TEST
https://doi.org/10.1007/s11749-017-0575-x

ORIGINAL PAPER

A central limit theorem for sums of functions of residuals in a high-dimensional regression model with an application to variance homoscedasticity test

Zhidong Bai[1] · **Guangming Pan**[2] · **Yanqing Yin**[3]

Received: 2 May 2017 / Accepted: 8 December 2017

Abstract We establish a joint central limit theorem for sums of squares and the fourth powers of residuals in a high-dimensional regression model. We then apply this CLT to detect the existence of heteroscedasticity for linear regression models without assuming randomness of covariates when the sample size n tends to infinity and the number of covariates p may be fixed or tend to infinity.

Keywords CLT · Dependent random variables · Breusch and Pagan test · White's test · Heteroscedasticity · Homoscedasticity · High-dimensional regression · Design matrix

---

Zhidong Bai is partially supported by a Grant NSF China 11571067 and 11471140. Guangming Pan was partially supported by a MOE Tier 2 Grant 2014-T2-2-060 and by a MOE Tier 1 Grant RG25/14 at the Nanyang Technological University, Singapore. Yanqing Yin was partially supported by a Grant NSF China 11701234, the Priority Academic Program Development of Jiangsu Higher Education Institutions and a project of China Scholarship Council.

✉ Yanqing Yin
yinyq@jsnu.edu.cn; yinyq799@nenu.edu.cn

Zhidong Bai
baizd@nenu.edu.cn

Guangming Pan
gmpan@ntu.edu.sg

1 KLASMOE and School of Mathematics and Statistics, Northeast Normal University, Changchun 130024, People's Republic of China

2 School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore

3 School of Mathematics and Statistics, Jiangsu Normal University, Xuzhou 221116, People's Republic of China

Published online: 23 December 2017

Mathematics Subject Classification Primary 62J05 · 62H15; Secondary 60F05

1 Introduction

Suppose $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)'$, where $\{\varepsilon_i\}_{i=1}^n$ is a sample of size n drawn from a certain standardized population. Let $\mathbf{A} = (a_{ij})_{n\times n}$ be a matrix. Define $\hat{\boldsymbol{\varepsilon}} = \left(\hat{\varepsilon}_1, \hat{\varepsilon}_2, \ldots, \hat{\varepsilon}_n\right)' = \mathbf{A}\boldsymbol{\varepsilon}$, a linear transformation of a random vector $\boldsymbol{\varepsilon}$ of i.i.d. entries. Notice that in this paper, the random variables ε_i's as well as the elements a_{ij}'s of the matrix $\mathbf{A}$ may depend on n, but this dependency is suppressed from the notations for brevity when there is no confusion. Consider a statistic

$$\mathfrak{A}_2(\boldsymbol{\varepsilon}) = \sum_{i=1}^n \hat{\varepsilon}_i^2 = \boldsymbol{\varepsilon}' \left(\mathbf{A}'\mathbf{A}\right) \boldsymbol{\varepsilon}.$$

Its exact or asymptotic distributions have been considered by many authors especially when the population distribution is Gaussian. It arises in many statistical applications, such as the power analysis of a test procedure when the test statistic (e.g., Pearson's Chi-square statistic) takes the form of $\mathfrak{A}_2(\boldsymbol{\varepsilon})$. We refer to Liu et al. (2009), Jensen and Solomon (1972), de Jong (1987), Gotze and Tikhomirov (1999), Whittle (1964), Nourdin et al. (2010), Deya and Nourdin (2014), Nourdin et al. (2016) and references therein for this problem and related problems.

Denote the spectral decomposition of $\mathbf{A}'\mathbf{A}$ by $\mathbf{P}'\boldsymbol{\Lambda}\mathbf{P}$, where $\mathbf{P}$ is an orthogonal matrix and $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ being the eigenvalues of $\mathbf{A}'\mathbf{A}$, here and in the following $\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ stands for a diagonal matrix with its diagonal entries being $\lambda_1, \lambda_2, \ldots, \lambda_n$. When the population distribution is standard Gaussian $\mathfrak{A}_2(\boldsymbol{\varepsilon})$ follows a generalized Chi-square distribution. It can be expressed as a weighted sum of Chi-square variables

$$\mathfrak{A}_2(\boldsymbol{\varepsilon}) = \sum_{i=1}^n \lambda_i \chi_i^2(1),$$

where $\{\chi_i(1)\}$ are n independent χ^2 random variables with degrees of freedom 1. In the situation where $\lambda_i = \lambda > 0$ for all $i = 1, \ldots, n$, $\mathfrak{A}_2(\boldsymbol{\varepsilon})$ has a $\chi^2(n)$ distribution after scaling by λ. If not all the λ_i's are equal, divide nonzero λ_i's into M sets according to their repetition and define $\mathbf{r} = (r_1, r_2, \ldots, r_M)$ to be the frequencies of the distinct values of eigenvalues λ_i. That is, assume

$$\begin{aligned}\lambda_1 = \cdots = \lambda_{r_1} &\triangleq \pi_1 > \lambda_{r_1+1} = \cdots = \lambda_{r_1+r_2} \triangleq \pi_2\\ &> \cdots > \lambda_{\sum_{j=1}^{M-1} r_j+1} = \cdots = \lambda_{\sum_{j=1}^{M} r_j} \triangleq \pi_M > 0.\end{aligned} \tag{1.1}$$

Then the pdf of $\mathfrak{A}_2(\boldsymbol{\varepsilon})$ was derived in Amari and Misra (1997) as follows:

$$f(x; \mathbf{r}, \pi_1, \ldots, \pi_M) = \prod_{m=1}^M \frac{1}{\pi_m^{r_m}} \sum_{k=1}^M \sum_{l=1}^{r_k} \frac{\Psi_{k,k,\mathbf{r}}}{(r_k - l)!} (-x)^{r_k - l} e^{-\frac{x}{\pi_k}}, \quad \text{for} \quad x \geq 0,$$

where

$$\Psi_{k,l,\mathbf{r}} = (-1)^{r_k-1} \sum_{\mathbf{i}\in\Omega_{k,l}} \prod_{j\neq k} \binom{i_j + r_j - 1}{i_j} \left(\frac{1}{\pi_j} - \frac{1}{\pi_k}\right)^{-(r_i+r_j)},$$

with $\mathbf{i} = (i_1, \ldots, i_M)'$ from the set $\Omega_{k,l}$ of all partitions of $l-1$ defined as

$$\Omega_{k,l} = \left\{ (i_1, \ldots, i_M) \in \mathbb{Z}^M; \sum_{j=1}^{M} i_j = l-1, i_k = 0, i_j \geq 0 \quad \text{for all } j \right\}.$$

However, in most cases, one may be interested in the situation where the populations are not Gaussian, even without explicit pdfs. Moreover, some statistics may take the form of

$$\mathfrak{A}_4(\boldsymbol{\varepsilon}) = \sum_{i=1}^{n} \hat{\varepsilon}_i^4, \quad \mathfrak{A}_6(\boldsymbol{\varepsilon}) = \sum_{i=1}^{n} \hat{\varepsilon}_i^6, \ldots,$$

or their functions. Note that unlike $\mathfrak{A}_2(\boldsymbol{\varepsilon})$, $\mathfrak{A}_4(\boldsymbol{\varepsilon})$ cannot be expressed as a quadratic form of independent random variables. Finding the asymptotic distribution of $\mathfrak{A}_4(\boldsymbol{\varepsilon})$ is far more than trivial. We establish the central limit theorem of the pair $(\mathfrak{A}_2(\boldsymbol{\varepsilon}), \mathfrak{A}_4(\boldsymbol{\varepsilon}))$ and explore its application in testing homoscedasticity in fixed design linear regressions.

