big) \qquad (5.15)\\
&=[1-\Phi(\frac{2\sqrt{10}s_n}{5})]\cdot[1+o(1)]\\
&=\frac{\sqrt{5}}{4\sqrt{\pi}s_n}e^{-\frac{4s_n^2}{5}}[1+o(1)]=o(\frac{1}{n^3}) \qquad (5.16)
\end{aligned}$$

as $n\to\infty$. In summary, $\lim_{n\to\infty}u_2=0$. Thus,

$$\begin{aligned}
\lim_{n\to\infty}P\Big(\max_{\alpha\in I'}X_\alpha\le s_n\Big)=&\lim_{n\to\infty}P\Big(\max_{1\le i<j\le n,i<j-1}\frac{1}{\sqrt{p}\sigma_{n1}}\sum_{k=1}^{p}(\eta_{ijk}-2)\le s_n\Big)\\
=&\exp(-\frac{1}{4\sqrt{2\pi}}e^{-\frac{x}{2}})
\end{aligned}$$

for any $x\in\mathbb{R}$. Then, by Lemma 2.4,

$$\max_{1\le i<j\le n,i<j-1}\frac{1}{\sqrt{p}\sigma_{n1}}\sum_{k=1}^{p}(\eta_{ijk}-2)=2\sqrt{\log n}-\frac{\log\log n}{4\sqrt{\log n}}+\frac{U_{n1}}{4\sqrt{\log n}},$$

where $U_{n1}\xrightarrow{d}\xi$ with distribution function $F_\xi(x)$ in (3.1). Reviewing (5.6), we have

$$\begin{aligned}
M_{n1}^2=&\max_{1\le i<j\le n,i<j-1}|\boldsymbol{X}_i-\boldsymbol{X}_j|^2=\max_{1\le i<j\le n,i<j-1}\sum_{k=1}^{p}\eta_{ijk}\\
=&\sqrt{p}\sigma_{n1}(2\sqrt{\log n}-\frac{\log\log n}{4\sqrt{\log n}}+\frac{U_{n1}}{4\sqrt{\log n}})+2p.
\end{aligned}$$

It follows that

$$4\sqrt{\log n}(\frac{M_{n1}^2-2p}{\sqrt{p(8-6\rho_n^2)}}-2\sqrt{\log n}+\frac{\log\log n}{4\sqrt{\log n}})\xrightarrow{d}\xi,$$

where the distribution function of ξ is given in (3.1). □

5.2. Proof of Theorem 3.2

The main goal of this section is to prove Theorem 3.2.

Proof of Theorem 3.2 Set $I'' = \{(i,j); 1 \le i < j \le n, i = j-1\}$. For $\alpha = (i,j) \in I''$, define

$$Y_\alpha = \frac{1}{\sqrt{p}\sigma_{n3}} \sum_{k=1}^{p} (\eta_{ijk} - 2(1-\rho_n))$$

and

$$B''_\alpha = \{(k,l) \in I'';\ k \in \{i-2, i-1, i, i+1, i+2\} \text{ but } (k,l) \ne \alpha\},$$

where $\sigma_{n3}^2 := 2\rho_n^2 - 12\rho_n + 8$. Note that random variable Y_α is independent of $\{Y_\beta; \beta \notin B''_\alpha\}$. By Lemma 2.1, we have

$$\Big|P\Big(\max_{\alpha \in I''} Y_\alpha \le s'_n\Big) - e^{-\lambda_{p2}}\Big| \le v_1 + v_2,$$

where

$$\lambda_{p2} = \sum_{\alpha \in I''} P(Y_\alpha > s'_n) = (n-1)P\Big(\frac{1}{\sqrt{p}\sigma_{n3}} \sum_{k=1}^{p} (\eta_{12k} - 2(1-\rho_n)) > s'_n\Big),$$

$$v_1 = \sum_{\alpha \in I''} \sum_{\beta \in B''_\alpha} P(Y_\alpha > s'_n) P(Y_\beta > s'_n) \le (n-1) \cdot 5 \cdot P\Big(\frac{1}{\sqrt{p}\sigma_{n3}} \sum_{k=1}^{p} (\eta_{12k} - 2(1-\rho_n)) > s'_n\Big)^2$$

and

$$\begin{aligned} v_2 &= \sum_{\alpha \in I''} \sum_{\alpha \ne \beta \in B''_\alpha} P(Y_\alpha > s'_n, Y_\beta > s'_n) \\ &\le (n-1) \cdot 5 \cdot P\Big(\frac{1}{\sqrt{p}\sigma_{n3}} \sum_{k=1}^{p} (\eta_{12k} - 2(1-\rho_n)) > s'_n, \frac{1}{\sqrt{p}\sigma_{n3}} \sum_{k=1}^{p} (\eta_{23k} - 2(1-\rho_n)) > s'_n\Big). \end{aligned}$$

Write

$$\begin{aligned} \eta_{12k} =& a_n(\xi_{0,k}^2 + \xi_{1,k}^2 - 2\xi_{0,k}\xi_{1,k}) + b_n(\xi_{1,k}^2 + \xi_{2,k}^2 - 2\xi_{1,k}\xi_{2,k}) + \\ & \operatorname{sgn}(\rho_n) \cdot c_n(\xi_{0,k}\xi_{1,k} + \xi_{1,k}\xi_{2,k} - 2\xi_{0,k}\xi_{2,k} - 2\xi_{1,k}^2) \end{aligned}$$

for each k. We have

$$E(\eta_{12k}) = 2(1-\rho_n) \text{ and } \operatorname{Var}(\eta_{12k}) = 2\rho_n^2 - 12\rho_n + 8 := \sigma_{n3}^2, \tag{5.17}$$

where $\rho_n = \operatorname{sgn}(\rho_n) \cdot c_n \in (-1/2, 1/2)$. Then, $\sigma_{n3}^2 \in (5/2, 29/2)$. Denote

$$a' = \frac{a_n}{\sqrt{p}\sigma_{n3}},\ b' = \frac{b_n}{\sqrt{p}\sigma_{n3}},\ c' = \frac{\operatorname{sgn}(\rho_n) \cdot c_n}{\sqrt{p}\sigma_{n3}},\ d' = -\frac{2(1-\rho_n)}{\sqrt{p}\sigma_{n3}}.$$