2 A joint CLT for $\mathfrak{A}_2(\boldsymbol{\varepsilon})$ and $\mathfrak{A}_4(\boldsymbol{\varepsilon})$

To present the main results, we first introduce the following notation. Use Diag $(\mathbf{B}) = \left(b_{1,1}, b_{2,2}, \ldots, b_{n,n}\right)'$ to stand for the vector formed by the diagonal entries of $\mathbf{B}$ and Diag$'$ $(\mathbf{B})$ as its transpose, use $\mathbf{D_B}$ to denote the diagonal matrix of $\mathbf{B}$ (replacing all off-diagonal entries with zero), and use $\mathbf{1}$ to stand for the n-vector $(1, 1, \ldots, 1)'$. We also use $\mathbf{A} \circ \mathbf{B} = (a_{ij} b_{ij})$ to denote the Hadamard product of two matrices $\mathbf{A} = (a_{ij})$ and $\mathbf{B} = (b_{ij})$ and use $\mathbf{A}^{\circ k}$ to denote the Hadamard product of k matrices $\mathbf{A}$. We are interested in a bivariate CLT for the pair $(T_1, T_2) = (\mathfrak{A}_2(\boldsymbol{\varepsilon}), \mathfrak{A}_4(\boldsymbol{\varepsilon}))$ when $n \to \infty$.

We now make the following assumptions.

(1) $\{\varepsilon_i\}_{1\leq i\leq n}$ are standard i.i.d. random variables (mean zero and variance one) with $E\varepsilon_1^8 < \infty$.
(2) The distribution of ε_1 is *symmetric*.
(3) The spectral norm (i.e., the largest singular value) of the $n \times n$ square matrix $\mathbf{A}$ is bounded.

Denote by $(\nu_{2k})_{k\geq 1}$ the cumulants (whenever they exist) of ε_1. In particular, $\nu_4 = M_4 - 3 \geq -2$, $\nu_6 = M_6 - 15M_4 + 30$ and $\nu_8 = M_8 - 28M_6 - 35M_4^2 + 420M_4 - 630$ where $M_j = \mathrm{E}\varepsilon_1^j$.

First of all, it is not difficult to calculate the means of T_1 and T_2 as follows:

$$\begin{aligned} ET_1 &= \sum_{1\le i,j\le n} a_{ij}^2 = \operatorname{tr}\left(\mathbf{A}\mathbf{A}'\right), \\ ET_2 &= 3\sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}^2\right)^2 + \nu_4 \sum_{1\le i,j\le n} a_{ij}^4 \\ &= 3\operatorname{tr}\left(\mathbf{A}\mathbf{A}'\right)\circ\left(\mathbf{A}\mathbf{A}'\right) + \nu_4 \operatorname{tr}\left(\mathbf{A}^{\circ 2}\mathbf{A}'^{\circ 2}\right). \end{aligned}$$

We are now in a position to state the main result.

Theorem 2.1 *In addition to assumptions 1–3, suppose that $\nu_4 > -2$ and*

$$\liminf_{n\to\infty} \frac{\operatorname{tr}\left(\mathbf{A}'\mathbf{A}\right)^2}{n} > 0. \tag{2.1}$$

Then the centered pair $(T_1 - \mathrm{E}T_1, T_2 - \mathrm{E}T_2)$ is asymptotically normal with variance–covariance matrix $\boldsymbol{\Sigma} = (\boldsymbol{\Sigma}_{ij})$ where

$$\begin{aligned} \Sigma_{11} &= 2\sum_{i_1 i_2 j_1 j_2} a_{i_1 j_1} a_{i_1 j_2} a_{i_2 j_1} a_{i_2 j_2} + \nu_4 \sum_{i_1 i_2 j} a_{i_1 j}^2 a_{i_2 j}^2 \\ &= 2\operatorname{tr}\left(\left(\mathbf{A}\mathbf{A}'\right)^2\right) + \nu_4 \operatorname{tr}\left(\mathbf{A}'\mathbf{A}\right)\circ\left(\mathbf{A}'\mathbf{A}\right), \qquad (2.2) \\ \Sigma_{22} &= 72 \sum_{i_1, i_2, j_1, \ldots, j_4} a_{i_1 j_1}^2 a_{i_1 j_2} a_{i_1 j_3} a_{i_2 j_2} a_{i_2 j_3} a_{i_2 j_4}^2 \\ &\quad + 24 \sum_{i_1, i_2, j_1, \ldots, j_4} a_{i_1 j_1} a_{i_1 j_2} a_{i_1 j_3} a_{i_1 j_4} a_{i_2 j_1} a_{i_2 j_2} a_{i_2 j_3} a_{i_2 j_4} \\ &\quad + \nu_4 \left(96 \sum_{i_1 i_2, j_1 j_2 j_3} a_{i_1 j_1}^3 a_{i_1 j_2} a_{i_2 j_1} a_{i_2 j_2} a_{i_2 j_3}^2 \right. \\ &\quad + 72 \sum_{i_1 i_2, j_1 j_2 j_3} a_{i_1 j_1}^2 a_{i_1 j_2} a_{i_1 j_3} a_{i_2 j_1}^2 a_{i_2 j_2} a_{i_2 j_3} \\ &\quad \left. + 36 \sum_{i_1 i_2, j_1 j_2 j_3} a_{i_1 j_1}^2 a_{i_1 j_2}^2 a_{i_2 j_1}^2 a_{i_2 j_3}^2 \right) \\ &\quad + \nu_4^2 \left(12 \sum_{i_1 i_2 j_1 j_2} a_{i_1 j_1}^2 a_{i_1 j_2}^2 a_{i_2 j_1}^2 a_{i_2 j_2}^2 + 16 \sum_{i_1 i_2 j_1 j_2} a_{i_1 j_1}^3 a_{i_1 j_2} a_{i_2 j_1} a_{i_2 j_2}^3 \right) \\ &\quad + \nu_6 \left(12 \sum_{i_1 i_2 j_1 j_2} a_{i_1 j_1}^4 a_{i_2 j_1}^2 a_{i_2 j_2}^2 + 16 \sum_{i_1 i_2 j_1 j_2} a_{i_1 j_1}^3 a_{i_1 j_2} a_{i_2 j_1} a_{i_2 j_2}^3 \right) \end{aligned}$$

$$
\begin{aligned}
&+ \nu_8 \sum_{i_1 i_2 j} a_{i_1 j}^4 a_{i_2 j}^4 \\
&= 72\text{Diag}'\left(\mathbf{AA}'\right)\left(\left(\mathbf{AA}'\right)\circ\left(\mathbf{AA}'\right)\right)\text{Diag}\left(\mathbf{AA}'\right) + 24\text{tr}\left(\left(\mathbf{AA}'\right)\circ\left(\mathbf{AA}'\right)\right)^2 \\
&+ \nu_4\left(96\text{tr}\left(\mathbf{AA}'\right)\mathbf{D}_{(\mathbf{AA}')}\mathbf{A}\left(\mathbf{A}'\right)^{\circ 3}\right. \\
&+ 72\text{tr}\left(\left(\left(\mathbf{AA}'\right)\circ\left(\mathbf{AA}'\right)\right)\left(\mathbf{A}\circ\mathbf{A}\right)\left(\mathbf{A}'\circ\mathbf{A}'\right)\right) \\
&\left.+ 36\text{Diag}'\left(\mathbf{AA}'\right)\left(\mathbf{A}^{\circ 2}\left(\mathbf{A}'\right)^{\circ 2}\right)\text{Diag}\left(\mathbf{AA}'\right)\right) \\
&+ \nu_4^2\left(18\text{tr}\left(\mathbf{A}^{\circ 2}\left(\mathbf{A}'\right)^{\circ 2}\right)^2 + 16\text{tr}\left(\mathbf{A}^{\circ 3}\mathbf{A}'\mathbf{A}\left(\mathbf{A}'\right)^{\circ 3}\right)\right) \\
&+ \nu_6\left(12\text{tr}\left(\left(\mathbf{A}'\mathbf{D}_{\mathbf{AA}'}\mathbf{A}\right)\circ\left(\mathbf{A}'^{\circ 2}\mathbf{A}^{\circ 2}\right)\right) + 16\text{tr}\left(\mathbf{AA}'\right)\mathbf{A}^{\circ 3}\left(\mathbf{A}'\right)^{\circ 3}\right) \\
&+ \nu_8\mathbf{1}'\left(\mathbf{A}^{\circ 4}\left(\mathbf{A}'\right)^{\circ 4}\right)\mathbf{1},
\end{aligned} \tag{2.3}
$$

$$
\begin{aligned}
\Sigma_{12} = \Sigma_{21} = &\left(12\sum_{i_1 i_2 j_1 j_2 j_3} a_{i_1 j_1}^2 a_{i_1 j_2} a_{i_1 j_3} a_{i_2 j_2} a_{i_2 j_3} + 8\nu_4 \sum_{i_1 i_2 j_1 j_2} a_{i_1 j_1}^3 a_{i_1 j_2} a_{i_2 j_1} a_{i_2 j_2}\right. \\
&\left.+ 6\nu_4 \sum_{i_1 i_2 j_1 j_2} a_{i_1 j_1}^2 a_{i_1 j_2}^2 a_{i_1 j_1}^2 + \nu_6 \sum_{i_1 i_2 j} a_{i_1 j}^4 a_{i_2 j}^2\right) \\
= &\left(12\text{tr}\left(\left(\mathbf{AA}'\right)^2\circ\left(\mathbf{AA}'\right)\right) + 8\nu_4\text{tr}\left(\left(\mathbf{AA}'\mathbf{A}\right)\mathbf{A}'^{\circ 3}\right)\right. \\
&\left.+ 6\nu_4\text{tr}\left(\left(\mathbf{A}'\mathbf{D}_{\mathbf{AA}'}\mathbf{A}\right)\circ\left(\mathbf{A}'\mathbf{A}\right)\right) + \nu_6\left[\text{Diag}'\left(\mathbf{A}'\mathbf{A}\right)\left(\mathbf{A}'^{\circ 4}\right)\mathbf{1}\right]\right).
\end{aligned} \tag{2.4}
$$

When $\nu_4 = -2$, that is, the ε_i's are drawn from a two-point distribution with the same masses $1/2$ at -1 and 1, the above CLT is still valid if the following assumption holds:

$$
\limsup_{n\to\infty} \frac{\text{tr}\left(\mathbf{A}'\mathbf{A}\right)\circ\left(\mathbf{A}'\mathbf{A}\right)}{\text{tr}\left(\mathbf{A}'\mathbf{A}\right)^2} < 1. \tag{2.5}
$$

2.1 Calculation of the variances and covariance of T_1 and T_2

To calculate the mean and the variance Σ_{11} of T_1, draw two parallel lines, called the I-line and the J-line, and then for given i, j_1, j_2, draw a simple graph $G_1(i, j_1 j_2)$ from i on the I-line to j_1 and j_2 on the J-line, respectively. The edges (i, j_1) and (i, j_2) correspond to two entries a_{i,j_1} and a_{i,j_2}, and the two J-vertices j_1 and j_2 correspond to random variables ε_{j_1} and ε_{j_2}, respectively. We say the graph $G_1(i, j_1, j_2)$ corresponds to the term $A_{G_1(i,j_1,j_2)}\varepsilon_{G_1(i,j_1,j_2)}$, where $A_{G_1(i,j_1,j_2)} = \prod_{t=1}^2 a_{ij_t}$ and $\varepsilon_{G_1(i,j_1,j_2)} = \prod_{t=1}^2 \varepsilon_{j_t}$. One can easily show that

$$
ET_1 = \sum_{i,j_1,j_2=1}^n EG_1(i, j_1, j_2) = \sum_{i,j_1,j_2=1}^n EA_{G_1(i,j_1,j_2)}\varepsilon_{G_1(i,j_1,j_2)} = \sum_{i,j=1}^n a_{ij}^2 = \text{tr}\mathbf{AA}',
$$

and

$$\Sigma_{11} = \sum_{i_1 i_2, j_1, j_2, j_3, j_4=1}^{n} E(G_1(i_1, j_1, j_2) - EG_1(i_1, j_1, j_2))(G_1(i_2, j_3, j_4) \\ -EG_1(i_2, j_3, j_4)).$$

Note that the term is zero if the graph $G_1(i_1, j_1, j_2)$ does not have J-vertices coincident with that of $G_1(i_2, j_3, j_4)$ or there is a single J-vertex of the combined graph $G = G_1(i_1, j_1, j_2) \cup G_1(i_2, j_3, j_4)$. Therefore,

$$\begin{aligned}\Sigma_{11} &= 2\sum_{i_1 i_2 j_1 j_2} a_{i_1 j_1} a_{i_1 j_2} a_{i_2 j_1} a_{i_2 j_2} + \nu_4 \sum_{i_1 i_2 j} a_{i_1 j}^2 a_{i_2 j}^2 \\ &= 2\mathrm{tr}\left(\left(\mathbf{AA}'\right)^2\right) + \nu_4 \mathrm{Diag}'\left(\mathbf{A}'\mathbf{A}\right) \mathrm{Diag}\left(\mathbf{A}'\mathbf{A}\right),\end{aligned} \tag{2.6}$$

where the term after the second equality follows from the fact that there are two cases for j_1, j_2 marching j_3, j_4 as two pairs and the second term obtained for the case where the 4 J vertices are identical.

Next, we will calculate the mean and variance of T_2. Select a point i on the I-line and four points j_1, j_2, j_3, j_4 on the J line. Then draw a graph of four edges from i to j_1, j_2, j_3 and j_4, respectively. We call the graph $G_2(i, j_1, j_2, j_3, j_4)$ or simply $G_2(i, \mathbf{j})$ which corresponds to the term $\prod_{t=1}^{4} a_{ij_t} \varepsilon_{j_t} := A_{G_2(i,\mathbf{j})} \varepsilon_{G_2(i,\mathbf{j})}$, where $A_{G_2(i,\mathbf{j})} = \prod_{t=1}^{4} a_{ij_t}$ and $\varepsilon_{G_2(i,\mathbf{j})} = \prod_{t=1}^{4} \varepsilon_{j_t}$. We then write

$$T_2 = \sum_{i=1}^{n} \sum_{\mathbf{j}} G_2(i, \mathbf{j}).$$

Note that $E\varepsilon_{G_2(i,\mathbf{j})} = 0$ if $G_2(i, \mathbf{j})$ contains a J-vertex whose degree is one. Thus, there are two cases where the term is nonzero: Case I, $G_2(i, \mathbf{j})$ contains two non-coincident J-vertices of degree 2, and three possibilities to distribute the four J-vertices into the two non-coincident J vertices; Case II, $G_2(i, \mathbf{j})$ contains only one non-coincident J-vertex of degree 4. It follows that

$$\begin{aligned}ET_2 &= \sum_{i=1}^{n} \sum_{\mathbf{j}} A_{G_2(i,\mathbf{j})} E\varepsilon_{G_2(i,\mathbf{j})} \\ &= \sum_{i=1}^{n} \sum_{j_1 \neq j_2} 3A_{G_2(i,j_1,j_1,j_2,j_2)} + M_4 \sum_{i=1}^{n} \sum_{j=1}^{n} A_{G_2(i,j,j,j,j)} \\ &= 3\sum_{i=1}^{n} \sum_{j_1, j_2} a_{ij_1}^2 a_{ij_2}^2 + \nu_4 \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^4 \\ &= 3\mathrm{tr}\left(\mathbf{AA}'\right) \circ \left(\mathbf{AA}'\right) + \nu_4 \mathrm{tr}\left(\mathbf{A}^{\circ 2} \mathbf{A}'^{\circ 2}\right).\end{aligned}$$

The readers should note that the third equality follows by adding the terms for $j_1 = j_2$ to the first sum and subtracting the same terms from the second sum, which produces the fourth cumulant of the underlying variable.