Set

$$\begin{aligned} \eta'_k =& a'(\xi_{0,k}^2 + \xi_{1,k}^2 - 2\xi_{0,k}\xi_{1,k}) + b'(\xi_{1,k}^2 + \xi_{2,k}^2 - 2\xi_{1,k}\xi_{2,k}) + \\ & c'(\xi_{0,k}\xi_{1,k} + \xi_{1,k}\xi_{2,k} - 2\xi_{0,k}\xi_{2,k} - 2\xi_{1,k}^2) + d'. \end{aligned}$$

Then it follows from (5.17) that

$$E(\eta'_k) = 0 \text{ and } \sum_{k=1}^{p} \mathrm{Var}(\eta'_k) = 1. \tag{5.18}$$

Furthermore, we have

$$\max\{|a'|, |b'|, |c'|\} \leq \frac{1}{\sqrt{p}} \text{ and } |d'| \leq \frac{2}{\sqrt{p}}.$$

Then, use the Hölder inequality, the fact that $|\xi_{1,k}\xi_{2,k}| \leq \xi_{1,k}^2 + \xi_{2,k}^2$ and independence to see

$$\begin{aligned} Ee^{h|\eta'_k|} &\leq E\exp[\frac{4h}{\sqrt{p}\sigma_{n3}}(\xi_{0,k}^2 + \xi_{1,k}^2 + \xi_{2,k}^2)] \cdot e^{h|d'|} \\ &\leq E\exp(\frac{4h}{\sqrt{p}}\xi_{0,k}^2) \cdot E\exp(\frac{4h}{\sqrt{p}}\xi_{1,k}^2) \cdot E\exp(\frac{4h}{\sqrt{p}}\xi_{2,k}^2) \cdot e^{\frac{2h}{\sqrt{p}}} < \infty \end{aligned} \tag{5.19}$$

for all h, k, p satisfying $0 < h < h_p := \frac{\sqrt{p}}{16}$ and $1 \leq k \leq p$. By Lemma 2.3, we have

$$\gamma := \sum_{k=1}^{p} E(|\eta'_k|e^{s'_n|\eta'_k|}) \leq \sum_{k=1}^{p} C \cdot (|a'|^3 + |b'|^3 + |c'|^3)e^{s'_n/\sqrt{p}} \leq \sum_{k=1}^{p} \frac{C \cdot e^{s'_n/\sqrt{p}}}{p^{3/2}} \leq \frac{C \cdot e^{s'_n/\sqrt{p}}}{\sqrt{p}}.$$

We first estimate λ_{p2}. By Condition 1.5, $s_n'^3\gamma = O(s_n'^3 p^{-1/2} e^{s'_n/\sqrt{p}}) \to 0$ as $n \to \infty$. By (5.18), (5.19), Lemma 2.2 and the formula $P(N(0,1) \geq x) = \frac{1}{\sqrt{2\pi}x}e^{-x^2/2}(1 + o(1))$ as $x \to \infty$, we conclude

$$\begin{aligned} n \cdot P\Big(\frac{1}{\sqrt{p}\sigma_{n3}}\sum_{k=1}^{p}(\eta_{12k} - 2(1-\rho_n)) > s'_n\Big) &= n \cdot P\Big(\sum_{k=1}^{p}\eta'_k > s'_n\Big) \\ &= n \cdot [1 - \Phi(s'_n)] \cdot [1 + O(1)(1 + s_n'^3)\gamma e^{4s_n'^3\gamma}] \\ &= \frac{n}{\sqrt{2\pi}s'_n}e^{-\frac{s_n'^2}{2}}[1 + o(1)] \to \frac{1}{4\sqrt{2\pi}}e^{-\frac{x}{2}} \end{aligned} \tag{5.20}$$

as $n \to \infty$. Thus, $\lim_{n\to\infty}\lambda_{p2} = \frac{1}{4\sqrt{2\pi}}e^{-\frac{x}{2}}$ for any $x \in \mathbb{R}$.

Then, we can see from (5.20) that

$$P\Big(\frac{1}{\sqrt{p}\sigma_{n3}}\sum_{k=1}^{p}(\eta_{12k} - 2(1-\rho_n)) > s'_n\Big)^2 = O(\frac{1}{n^2}) \tag{5.21}$$

as $n \to \infty$, which means $\lim_{n\to\infty} v_1 = 0$.

Finally, for v_2, it is obvious to find that

$$\begin{aligned} v_2 &\leq 5(n-1)P\Big(\frac{1}{\sqrt{p}\sigma_{n3}}\sum_{k=1}^{p}(\eta_{12k} - 2(1-\rho_n)) > s'_n, \frac{1}{\sqrt{p}\sigma_{n3}}\sum_{k=1}^{p}(\eta_{23k} - 2(1-\rho_n)) > s'_n\Big) \\ &\leq 5(n-1)P\Big(\frac{1}{\sqrt{p}\sigma_{n3}}\sum_{k=1}^{p}(\eta_{12k} + \eta_{23k} - 4(1-\rho_n)) > 2s'_n\Big). \end{aligned}$$

Write

$$\begin{aligned} \eta_{23k} =& a_n(\xi_{1,k}^2 + \xi_{2,k}^2 - 2\xi_{1,k}\xi_{2,k}) + b_n(\xi_{2,k}^2 + \xi_{3,k}^2 - 2\xi_{2,k}\xi_{3,k}) + \\ & \mathrm{sgn}(\rho_n) \cdot c_n(\xi_{1,k}\xi_{2,k} + \xi_{2,k}\xi_{3,k} - 2\xi_{1,k}\xi_{3,k} - 2\xi_{2,k}^2). \end{aligned}$$