To calculate the variance of T_2, we similarly draw another G_2 graph from i_2 to $\mathbf{j}_2 = (j_5, j_6, j_7, j_8)$ such that

$$\begin{aligned}\operatorname{Var}(T_2) &= \sum_{i_1,i_2=1}^{n} \sum_{j_1,\ldots,j_8=1}^{n} E(G_2(i_1,\mathbf{j}_1) - EG_2(i_1,\mathbf{j}_1))(G_2(i_2,\mathbf{j}_2) - EG_2(i_2,\mathbf{j}_2)) \\ &= \sum_{i_1,i_2=1}^{n} \sum_{j_1,\ldots,j_8=1}^{n} E(G_2(i_1,\mathbf{j}_1)G_2(i_2,\mathbf{j}_2)) - EG_2(i_1,\mathbf{j}_1)EG_2(i_2,\mathbf{j}_2).\end{aligned} \tag{2.7}$$

It is easy to see that the term is zero if the two G_2 graphs do not have coincident J vertices due to the independence of the two factors $G_2(i_1, \mathbf{j}_1)$ and $G_2(i_2, \mathbf{j}_2)$. Also, the term is zero if the degree of a non-coincident J-vertex of the combined graph $G_2(i_1, \mathbf{j}_1) \cup G_2(i_2, \mathbf{j}_2)$ is odd due to the fact that the distribution of the underlying variables is symmetric.

We below investigate the expansion of (2.7) term by term. Firstly, consider the terms involving the second moment of ε only (that is, the degrees of non-coincident J vertices are all 2). The contribution to this case only comes from $E(G_2(i_1, \mathbf{j}_1)G_2(i_2, \mathbf{j}_2))$, because $G_2(i_1, \mathbf{j}_1)$ and $G_2(i_2, \mathbf{j}_2)$ have at least one coincident J-vertex of degree 1 and hence $E(G_2(i_1, \mathbf{j}_1)) = E(G_2(i_2, \mathbf{j}_2)) = 0$. This case consists of two sub-cases:

$$a_{i_1 j_1}^2 a_{i_2 j_2}^2 a_{i_1 j_3} a_{i_1 j_4} a_{i_2 j_3} a_{i_2 j_4} \quad \text{and} \quad a_{i_1 j_1} a_{i_1 j_2} a_{i_1 j_3} a_{i_1 j_4} a_{i_2 j_1} a_{i_2 j_2} a_{i_2 j_3} a_{i_2 j_4}.$$

The first sub-case has $\binom{4}{2}^2\binom{2}{1} = 72$ possibilities, and the second has $4! = 24$ possibilities. It follows from the inclusion–exclusion principle (IEP) that

$$\begin{aligned}&\Bigg(IEP : \Big\{ j_1, j_2, j_3, j_4 \text{ distinct}\} = \bigcap_{t<s} (\Omega - \{j_t = j_s\}\Big) \\ &\quad = \Omega - \sum\{\text{one equal sign}\} + \sum\{\text{two equal signs}\} \\ &\quad - \cdots - \sum\{6 \text{ equal signs}\} \\ &\quad \text{where } \Omega \text{ stands for there are no restrictions among } j_1, j_2, j_3 \text{ and } j_4\Bigg) \\ &\quad 72 \sum_{\substack{i_1,i_2 \\ j_1,j_2,j_3,j_4 \text{ distinct}}} a_{i_1 j_1}^2 a_{i_2 j_2}^2 a_{i_1 j_3} a_{i_1 j_4} a_{i_2 j_3} a_{i_2 j_4}\end{aligned}$$

$$
\begin{aligned}
=72 \Bigg(& \sum_{i_1,i_2,j_1,j_2,j_3,j_4} a_{i_1j_1}^2 a_{i_2j_2}^2 a_{i_1j_3} a_{i_1j_4} a_{i_2j_3} a_{i_2j_4} \\
& - \underbrace{\sum_{i_1,i_2,j_1j_2j_3} a_{i_1j_1}^2 a_{i_2j_1}^2 a_{i_1j_2} a_{i_1j_3} a_{i_2j_2} a_{i_2j_3}}_{j_1=j_2} \\
& - \underbrace{4 \sum_{i_1,i_2,j_1j_2j_3} a_{i_1j_1}^3 a_{i_2j_2}^2 a_{i_1j_3} a_{i_2j_1} a_{i_2j_3}}_{j_1=j_3 \text{ or } j_1=j_4 \text{ or } j_2=j_3 \text{ or } j_2=j_4} - \underbrace{\sum_{i_1,i_2,j_1j_2j_3} a_{i_1j_1}^2 a_{i_2j_2}^2 a_{i_1j_3}^2 a_{i_2j_3}^2}_{j_3=j_4} \\
& + \underbrace{4 \sum_{i_1i_2j_1j_2} a_{i_1j_1}^4 a_{i_2j_2}^2 a_{i_2j_1}^2}_{\substack{j_1=j_3=j_4;\ 6 \text{ from two equal signs} \\ -2 \text{ from three equal signs}}} + \underbrace{4 \sum_{i_1i_2j_1j_2} a_{i_1j_1}^3 a_{i_2j_1}^3 a_{i_1j_2} a_{i_2j_2}}_{\substack{j_1=j_2=j_3;\ 6 \text{ from two equal signs} \\ -2 \text{ from three equal signs}}} \\
& + \underbrace{2 \sum_{i_1i_2j_1j_2} a_{i_1j_1}^3 a_{i_2j_2}^3 a_{i_1j_2} a_{i_2j_1}}_{\{j_1=j_3 \text{ and } j_2=j_4\} \text{ or } \{j_1=j_4 \text{ and } j_2=j_3\}} \\
& + \underbrace{\sum_{i_1i_2j_1j_2} a_{i_1j_1}^2 a_{i_2j_1}^2 a_{i_1j_2}^2 a_{i_2j_2}^2}_{j_1=j_2 \text{ and } j_3=j_4} - \underbrace{6 \sum_{i_1i_2j} a_{i_1j}^4 a_{i_2j}^4}_{\substack{j_1=j_2=j_3=j_4;\ -16 \text{ from three; } 15 \text{ from four} \\ -6 \text{ from five } 1 \text{ from six equal signs}}} \Bigg)
\end{aligned}
$$

and that

$$
\begin{aligned}
24 & \sum_{\substack{i_1,i_2 \\ j_1,j_2,j_3,j_4 \text{ distinct}}} a_{i_1j_1} a_{i_1j_2} a_{i_1j_3} a_{i_1j_4} a_{i_2j_1} a_{i_2j_2} a_{i_2j_3} a_{i_2j_4} \\
= 24 \Bigg(& \sum_{i_1,i_2,j_1,j_2,j_3,j_4} a_{i_1j_1} a_{i_1j_2} a_{i_1j_3} a_{i_1j_4} a_{i_2j_1} a_{i_2j_2} a_{i_2j_3} a_{i_2j_4} \\
& - \underbrace{6 \sum_{i_1,i_2,j_1j_2j_3} a_{i_1j_1}^2 a_{i_2j_1}^2 a_{i_1j_2} a_{i_1j_3} a_{i_2j_2} a_{i_2j_3}}_{j_1=j_2 \text{ or } \cdots \text{ or } j_3=j_4}
\end{aligned}
$$

$$
\begin{aligned}
&+ \underbrace{8\sum_{i_1 i_2 j_1 j_2} a^3_{i_1 j_1} a^3_{i_2 j_1} a_{i_1 j_2} a_{i_2 j_2}}_{\substack{j_1=j_2=j_3;\ 12 \text{ from two equal signs} \\ -4 \text{ from three equal signs}}} \\
&+ \underbrace{3\sum_{i_1 i_2 j_1 j_2} a^2_{i_1 j_1} a^2_{i_2 j_1} a^2_{i_1 j_2} a^2_{i_2 j_2}}_{\{j_1=j_2 \text{ and} j_3=j_4\} \text{ or } j_1=j_3 \text{ and } j_2=j_4 \text{or } \{j_1=j_4 \text{ and } j_2=j_3\}} \\
&- \left. \underbrace{6\sum_{i_1 i_2 j} a^4_{i_1 j} a^4_{i_2 j}}_{\substack{j_1=j_2=j_3=j_4;\ -16 \text{ from three; } 15 \text{ from four} \\ -6 \text{ from five } 1 \text{ from six equal signs}}} \right)
\end{aligned}
$$

Secondly, consider the terms containing one 4th moment and two second moments of ε (the degrees of non-coincident J vertices are 4, 2 and 2), which may happen in the following three sub-cases: $a^3_{i_1 j_1} a_{i_1 j_2} a_{i_2 j_1} a_{i_2 j_2} a^2_{i_2 j_3}$, $a^2_{i_1 j_1} a_{i_1 j_2} a_{i_1 j_3} a^2_{i_2 j_1} a_{i_2 j_2} a_{i_2 j_3}$ and $a^2_{i_1 j_1} a^2_{i_1 j_2} a^2_{i_2 j_1} a^2_{i_2 j_3}$. The first sub-case has $\binom{4}{1}^2 \binom{3}{1} \binom{2}{1} = 96$ possibilities, the second has $\binom{4}{2}^2 \binom{2}{1} = 72$ possibilities, and the third has $\binom{4}{2}^2 = 36$ possibilities.

The contribution of the first case is

$$
\begin{aligned}
96 M_4 & \sum_{\substack{i_1 i_2 \\ j_1 j_2 j_3 \text{ distinct}}} a^3_{i_1 j_1} a_{i_1 j_2} a_{i_2 j_1} a_{i_2 j_2} a^2_{i_2 j_3} \\
&= 96 M_4 \left(\sum_{i_1 i_2 j_1 j_2 j_3} a^3_{i_1 j_1} a_{i_1 j_2} a_{i_2 j_1} a_{i_2 j_2} a^2_{i_2 j_3} - \underbrace{\sum_{i_1 i_2 j_1 j_2} a^4_{i_1 j_1} a^2_{i_2 j_1} a^2_{i_2 j_2}}_{j_1=j_2} \right. \\
&\quad - \underbrace{\sum_{i_1 i_2 j_1 j_2} a^3_{i_1 j_1} a_{i_1 j_2} a_{i_2 j_1} a^3_{i_2 j_2}}_{j_2=j_3} - \underbrace{\sum_{i_1 i_2 j_1 j_2} a^3_{i_1 j_1} a_{i_1 j_2} a^3_{i_2 j_1} a_{i_2 j_2}}_{j_1=j_3} \\
&\quad + \left. \underbrace{2\sum_{i_1 i_2 j} a^4_{i_1 j} a^4_{i_2 j}}_{\substack{j_1=j_2=j_3;\ 3 \text{ from two equal} \\ \text{signs and} -1 \text{ from three equal signs.}}} \right).
\end{aligned}
$$

The contribution of the second case is

$$72M_4 \sum_{\substack{i_1 i_2 \\ j_1 j_2 j_3 \text{ distinct}}} a_{i_1 j_1}^2 a_{i_1 j_2} a_{i_1 j_3} a_{i_2 j_1}^2 a_{i_2 j_2} a_{i_2 j_3}$$

$$= 72M_4 \left(\sum_{i_1 i_2 j_1 j_2 j_3} a_{i_1 j_1}^2 a_{i_1 j_2} a_{i_1 j_3} a_{i_2 j_1}^2 a_{i_2 j_2} a_{i_2 j_3} - \underbrace{2 \sum_{i_1 i_2 j_1 j_2} a_{i_1 j_1}^3 a_{i_2 j_1}^3 a_{i_1 j_2} a_{i_2 j_2}}_{j_1 = j_2 \text{ or } j_1 = j_3} \right.$$

$$\left. - \underbrace{\sum_{i_1 i_2 j_1 j_2} a_{i_1 j_1}^2 a_{i_1 j_2}^2 a_{i_2 j_1}^2 a_{i_2 j_2}^2}_{j_2 = j_3} + \underbrace{2 \sum_{i_1 i_2 j} a_{i_1 j}^4 a_{i_2 j}^4}_{\substack{j_1 = j_2 = j_3;\ 3 \text{ from two equal} \\ \text{signs and } -1 \text{ from three equal signs.}}} \right).$$

In the two cases above, the contributions come only from $EG_1(i_1, \mathbf{j}_1)G_1(i_2, \mathbf{j}_2)$ because $EG_1(i_1, \mathbf{j}_1) = EG_1(i_2, \mathbf{j}_2) = 0$. The contribution of the third case is

$$36(M_4 - 1) \sum_{\substack{i_1 i_2 \\ j_1 j_2 j_3 \text{ distinct}}} a_{i_1 j_1}^2 a_{i_1 j_2}^2 a_{i_2 j_1}^2 a_{i_2 j_3}^2$$

$$= 36(M_4 - 1) \left(\sum_{i_1 i_2 j_1 j_2 j_3} a_{i_1 j_1}^2 a_{i_1 j_2}^2 a_{i_2 j_1}^2 a_{i_2 j_3}^2 - \underbrace{2 \sum_{i_1 i_2 j_1 j_2} a_{i_1 j_1}^4 a_{i_2 j_1}^2 a_{i_2 j_2}^2}_{j_1 = j_2 \text{ or } j_1 = j_3} \right.$$

$$\left. - \underbrace{\sum_{i_1 i_2 j_1 j_2} a_{i_1 j_1}^2 a_{i_1 j_2}^2 a_{i_2 j_1}^2 a_{i_2 j_2}^2}_{j_2 = j_3} + \underbrace{2 \sum_{i_1 i_2 j} a_{i_1 j}^4 a_{i_2 j}^4}_{\substack{j_1 = j_2 = j_3;\ 3 \text{ from two equal} \\ \text{signs and } -1 \text{ from three equal signs.}}} \right),$$

where M_4 comes from $EG_2(i_1, \mathbf{j}_1)G_2(i_2, \mathbf{j}_2)$ and -1 comes from $-EG_2(i_1, \mathbf{j}_1)EG_2$ $(i_2, \mathbf{j}_2)$.

Thirdly, consider the term containing only two 4th moments of ε (the degrees of non-coincident J vertices are 4 and 4), which consists of two sub-cases: $a_{i_1 j_1}^2 a_{i_1 j_2}^2 a_{i_2 j_1}^2 a_{i_2 j_2}^2$

and $a^3_{i_1j_1}a_{i_1j_2}a_{i_2j_1}a^3_{i_2j_2}$. The first sub-case has $\frac{1}{2}\binom{4}{2}^2 = 18$ possibilities, and the second has $\binom{4}{1}^2 = 16$ possibilities. The contribution of the first sub-case is

$$18\left(M_4^2 - 1\right) \sum_{\substack{i_1i_2 \\ j_1j_2 \text{ distinct}}} a^2_{i_1j_1}a^2_{i_1j_2}a^2_{i_2j_1}a^2_{i_2j_2}$$

$$= 18\left(M_4^2 - 1\right)\left(\sum_{i_1i_2j_1j_2} a^2_{i_1j_1}a^2_{i_1j_2}a^2_{i_2j_1}a^2_{i_2j_2} - \sum_{i_1i_2j} a^4_{i_1j}a^4_{i_2j}\right).$$

The contribution of the second sub-case is

$$16M_4^2 \sum_{\substack{i_1i_2 \\ j_1j_2 \text{ distinct}}} a^3_{i_1j_1}a_{i_1j_2}a_{i_2j_1}a^3_{i_2j_2}$$

$$= 16M_4^2\left(\sum_{i_1i_2j_1j_2} a^3_{i_1j_1}a_{i_1j_2}a_{i_2j_1}a^3_{i_2j_2} - \sum_{i_1i_2j} a^4_{i_1j}a^4_{i_2j}\right).$$