By some calculations, we can get that

$$E(\eta_{12k}+\eta_{23k}-4(1-\rho_n))=0,\quad \mathrm{Var}(\eta_{12k}+\eta_{23k}-4(1-\rho_n))=14\rho_n^2-36\rho_n+20:=\sigma_{n4}^2.$$

Then, $\sigma_{n4}^2\in(11/2,83/2)$ and $\sigma_{n3}/\sigma_{n4}>\sqrt{\frac{29}{83}}$ due to $\rho_n\in(-1/2,1/2)$. Thus,

$$E(\frac{\eta_{12k}+\eta_{23k}-4(1-\rho_n)}{\sqrt{p}\sigma_{n4}})=0 \text{ and } \sum_{k=1}^{p}\mathrm{Var}(\frac{\eta_{12k}+\eta_{23k}-4(1-\rho_n)}{\sqrt{p}\sigma_{n4}})=1. \tag{5.22}$$

Furthermore,

$$E\exp(\frac{h|\eta_{12k}+\eta_{23k}-4(1-\rho_n)|}{\sqrt{p}\sigma_{n4}})<\infty \tag{5.23}$$

for all h,k,p satisfying $0<h<h_p:=\frac{\sqrt{p}}{16}$ and $1\le k\le p$. By the same argument as in (5.20), we obtain that

$$\begin{aligned}
&P\Big(\frac{1}{\sqrt{p}\sigma_{n3}}\sum_{k=1}^{p}(\eta_{12k}+\eta_{23k}-4(1-\rho_n))>2s_n'\Big)\\
&=P\Big(\frac{1}{\sqrt{p}\sigma_{n4}}\sum_{k=1}^{p}(\eta_{12k}+\eta_{23k}-4(1-\rho_n))>\frac{2\sigma_{n3}}{\sigma_{n4}}s_n'\Big)\\
&\le P\Big(\sum_{k=1}^{p}\frac{\eta_{12k}+\eta_{23k}-4(1-\rho_n)}{\sqrt{p}\sigma_{n4}}>2\sqrt{\frac{29}{83}}s_n'\Big)\\
&=[1-\Phi(2\sqrt{\frac{29}{83}}s_n')]\cdot[1+o(1)]\\
&=\frac{C}{s_n'}e^{-\frac{58s_n'^2}{83}}[1+o(1)]=o(\frac{1}{n})
\end{aligned} \tag{5.24}$$

as $n\to\infty$. In summary, $\lim_{n\to\infty}v_2=0$. Thus,

$$\begin{aligned}
\lim_{n\to\infty}P\Big(\max_{\alpha\in I''}Y_\alpha\le s_n'\Big)&=\lim_{n\to\infty}P\Big(\max_{1\le i<j\le n,i=j-1}\frac{1}{\sqrt{p}\sigma_{n3}}\sum_{k=1}^{p}(\eta_{ijk}-2(1-\rho_n))\le s_n'\Big)\\
&=\exp(-\frac{1}{4\sqrt{2\pi}}e^{-\frac{x}{2}})
\end{aligned}$$

for any $x\in\mathbb{R}$. Then, by Lemma 2.4,

$$\begin{aligned}
&\max_{1\le i<j\le n,i=j-1}\frac{1}{\sqrt{p}\sigma_{n3}}\sum_{k=1}^{p}(\eta_{ijk}-2(1-\rho_n))\\
&=\sqrt{2\log n}-\frac{\log\log n}{2\sqrt{2\log n}}+\frac{\log 8}{2\sqrt{2\log n}}+\frac{U_{n2}}{2\sqrt{2\log n}},
\end{aligned}$$

where $U_{n2}\xrightarrow{d}\xi$ with distribution function $F_\xi(x)$ in (3.1). Reviewing (5.6), we have

$$\begin{aligned}
M_{n2}^2&=\max_{1\le i<j\le n,i=j-1}|\boldsymbol{X}_i-\boldsymbol{X}_j|^2=\max_{1\le i<j\le n,i=j-1}\sum_{k=1}^{p}\eta_{ijk}\\
&=\sqrt{p}\sigma_{n3}(\sqrt{2\log n}-\frac{\log\log n}{2\sqrt{2\log n}}+\frac{\log 8}{2\sqrt{2\log n}}+\frac{U_{n2}}{2\sqrt{2\log n}})+2p(1-\rho_n).
\end{aligned}$$

It follows that

$$2\sqrt{2\log n}\Big(\frac{M_{n2}^2-2p(1-\rho_n)}{\sqrt{p(2\rho_n^2-12\rho_n+8)}}-\sqrt{2\log p}+\frac{\log\log p}{2\sqrt{2\log p}}-\frac{\log 8}{2\sqrt{2\log p}}\Big)\stackrel{d}{\to}\xi,$$

where the distribution function of ξ is given in (3.1). □

5.3. Proof of Theorem 3.3

Now we are in a position to prove Theorem 3.3.