Fourthly, look at the term containing one 6th moment and one second moment of ε (the degrees of non-coincident J vertices are 6 and 2), which includes two sub-cases: $a^4_{i_1j_1}a^2_{i_2j_1}a^2_{i_2j_2}$ and $a^3_{i_1j_1}a_{i_1j_2}a^2_{i_2j_1}a_{i_2j_2}$. The first sub-case has $\binom{4}{2}\binom{2}{1} = 12$ possibilities, and the second has $\binom{4}{1}^2 = 16$ possibilities. The contribution of the first sub-case is

$$12(M_6 - M_4) \sum_{\substack{i_1i_2 \\ j_1j_2 \text{ distinct}}} a^4_{i_1j_1}a^2_{i_2j_1}a^2_{i_2j_2}$$

$$= 12(M_6 - M_4)\left(\sum_{i_1i_2j_1j_2} a^4_{i_1j_1}a^2_{i_2j_1}a^2_{i_2j_2} - \sum_{i_1i_2j} a^4_{i_1j}a^4_{i_2j}\right).$$

The contribution of the second sub-case is

$$16M_6 \sum_{\substack{i_1i_2 \\ j_1j_2 \text{ distinct}}} a^3_{i_1j_1}a_{i_1j_2}a^3_{i_2j_1}a_{i_2j_2}$$

$$= 16M_6\left(\sum_{i_1i_2j_1j_2} a^3_{i_1j_1}a_{i_1j_2}a^3_{i_2j_1}a_{i_2j_2} - \sum_{i_1i_2j} a^4_{i_1j}a^4_{i_2j}\right).$$

Finally, there is one case when the term contains the 8th moment of ε (the degree of non-coincident J vertices is 8) $a^4_{i_1j}a^4_{i_2j}$. The arguments above show the validity of (2.3).

Finally, we calculate the covariance of T_1 and T_2

$$
\begin{aligned}
\Sigma_{12} = \Sigma_{21} &= \sum_{i_1 i_2, j_1, j_2, \mathbf{j}_3} E(G_1(i_1, j_1 j_2) - EG_1(i_1, j_1 j_2))(G_2(i_2, \mathbf{j}_3) - EG_2(i_2, \mathbf{j}_3)) \\
&= \Bigg(12 \sum_{i_1 i_2 j_1 j_2 j_3} a_{i_1 j_1}^2 a_{i_1 j_2} a_{i_1 j_3} a_{i_2 j_2} a_{i_2 j_3} + 8\nu_4 \sum_{i_1 i_2 j_1 j_2} a_{i_1 j_1}^3 a_{i_1 j_2} a_{i_1 j_1} a_{i_2 j_2} \\
&\quad + 6\nu_4 \sum_{i_1 i_2 j_1 j_2} a_{i_1 j_1}^2 a_{i_1 j_2}^2 a_{i_1 j_1}^2 + \nu_6 \sum_{i_1 i_2 j} a_{i_1 j}^4 a_{i_2 j}^2 \Bigg) \\
&= \left(12 \mathrm{tr} \left((\mathbf{AA}')^2 \circ (\mathbf{AA}') \right) + 8\nu_4 \mathrm{tr} \left((\mathbf{AA}'\mathbf{A}) \mathbf{A}'^{\circ 3} \right) \right. \\
&\quad \left. + 6\nu_4 \mathrm{tr} \left((\mathbf{A}'\mathbf{D}_{\mathbf{AA}'}\mathbf{A}) \circ (\mathbf{A}'\mathbf{A}) \right) + \nu_6 \left[\mathrm{Diag}' (\mathbf{A}'\mathbf{A}) \left(\mathbf{A}'^{\circ 4} \right) \mathbf{1} \right] \right). \qquad (2.8)
\end{aligned}
$$

2.2 Proof of Theorem 2.1

This subsection is to prove the main theorem.

2.2.1 Truncation

Since $E\varepsilon_1^8 < \infty$, for any $\eta > 0$ we have

$$\eta^{-8} n P \left(|\varepsilon_1| \geq \eta n^{1/8} \right) \to 0.$$

Hence, we may select a sequence $\eta_n \to 0$ such that $\eta_n^{-8} n P(|\varepsilon| > \eta_n n^{1/8}) \to 0$. Let $\tilde{\varepsilon}_j = \varepsilon_j I(|\varepsilon_j| < \eta_n n^{1/8}$. Define $(\widetilde{T}_1, \widetilde{T}_2)$ to be the analogue of (T_1, T_2) with ε_j replaced by $\tilde{\varepsilon}_j$. Then

$$P\left((T_1, T_2) \neq (\widetilde{T}_1, \widetilde{T}_2) \right) \leq \sum_{j=1}^{n} P\left(\varepsilon_j \neq \tilde{\varepsilon}_j \right) = nP\left(|\varepsilon| \geq \eta_n n^{1/8} \right) \to 0.$$

Therefore, (T_1, T_2) has the same limiting distribution as $(\widetilde{T}_1, \widetilde{T}_2)$.

2.2.2 Normalization

Denote by $d^2(X, Y) = E|X - Y|^2$ the distance $d(X, Y)$ between random variables X and Y. Since the distribution of ε_j's is symmetric, we do not need to centralize them. Define $(\overline{T}_1, \overline{T}_2)$ to be the one obtained from (T_1, T_2) with ε_j replaced by $\bar{\varepsilon}_j = \frac{\tilde{\varepsilon}_j}{\sqrt{E\tilde{\varepsilon}_j^2}}$. It follows that

$$d(\widetilde{T}_1, \overline{T}_1) \leq \sum_{i=1}^{n} d\left((A_i \tilde{\varepsilon})^4, (A_i \bar{\varepsilon})^4 \right)$$

$$= \sum_{i=1}^{n} E^{1/2} \left|(A_i\tilde{\varepsilon})^3(A_i(\tilde{\varepsilon}-\bar{\varepsilon})+(A_i\tilde{\varepsilon})^2(A_i(\tilde{\varepsilon}-\bar{\varepsilon})^2+(A_i\tilde{\varepsilon})(A_i(\tilde{\varepsilon}-\bar{\varepsilon})^3\right|^2$$
$$\leq \sum_{i=1}^{n} \left(E^{1/2} \left|(A_i\tilde{\varepsilon})^3(A_i(\tilde{\varepsilon}-\bar{\varepsilon})\right|^2 + E^{1/2} \left|(A_i\tilde{\varepsilon})^2(A_i(\tilde{\varepsilon}-\bar{\varepsilon})^2\right|^2 \right.$$
$$\left. + E^{1/2} \left|(A_i\tilde{\varepsilon})(A_i(\tilde{\varepsilon}-\bar{\varepsilon})^3\right|^2 \right), \tag{2.9}$$

where A_i stands for the i-th row of matrix $\mathbf{A}$.

To estimate the above quantity, note that

$$E\left((A_i\tilde{\varepsilon})^3(A_i(\tilde{\varepsilon}-\bar{\varepsilon})\right)^2 = E(A_i\tilde{\varepsilon})^8\left(1-\frac{1}{\sqrt{E\tilde{\varepsilon}^2}}\right)^2$$
$$\leq \left(\sum_{j=1}^{n} a_{ij}^8 E\tilde{\varepsilon}_1^8 + 28\sum_{j_1,j_2=1}^{n} a_{ij_1}^6 a_{ij_2}^2 E\tilde{\varepsilon}_1^6 E\tilde{\varepsilon}_1^2 + 35\sum_{j_1,j_2=1}^{n} a_{ij_1}^4 a_{ij_2}^4 E\tilde{\varepsilon}_1^4 E\tilde{\varepsilon}_1^4\right.$$
$$\left. + 210\sum_{j_1j_2j_3j_4=1}^{n} a_{ij_1}^2 a_{ij_2}^2 a_{ij_3}^2 a_{ij_4}^2\left(E\tilde{\varepsilon}_1^2\right)^4\right)\left(E\varepsilon^2 I\left(|\varepsilon|\geq \eta_n n^{1/8}\right)\right)^2$$
$$\leq K\eta_n^2 n^{-5/4}.$$

It is easy to see that the same bound holds for the other two terms on the right-hand side of (2.9). Therefore,

$$d(\widetilde{T}_1, \overline{T}_1) \leq K\eta_n n^{3/8}. \tag{2.10}$$

Similarly, one can prove that $d(\widetilde{T}_2, \overline{T}_2) = o(n^{3/8})$. Noting that the covariance matrix Σ has at most the same order as n, we conclude that $n^{-1/2}(T_1, T_2)$ has the same limiting distribution as $n^{-1/2}(\overline{T}_1, \overline{T}_2)$.