Proof of Theorem 3.3 Set $I=\{(i,j);1\le i<j\le n\}$. For $\alpha=(i,j)\in I$, define $Z_\alpha=\frac{1}{\sqrt{p}}\sum_{k=1}^p\eta_{ijk}$ and

$$B_\alpha=\{(k,l)\in I;\ \{k,l\}\cap\{i-1,i,i+1,j-1,j,j+1\}\neq\emptyset \text{ but } (k,l)\neq\alpha\}.$$

Then we find that Z_α is independent of $\{Z_\beta;\beta\notin B_\alpha\}$. Review the definitions of I', I'', B'_α and B''_α. By Lemma 2.1, we have $|P(\max_{\alpha\in I}Z_\alpha\le\tau)-e^{-\lambda_{p3}}|\le w_1+w_2$, where

$$\begin{aligned}
\lambda_{p3}=&\sum_{\alpha\in I}P(Z_\alpha>\tau)=\sum_{\alpha\in I'}P(Z_\alpha>\tau)+\sum_{\alpha\in I''}P(Z_\alpha>\tau)\\
=&(n-1)P\Big(\frac{1}{\sqrt{p}}\sum_{k=1}^p\eta_{12k}>\tau\Big)+\frac{(n-1)(n-2)}{2}P\Big(\frac{1}{\sqrt{p}}\sum_{k=1}^p\eta_{13k}>\tau\Big),\\
w_1=&\sum_{\alpha\in I}\sum_{\beta\in B_\alpha}P(Z_\alpha>\tau)P(Z_\beta>\tau)\\
=&\sum_{\alpha\in I'}\sum_{\beta\in B'_\alpha}P(Z_\alpha>\tau)P(Z_\beta>\tau)+\sum_{\alpha\in I''}\sum_{\beta\in B''_\alpha}P(Z_\alpha>\tau)P(Z_\beta>\tau)\\
\le&(n-1)\cdot 5\cdot P\Big(\frac{1}{\sqrt{p}}\sum_{k=1}^p\eta_{12k}>\tau\Big)^2+\frac{(n-1)(n-2)}{2}\cdot(6n)\cdot P\Big(\frac{1}{\sqrt{p}}\sum_{k=1}^p\eta_{13k}>\tau\Big)^2,\\
w_2=&\sum_{\alpha\in I}\sum_{\alpha\neq\beta\in B_\alpha}P(Z_\alpha>\tau,Z_\beta>\tau)\\
=&\sum_{\alpha\in I'}\sum_{\beta\in B'_\alpha}P(Z_\alpha>\tau,Z_\beta>\tau)+\sum_{\alpha\in I''}\sum_{\beta\in B''_\alpha}P(Z_\alpha>\tau,Z_\beta>\tau)\\
\le&(n-1)\cdot 5\cdot P\Big(\frac{1}{\sqrt{p}}\sum_{k=1}^p\eta_{12k}>\tau,\frac{1}{\sqrt{p}}\sum_{k=1}^p\eta_{23k}>\tau\Big)+\\
&\frac{(n-1)(n-2)}{2}\cdot(6n)\cdot P\Big(\frac{1}{\sqrt{p}}\sum_{k=1}^p\eta_{13k}>\tau,\frac{1}{\sqrt{p}}\sum_{k=1}^p\eta_{15k}>\tau\Big).
\end{aligned}$$

Case 1. $\lambda\in[-\infty,2-2\sqrt{2})$. Set $\tau=\sigma_{n3}s'_n+2\sqrt{p}(1-\rho_n)$. We first estimate λ_{p3}. When $\lim_{n\to\infty}\frac{\rho_n\sqrt{p}}{\sqrt{\log n}}=\lambda\in(-\infty,2-2\sqrt{2})$, we know that $\rho_n<0$, $\frac{\sigma_{n3}}{\sigma_{n1}}\ge 1$ and $\frac{1}{\sigma_{n1}}\ge\frac{1}{2\sqrt{2}}$. Review (5.8), (5.10) and the proof Theorem 3.1. By the same argument as in (5.11), we arrive at

$$n^2\cdot P\Big(\frac{1}{\sqrt{p}}\sum_{k=1}^p\eta_{13k}>\sigma_{n3}s'_n+2\sqrt{p}(1-\rho_n)\Big)$$

$$
\begin{aligned}
&= n^2 \cdot P\Big(\frac{1}{\sqrt{p}\sigma_{n1}}\sum_{k=1}^{p}(\eta_{13k}-2) > \frac{\sigma_{n3}s'_n}{\sigma_{n1}} - \frac{2\rho_n\sqrt{p}}{\sigma_{n1}}\Big)\\
&\le n^2 \cdot P\Big(\sum_{k=1}^{p}\eta_k > s'_n - \frac{\rho_n\sqrt{p}}{\sqrt{2}}\Big) = \frac{n^2}{\sqrt{2\pi}(s'_n - \frac{\rho_n\sqrt{p}}{\sqrt{2}})} e^{-(s'_n - \frac{\rho_n\sqrt{p}}{\sqrt{2}})^2/2}\\
&= \frac{n^2}{\sqrt{2\pi}(s'_n - \frac{\rho_n\sqrt{p}}{\sqrt{2}})}\exp\Big(-\frac{s'^2_n}{2} - \frac{\rho_n^2 p}{4} + \frac{s'_n\rho_n\sqrt{p}}{\sqrt{2}}\Big) \to 0
\end{aligned}
$$

as $n \to \infty$. When $\lim_{n\to\infty}\frac{\rho_n\sqrt{p}}{\sqrt{\log n}} = -\infty$, we have $\rho_n < 0$. By the same argument as in (5.11), we obtain that

$$
\begin{aligned}
&n^2 \cdot P\Big(\frac{1}{\sqrt{p}}\sum_{k=1}^{p}\eta_{13k} > \sigma_{n3}s'_n + 2\sqrt{p}(1-\rho_n)\Big)\\
&\quad = n^2 \cdot P\Big(\frac{1}{\sqrt{p}\sigma_{n1}}\sum_{k=1}^{p}(\eta_{13k}-2) > \frac{\sigma_{n3}s'_n}{\sigma_{n1}} - \frac{2\rho_n\sqrt{p}}{\sigma_{n1}}\Big)\\
&\quad \le n^2 \cdot P\Big(\sum_{k=1}^{p}\eta_k > 2s'_n\Big) = \frac{n^2}{2\sqrt{2\pi}s'_n}e^{-2s'^2_n} \to 0
\end{aligned}
$$

as $n \to \infty$. Thus

$$
\lim_{n\to\infty} n^2 \cdot P\Big(\frac{1}{\sqrt{p}}\sum_{k=1}^{p}\eta_{13k} > \sigma_{n3}s'_n + 2\sqrt{p}(1-\rho_n)\Big) = 0. \tag{5.25}
$$

Reviewing (5.20), we have $\lim_{n\to\infty}\lambda_{p3} = \frac{1}{4\sqrt{2\pi}}e^{-\frac{x}{2}}$ for any $x \in \mathbb{R}$.