Therefore, we below assume that the addition condition that $|\varepsilon_j| \leq \eta_n n^{1/8}$ holds in the proof of the CLT of (T_1, T_2).

2.2.3 Proof of Theorem 2.1

The proof is to use the moment method. Under the assumption that the 8th moment of ε_1 is finite and the distribution of ε_1 is symmetric, we may truncate the underlying variables at $\eta_n n^{1/8}$ with $\eta_n \to 0$ and renormalize them to have mean 0 and variance 1.

Let $(\alpha, \beta) \neq (0, 0)$ be two real numbers. The aim is to show that for any integer k

$$E(\alpha(T_1 - ET_1) + \beta(T_2 - ET_2))^k$$
$$= \begin{cases} (k-1)!!\left(\alpha^2\Sigma_{11} + 2\alpha\beta\Sigma_{12} + \beta^2\Sigma_{22}\right)^{k/2}(1+o(1)) & \text{if } k \text{ is even,} \\ o\left(n^{k/2}\right) & \text{if } k \text{ is odd.} \end{cases} \tag{2.11}$$

Write

$$\begin{aligned}&\mathrm{E}\left(\alpha(T_1-ET_1)+\beta(T_2-ET_2)\right)^k\\&\quad=\sum_{k_1+k_2=k}\frac{k!}{k_1!k_2!}\alpha^{k_1}\beta^{k_2}\mathrm{E}\left((T_1-ET_1)^{k_1}(T_2-ET_2)^{k_2}\right).\end{aligned}\tag{2.12}$$

We will complete our proof by employing graphs that are introduced in Section 2.1. For this purpose, we draw k_1 G_1 graphs and k_2 G_2 graphs as follows. Draw parallel an I line and a J line, select k_1 points $i_1,\ldots,i_{k_1}$ on the I line and $2k_1$ points $j_{\ell,t}$, $\ell=1,\ldots,k_1, t=1,2$ on the J line, and then draw k_1 G_1 graphs of 2 edges each from i_ℓ to $j_{\ell,s}$ $s=1,2$. Denote them by $G_{1\ell}, \ell=1,\ldots,k_1$. Similarly, draw k_2 four-edge G_2 graph between the I line and J line and denote them by $G_{2l}(i_l,\mathbf{j}_l), l=1,\ldots,k_2$.

Let

$$\pi_{G_{1\ell}}=\prod_{t=1}^{2}a_{i_\ell,j_{\ell,t}}\varepsilon_{j_{\ell,t}}=\mathbf{A}_{G_{1\ell}}\boldsymbol{\varepsilon}_{G_{1\ell}},\qquad \pi_{G_{2l}}=\prod_{t=1}^{4}a_{i_l,j_{l,t}}\varepsilon_{j_{l,t}}=\mathbf{A}_{G_{2l}}\boldsymbol{\varepsilon}_{G_{2l}},$$

where a_{ij} is the (i,j)th entry of the matrix $\mathbf{A}$.

We then further expand

$$\begin{aligned}&\mathrm{E}\left((T_1-\mathrm{E}T_1)^{k_1}(T_2-\mathrm{E}T_2)^{k_2}\right)\\&\quad=\sum_{G_1,G_2}\mathrm{E}\left(\prod_{\ell=1}^{k_1}\left(\pi_{G_{1\ell}}-\mathrm{E}\pi_{G_{1\ell}}\right)\prod_{l=1}^{k_2}\left(\pi_{G_{2l}}-\mathrm{E}\pi_{G_{2l}}\right)\right),\end{aligned}\tag{2.13}$$

where the summation runs over all possibilities of the G_1 and G_2 graphs. Now, we have associated the terms in the expatiation of $\mathrm{E}\left((T_1-\mathrm{E}T_1)^{k_1}(T_2-\mathrm{E}T_2)^{k_2}\right)$ with graphs. By independence of the underlying variables, it is easy to see the following facts:

(i) If one subgraph G does not have any J vertices coincident with J vertices of the other subgraphs, then the term is zero.
(ii) If the degree of a non-coincident J-vertices of the combined graph $G=\bigcup_{\ell=1}^{k_1}G_{1\ell}\bigcup_{l=1}^{k_2}G_{2l}$ is odd, the term is zero because the underlying distribution is assumed to be symmetric.

Therefore, we only need to evaluate the sum of terms with graphs which do not fall into the above cases. Suppose the combined graph by G which contains μ connected pieces $G_1,\ldots,G_\mu$ consisting of $\iota_1,\ldots,\iota_\mu$ subgraphs. It is easy to see that $\iota_h\geq 2$ $(1\leq h\leq\mu)$ and hence $\mu\leq k/2$ because $\iota_1+\cdots+\iota_\mu=k$.

Lemma 2.2 *For the h-th connected graph, if $\iota_h>2$ for some $h\leq\mu$, then*

$$\sum_{G_h}\mathrm{E}\prod_{t=1}^{\iota_h}\left(\mathbf{A}_{G_{ht}}\boldsymbol{\varepsilon}_{G_{ht}}-\mathrm{E}\mathbf{A}_{G_{ht}}\boldsymbol{\varepsilon}_{G_{ht}}\right)=o\left(n^{\iota_h/2}\right)\tag{2.14}$$

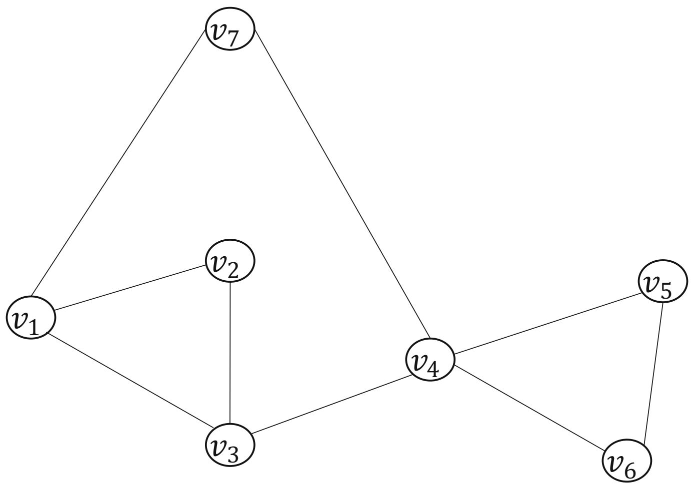

Fig. 1 An example of two-edge-connected graph

and if $\iota_h = 2$ *for* $h \leq \mu$, *then*

$$\sum_{G_h} E \prod_{t=1}^{2} \left(\mathbf{A}_{G_{ht}} \boldsymbol{\varepsilon}_{G_{ht}} - \mathrm{E}\mathbf{A}_{G_{ht}} \boldsymbol{\varepsilon}_{G_{ht}} \right) = O(n), \tag{2.15}$$

where the summation $\sum_{G_h}$ *runs over all possibilities of graphs homomorphic to* G_h, *the constructing graph* G_{ht} *is either a* G_1 *or a* G_2, *subgraph of* G_h, $\mathbf{A}_{G_{ht}}$ *denotes the product of the entries corresponding to the edges of* G_{ht}, *and* $\varepsilon_{G_{ht}}$ *stands for the product of the underlying variables* ε *corresponding to the* J*-vertices of the graph* G_{ht}.

To prove the above lemma, we need the next lemma about graph-associated multiple matrices. For the details of this lemma, one can refer to section **A**.4.2 in Bai and Silverstein (2010).

We first give some necessary definitions:

Definition 2.3 A graph G is called two-edge-connected if the resulting subgraph is still connected after removing any edge from G.