Then, we can see from (5.25) that

$$
P\Big(\frac{1}{\sqrt{p}}\sum_{k=1}^{p}\eta_{13k} > \sigma_{n3}s'_n + 2\sqrt{p}(1-\rho_n)\Big)^2 = o\big(\frac{1}{n^4}\big) \tag{5.26}
$$

as $n \to \infty$. Recalling (5.21), we have $\lim_{n\to\infty} w_1 = 0$.

Finally, for w_2, it is obvious that

$$
\begin{aligned}
&P\Big(\frac{1}{\sqrt{p}}\sum_{k=1}^{p}\eta_{13k} > \sigma_{n3}s'_n + 2\sqrt{p}(1-\rho_n), \frac{1}{\sqrt{p}}\sum_{k=1}^{p}\eta_{15k} > \sigma_{n3}s'_n + 2\sqrt{p}(1-\rho_n)\Big)\\
&\quad \le P\Big(\frac{1}{\sqrt{p}\sigma_{n2}}\sum_{k=1}^{p}(\eta_{13k}+\eta_{15k}-4) > \frac{2\sigma_{n3}s'_n}{\sigma_{n2}} - \frac{4\rho_n\sqrt{p}}{\sigma_{n2}}\Big).
\end{aligned}
$$

When $\lim_{n\to\infty}\frac{\rho_n\sqrt{p}}{\sqrt{\log n}} = \lambda \in [-\infty, 2-2\sqrt{2})$, we have $\frac{\sigma_{n3}}{\sigma_{n2}} \ge \sqrt{\frac{2}{5}}$, $\frac{1}{\sigma_{n2}} \ge \frac{1}{2\sqrt{5}}$ and $\rho_n\sqrt{p} < (2-2\sqrt{2})\sqrt{\log n} + o(\sqrt{\log n}) < (\sqrt{2}-2)s'_n + o(\sqrt{\log n})$. By (5.13), (5.14) and Lemma 2.2, we obtain

$$
\begin{aligned}
&n^3 \cdot P\Big(\frac{1}{\sqrt{p}\sigma_{n2}}\sum_{k=1}^{p}(\eta_{13k}+\eta_{15k}-4) > \frac{2\sigma_{n3}s'_n}{\sigma_{n2}} - \frac{4\rho_n\sqrt{p}}{\sigma_{n2}}\Big)\\
&\quad \le n^3 \cdot P\Big(\frac{1}{\sqrt{p}\sigma_{n2}}\sum_{k=1}^{p}(\eta_{13k}+\eta_{15k}-4) > \frac{4\sqrt{5}}{5}s'_n + o(\sqrt{\log n})\Big)
\end{aligned}
$$

$$= \frac{n^3}{\sqrt{2\pi}(\frac{4\sqrt{5}}{5}s_n' + o(\sqrt{\log n}))} e^{-(\frac{4\sqrt{5}}{5}s_n' + o(\sqrt{\log n}))^2/2} \to 0$$

as $n \to \infty$. Therefore,

$$P\Big(\frac{1}{\sqrt{p}}\sum_{k=1}^{p}\eta_{13k} > \sigma_{n3}s_n' + 2\sqrt{p}(1-\rho_n), \frac{1}{\sqrt{p}}\sum_{k=1}^{p}\eta_{15k} > \sigma_{n3}s_n' + 2\sqrt{p}(1-\rho_n)\Big) = o(\frac{1}{n^3})$$

as $n \to \infty$. Reviewing (5.24), we have $\lim_{n\to\infty} w_2 = 0$. Thus, when $\lambda \in [-\infty, 2-2\sqrt{2})$,

$$\begin{aligned}
&\lim_{n\to\infty} P\Big(\max_{\alpha\in I} Z_\alpha \le \sigma_{n3}s_n' + 2\sqrt{p}(1-\rho_n)\Big)\\
&= \lim_{n\to\infty} P\Big(\max_{1\le i<j\le p} \frac{1}{\sqrt{p}\sigma_{n3}}\sum_{k=1}^{p}(\eta_{ijk} - 2(1-\rho_n)) \le s_n'\Big)\\
&= \exp(-\frac{1}{4\sqrt{2\pi}}e^{-\frac{x}{2}})
\end{aligned}$$

for any $x \in \mathbb{R}$. Then, by Lemma 2.4,

$$\max_{1\le i<j\le n} \frac{1}{\sqrt{p}\sigma_{n3}}\sum_{k=1}^{p}(\eta_{ijk} - 2(1-\rho_n)) = \sqrt{2\log n} - \frac{\log\log n}{2\sqrt{2\log n}} + \frac{\log 8}{2\sqrt{2\log n}} + \frac{U_{n3}}{2\sqrt{2\log n}},$$

where $U_{n3} \xrightarrow{d} \xi$ with distribution function $F_\xi(x)$ in (3.1). Reviewing (1.1) and (5.6), we have

$$\begin{aligned}
M_n^2 &= \max_{1\le i<j\le n} |\boldsymbol{X}_i - \boldsymbol{X}_j|^2 = \max_{1\le i<j\le n} \sum_{k=1}^{p}\eta_{ijk}\\
&= \sqrt{p}\sigma_{n3}(\sqrt{2\log n} - \frac{\log\log n}{2\sqrt{2\log n}} + \frac{\log 8}{2\sqrt{2\log n}} + \frac{U_{n3}}{2\sqrt{2\log n}}) + 2p(1-\rho_n).
\end{aligned}$$

It follows that

$$2\sqrt{2\log n}(\frac{M_n^2 - 2p(1-\rho_n)}{\sqrt{p(2\rho_n^2 - 12\rho_n + 8)}} - \sqrt{2\log p} + \frac{\log\log p}{2\sqrt{2\log p}} - \frac{\log 8}{2\sqrt{2\log p}}) \xrightarrow{d} \xi,$$

where the distribution function of ξ is given in (3.1).

Case 2. $\lambda \in [2-2\sqrt{2}, \infty]$. Set $\tau = \sigma_{n1}s_n + 2\sqrt{p}$. We first estimate λ_{p3}. When $\lim_{n\to\infty}\frac{\rho_n\sqrt{p}}{\sqrt{\log n}} = \lambda \in [2-2\sqrt{2}, \infty)$, we get that $\frac{\sigma_{n1}}{\sigma_{n3}} \ge 1$ and $\frac{1}{\sigma_{n3}} = \frac{\sqrt{2}}{4} + o(\frac{1}{\sqrt{\log n}})$. Reviewing (5.18), (5.19) and the proof of Theorem 3.2, by the same argument as in (5.20), we have

$$\begin{aligned}
&n \cdot P\Big(\frac{1}{\sqrt{p}}\sum_{k=1}^{p}\eta_{12k} > \sigma_{n1}s_n + 2\sqrt{p}\Big)\\
&= n \cdot P\Big(\frac{1}{\sqrt{p}\sigma_{n3}}\sum_{k=1}^{p}(\eta_{12k} - 2(1-\rho_n)) > \frac{\sigma_{n1}s_n}{\sigma_{n3}} + \frac{2\rho_n\sqrt{p}}{\sigma_{n3}}\Big)\\
&\le n \cdot P\Big(\sum_{k=1}^{p}\eta_k' > s_n + 2\rho_n\sqrt{p}[\frac{\sqrt{2}}{4} + o(\frac{1}{\sqrt{\log n}})]\Big)\\
&= n \cdot P\Big(\sum_{k=1}^{p}\eta_k' > s_n + \frac{\rho_n\sqrt{2p}}{2} + o(1)\Big)\\
&= \frac{n}{\sqrt{2\pi}(s_n + \frac{\rho_n\sqrt{2p}}{2} + o(1))} e^{-(s_n + \frac{\rho_n\sqrt{2p}}{2} + o(1))^2/2}
\end{aligned}$$

$$= \frac{n}{\sqrt{2\pi}(s_n + \frac{\rho_n\sqrt{2p}}{2} + o(1))} \exp(-\frac{s_n^2}{2} - \frac{\rho_n^2 p}{4} - \frac{s_n\rho_n\sqrt{2p}}{2} - o(s_n + \frac{\rho_n\sqrt{2p}}{2})) \to 0$$

as $n \to \infty$. When $\lim_{n\to\infty} \frac{\rho_n\sqrt{p}}{\sqrt{\log n}} = \infty$, we know that $\rho_n > 0$. By the same argument as in (5.20), we obtain that

$$\begin{aligned} & n \cdot P\Big(\frac{1}{\sqrt{p}} \sum_{k=1}^{p} \eta_{12k} > \sigma_{n1}s_n + 2\sqrt{p}\Big) \\ & = n \cdot P\Big(\frac{1}{\sqrt{p}\sigma_{n3}} \sum_{k=1}^{p} (\eta_{12k} - 2(1-\rho_n)) > \frac{\sigma_{n1}s_n}{\sigma_{n3}} + \frac{2\rho_n\sqrt{p}}{\sigma_{n3}}\Big) \\ & \le n \cdot P\Big(\sum_{k=1}^{p} \eta'_k > s_n\Big) = \frac{n}{\sqrt{2\pi}s_n} e^{-\frac{s_n^2}{2}} \to 0 \end{aligned}$$

as $n \to \infty$. Thus

$$\lim_{n\to\infty} n \cdot P\Big(\frac{1}{\sqrt{p}} \sum_{k=1}^{p} \eta_{12k} > \sigma_{n1}s_n + 2\sqrt{p}\Big) = 0. \tag{5.27}$$

Reviewing (5.11), we have $\lim_{n\to\infty} \lambda_{p3} = \frac{1}{4\sqrt{2\pi}} e^{-\frac{x}{2}}$ for any $x \in \mathbb{R}$.

Then, we can see from (5.27) that

$$P\Big(\frac{1}{\sqrt{p}} \sum_{k=1}^{p} \eta_{12k} > \sigma_{n1}s_n + 2\sqrt{p}\Big)^2 = o(\frac{1}{n^2}) \tag{5.28}$$

as $n \to \infty$. Recalling (5.12), we have $\lim_{n\to\infty} w_1 = 0$.

Finally, for w_2, it is easy to find that

$$\begin{aligned} & P\Big(\frac{1}{\sqrt{p}} \sum_{k=1}^{p} \eta_{12k} > \sigma_{n1}s_n + 2\sqrt{p}, \frac{1}{\sqrt{p}} \sum_{k=1}^{p} \eta_{23k} > \sigma_{n1}s_n + 2\sqrt{p}\Big) \\ & \le P\Big(\frac{1}{\sqrt{p}\sigma_{n4}} \sum_{k=1}^{p} (\eta_{12k} + \eta_{23k} - 4(1-\rho_n)) > \frac{2\sigma_{n1}s_n}{\sigma_{n4}} + \frac{4\rho_n\sqrt{p}}{\sigma_{n4}}\Big). \end{aligned}$$

When $\lim_{n\to\infty} \frac{\rho_n\sqrt{p}}{\sqrt{\log n}} = \lambda \in [2 - 2\sqrt{2}, \infty)$, we get that $\frac{\sigma_{n1}}{\sigma_{n4}} \ge \sqrt{\frac{2}{5}}$ and $\frac{1}{\sigma_{n4}} = \frac{1}{2\sqrt{5}} + o(\frac{1}{\sqrt{\log n}})$. By (5.22), (5.23), Lemma 2.2 and the formula $P(N(0,1) \ge x) = \frac{1}{\sqrt{2\pi}x} e^{-x^2/2}(1 + o(1))$ as $x \to \infty$, we obtain

$$\begin{aligned} & n \cdot P\Big(\frac{1}{\sqrt{p}\sigma_{n4}} \sum_{k=1}^{p} (\eta_{12k} + \eta_{23k} - 4(1-\rho_n)) > \frac{2\sigma_{n1}s_n}{\sigma_{n4}} + \frac{4\rho_n\sqrt{p}}{\sigma_{n4}}\Big) \\ & \le n \cdot P\Big(\frac{1}{\sqrt{p}\sigma_{n4}} \sum_{k=1}^{p} (\eta_{12k} + \eta_{23k} - 4(1-\rho_n)) > \frac{2}{\sqrt{5}}(\sqrt{2}s_n + \rho_n\sqrt{p} + o(1))\Big) \\ & = \frac{\sqrt{5}n}{2\sqrt{2\pi}(\sqrt{2}s_n + \rho_n\sqrt{p} + o(1))} e^{-2(\sqrt{2}s_n + \rho_n\sqrt{p} + o(1))^2/5} \\ & = \frac{\sqrt{5}n}{2\sqrt{2\pi}(\sqrt{2}s_n + \rho_n\sqrt{p} + o(1))} \exp(-\frac{2}{5}(2s_n^2 + \rho_n^2 p - 2s_n\rho_n\sqrt{2p}) + o(\sqrt{2}s_n + \rho_n\sqrt{p})) \to 0 \end{aligned}$$

as $n \to \infty$. When $\lim_{n\to\infty} \frac{\rho_n\sqrt{p}}{\sqrt{\log n}} = \infty$, we get that $\rho_n > 0$. By (5.22), (5.23), Lemma 2.2 and the formula $P(N(0,1) \geq x) = \frac{1}{\sqrt{2\pi}x} e^{-x^2/2}(1+o(1))$ as $x \to \infty$, we obtain that

$$
\begin{aligned}
& n \cdot P\Big(\frac{1}{\sqrt{p}\sigma_{n4}} \sum_{k=1}^{p} (\eta_{12k} + \eta_{23k} - 4(1-\rho_n)) > \frac{2\sigma_{n1} s_n}{\sigma_{n4}} + \frac{4\rho_n \sqrt{p}}{\sigma_{n4}}\Big) \\
& \leq n \cdot P\Big(\frac{1}{\sqrt{p}\sigma_{n4}} \sum_{k=1}^{p} (\eta_{12k} + \eta_{23k} - 4(1-\rho_n)) > \sqrt{2} s_n\Big) \\
& = \frac{n}{2\sqrt{\pi} s_n} e^{-s_n^2} \to 0
\end{aligned}
$$

as $n \to \infty$. In summary,

$$
P\Big(\frac{1}{\sqrt{p}\sigma_{n4}} \sum_{k=1}^{p} (\eta_{12k} + \eta_{23k} - 4(1-\rho_n)) > \frac{2\sigma_{n1} s_n}{\sigma_{n4}} + \frac{4\rho_n \sqrt{p}}{\sigma_{n4}}\Big) = o(\frac{1}{n})
$$

as $n \to \infty$. Reviewing (5.15), we have $\lim_{n\to\infty} w_2 = 0$. Thus, when $\lambda \in [2 - 2\sqrt{2}, \infty]$,

$$
\begin{aligned}
& \lim_{n\to\infty} P\Big(\max_{\alpha \in I} Z_\alpha \leq \sigma_{n1} s_n + 2\sqrt{p}\Big) = \lim_{n\to\infty} P\Big(\max_{1\leq i<j\leq n} \frac{1}{\sqrt{p}\sigma_{n1}} \sum_{k=1}^{p} (\eta_{ijk} - 2) \leq s_n\Big) \\
& = \exp(-\frac{1}{4\sqrt{2\pi}} e^{-\frac{x}{2}})
\end{aligned}
$$

for any $x \in \mathbb{R}$. Then, by Lemma 2.4,

$$
\max_{1\leq i<j\leq n} \frac{1}{\sqrt{p}\sigma_{n1}} \sum_{k=1}^{p} (\eta_{ijk} - 2) = 2\sqrt{\log n} - \frac{\log\log n}{4\sqrt{\log n}} + \frac{U_{n4}}{4\sqrt{\log n}},
$$

where $U_{n4} \xrightarrow{d} \xi$ with distribution function $F_\xi(x)$ in (3.1). Reviewing (1.1) and (5.6), we have

$$
\begin{aligned}
M_n^2 &= \max_{1\leq i<j\leq n} |\boldsymbol{X}_i - \boldsymbol{X}_j|^2 = \max_{1\leq i<j\leq n} \sum_{k=1}^{p} \eta_{ijk} \\
&= \sqrt{p}\sigma_{n1}(2\sqrt{\log n} - \frac{\log\log n}{4\sqrt{\log n}} + \frac{U_{n4}}{4\sqrt{\log n}}) + 2p.
\end{aligned}
$$

It follows that

$$
4\sqrt{\log n}(\frac{M_n^2 - 2p}{\sqrt{p(8 - 6\rho_n^2)}} - 2\sqrt{\log n} + \frac{\log\log n}{4\sqrt{\log n}}) \xrightarrow{d} \xi,
$$

where the distribution function of ξ is given in (3.1).

Combining Case 1 with Case 2, the proof is completed. □

Acknowledgements We thank the referees for their time and comments.

Conflict of Interest The authors declare no conflict of interest.

References

[1] P. C. MATTHEWS, A. L. RUKHIN. *Asymptotic distribution of the normal sample range*. Ann. Appl. Probab., 1993, **3**(2): 454–466.

[2] H. A. DAVID. *Order Statistics, 2nd ed.* Wiley, New York, 1981.

[3] S. R. JAMMALAMADAKA, S. JANSON. *Asymptotic distribution of the maximum interpoint distance in a sample of random vectors with a spherically symmetric distribution*. Ann. Appl. Probab., 2015, **25**(6): 3571–3591.

[4] N. HENZE, T. KLEIN. *The limit distribution of the largest interpoint distance from a symmetric kotz sample*. J. Multivariate Anal., 1996, **57**(2): 228–239.

[5] N. HENZE, W. LAO. *The Limit Distribution of the Largest Interpoint Distance for Power-Tailed Spherically Decomposable Distributions and Their Affine Images*. In Press, Karlsruhe Institute of Technology, 2010.

[6] M. J. APPEL, C. A. NAJIM, R. P. RUSSO. *Limit laws for the diameter of a random point set*. Adv. Appl. Probab., 2002, **34**(1): 1–10.

[7] M. J. APPEL, R. P. RUSSO. *Limiting distributions for the maximum of a symmetric function on a random point set*. J. Theoret. Probab., 2006, **19**(2): 365–375.

[8] M. MAYER, I. MOLCHANOV. *Limit theorems for the diameter of a random sample in the unit ball*. Extremes, 2007, **10**(3): 129–150.

[9] W. LAO. *Some Weak Limit Laws for the Diameter of Random Point Sets in Bounded Regions*. KIT Scientific Publishing, Karlsruhe, 2010.

[10] M. SCHREMPP. *The limit distribution of the largest interpoint distance for distributions supported by a d-dimensional ellipsoid and generalizations*. Adv. Appl. Probab., 2015, **8**(4): 1256–1270.

[11] M. SCHREMPP. *Limit laws for the diameter of a set of random points from a distribution supported by a smoothly bounded set*. Extremes, 2019, **22**(1): 167–191.

[12] Y. DEMICHEL, A. K. FERMIN, P. SOULIER. *The diameter of an elliptical cloud*. Electron. J. Probab., 2015, **20**(27): 1–32.

[13] Ping TANG, Rongrong LU, Junshan XIE. *Asymptotic distribution of the maximum interpoint distance for high-dimensional data*. Statist. Probab. Lett., 2022, **190**: Paper No. 109567, 7 pp.

[14] J. HEINY, C. KLEEMANN. *Maximum interpoint distance of high-dimensional random vectors*. Bernoulli, 2025, **31**(1): 537–560.

[15] Jun LI. *Asymptotic distribution-free change-point detection based on interpoint distances for high-dimensional data*. J. Nonparametr. Stat., 2020, **32**(1): 157–184.

[16] R. MODARRES. *Multinomial interpoint distances*. Statist. Papers, 2018, **59**(1): 341–360.

[17] Yu SONG, R. MODARRES. *Interpoint distance test of homogeneity for multivariate mixture models*. Int. Stat. Rev., 2019, **87**(3): 613–638.

[18] R. MODARRES, Yu SONG. *Interpoint distances: applications, properties, and visualization*. Appl. Stoch. Models Bus. Ind., 2020, **36**(6): 1147–1168.

[19] Guowei YAN, Long FENG. *Limit law for the maximum interpoint distance of high dimensional dependent variables*. arXiv: 2312.10875, 2022.

[20] Weidong LIU, Zhengyan LIN. *Asymptotic distributions of the largest entries of sample correlation matrices under dependence assumptions*. Chinese Ann. Math. Ser. A, 2008, **29**(4): 543–556. (in Chinese)

[21] R. A. HORN, C. R. JOHNSON. *Matrix Analysis*. Cambridge University Press, 2012.

[22] L. H. Y. CHEN, Xiao FANG, Qiman SHAO. *From Stein identities to moderate deviations*. Ann. Probab., 2013, **41**(1): 262–293.

[23] R. ARRATIA, L. GOLDSTEIN, L. GORDON. *Two moments suffice for Poisson approximations: the Chen-Stein method*. Ann. Probab., 1989, **17**(1): 9–25.

[24] T. CAI, Tiefeng JIANG. *Limiting laws of coherence of random matrices with applications to testing covariance structure and construction of compressed sensing matrices*. Ann. Statist., 2011, **39**(3): 1496–1525.

[25] Jianqing FAN, Tiefeng JIANG. *Largest entries of sample correlation matrices from equi-correlated normal populations*. Ann. Probab., 2019, **47**(5): 3321–3374.

[26] Tiefeng JIANG. *The asymptotic distributions of the largest entries of sample correlation matrices*. Ann. Appl. Probab., 2004, **14**(2): 865–880.

[27] Deli LI, Yongcheng QI, A. ROSALSKY. *On Jiang's asymptotic distribution of the largest entry of a sample correlation matrix*. J. Multivariate Anal., 2012, **111**: 256–270.

[28] Haozhu ZHAO, Yong ZHANG. *The asymptotic distributions of the largest entries of sample correlation matrices under an α-mixing assumption*. Acta Math. Sin. (Engl. Ser.), 2022, **38**(11): 2039–2056.

[29] Jingcai YANG, Dawei LU. *A general solution of the type I McKay distribution's Stein equation*. J. Math. Res. Appl., 2023, **43**(5): 607–618.

[30] Wang ZHOU. *Asymptotic distribution of the largest off-diagonal entry of correlation matrices*. Trans. Amer. Math. Soc., 2007, **359**(11): 5345–5363.