Definition 2.4 An edge e in a graph G is called a cutting edge if deleting this edge results in a disconnected subgraph.

Figure 1 shows an example of two-edge-connected graph, while the graph shown in Fig. 2 is not a two-edge-connected graph since e is a cutting edge.

Lemma 2.5 *Suppose that* $\mathbf{G} = (\mathbf{V}, \mathbf{E}, \mathbf{F})$ *is a two-edge-connected graph with* t *vertices and* k *edges. Each vertex* i *corresponds to an integer* $m_i \geq 2$ *and each edge* e_j

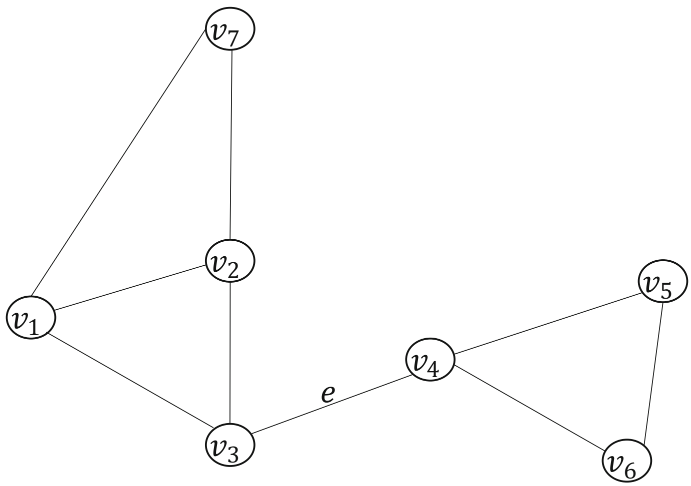

Fig. 2 An example of graph which is not two-edge-connected. The edge e is a cutting edge

corresponds to a matrix $\mathbf{T}^{(j)} = \left(t_{\alpha,\beta}^{(j)}\right)$, $j = 1, \dots, k$, *with consistent dimensions, that is, if* $F(e_j) = (f_i(e_j), f_e(e_j)) = (g, h)$, *then the matrix* $\mathbf{T}^{(j)}$ *has dimensions* $m_g \times m_h$. *Define* $\mathbf{v} = (v_1, v_2, \dots, v_t)$ *and*

$$T' = \sum_{\mathbf{v}} \prod_{j=1}^{k} t_{v_{f_i(e_j)}, v_{f_e(e_j)}}^{(j)}, \tag{2.16}$$

where the summation $\sum_{\mathbf{v}}$ *is taken for* $v_i = 1, 2, \dots, m_i, \; i = 1, 2, \cdots, t$. *Then for any* $i \le t$, *we have*

$$|T'| \le m_i \prod_{j=1}^{k} \|\mathbf{T}^{(j)}\|.$$

Remark 2.6 For the reader's convenience, we shall look into the first term in (2.8),

$$\sum_{i_1 i_2 j_1 j_2 j_3} a_{i_1 j_1}^2 a_{i_1 j_2} a_{i_1 j_3} a_{i_2 j_2} a_{i_2 j_3}$$

as an example. Lemma 2.5 indicates that if we let $\mathbf{T}^{(j)} = \mathbf{A}$ for $j = 1, \dots, k = 6$ and $m_i = n$ for the 5 vertices $i = 1, \dots, 5$, then we have

$$\sum_{i_1 i_2 j_1 j_2 j_3} a_{i_1 j_1}^2 a_{i_1 j_2} a_{i_1 j_3} a_{i_2 j_2} a_{i_2 j_3} = O(n).$$

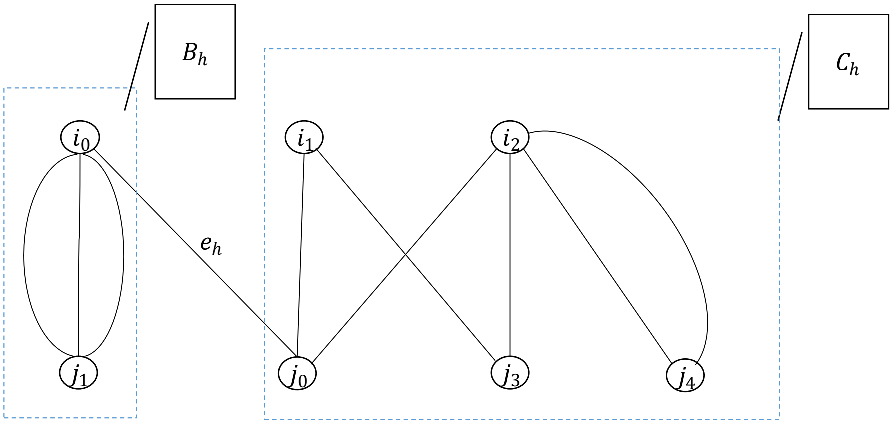

Fig. 3 An example of G_h

We now prove Lemma 2.2. Note that $\mathrm{E}\prod_{t=1}^{\iota_h}(\mathbf{A}_{G_{ht}}\boldsymbol{\varepsilon}_{G_{ht}} - \mathrm{E}\mathbf{A}_{G_{ht}}\boldsymbol{\varepsilon}_{G_{ht}}) = \mathbf{A}_{G_h}\mathrm{E}\prod_{t=1}^{\iota_h}(\boldsymbol{\varepsilon}_{G_{ht}} - \mathrm{E}\boldsymbol{\varepsilon}_{G_{ht}})$. We first show that the subgraph G_h is two-edge-connected. Otherwise, there exist subgraphs B_h, C_h and a cutting edge $e_h = (i_0, j_0)$ such that $G_h = B_h \cup C_h \cup e_h$ and $i_0 \in B_h$, $j_0 \in C_h$ and $B_h \cap C_h = \emptyset$ (see Fig. 3). If e_h is an edge of a constructing subgraph G_0 of G_h, then $G_0 \backslash e_h \subset B_h$. Therefore, the total degrees of J vertices of B_h must be odd and hence G_h must have a J vertex whose degree is odd. This is impossible because the term is not zero.

It follows from Lemma 2.5 that

$$\sum_{G_h} \mathbf{A}_{G_h} = O(n).$$

Furthermore, by the truncation, one can see that

$$|E\varepsilon_{G_h}| \le E|\varepsilon_1|^{4\iota_h} \begin{cases} \le \left(\eta_n^4\sqrt{n}\right)^{\iota_h-2} E|\varepsilon|^8 = o\left(n^{(\iota_h-2)/2}\right). & \text{if } \iota_h > 2 \\ = O(1), & \text{if } \iota_h = 2. \end{cases}$$

The two conclusions of Lemma 2.2 follow from these two estimates. The proof of the lemma is complete.

We next turn back to the proof of the CLT of (T_1, T_2). From Lemma 2.2, we conclude that the second conclusion of (2.11) holds if k is odd.

When k is even, consider a homomorphic class G which consists of u_1 connected subgraphs composed of two G_1 subgraphs, u_2 connected subgraphs composed of two G_2 graphs and u_3 connected subgraphs composed of one G_1 and one G_2 graphs. Comparing

$$\sum_G \mathrm{E}\prod_{\ell=1}^{k_1}\left(\pi_{G_{1\ell}} - \mathrm{E}\pi_{G_{1\ell}}\right)\prod_{l=1}^{k_2}\left(\pi_{G_{2l}} - \mathrm{E}\pi_{G_{2l}}\right) \tag{2.17}$$

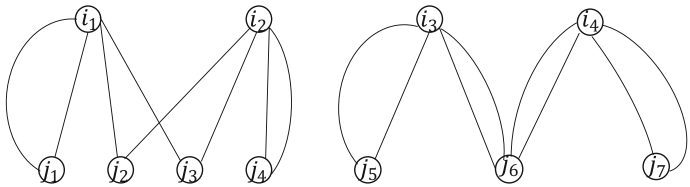

Fig. 4 An example of G_h

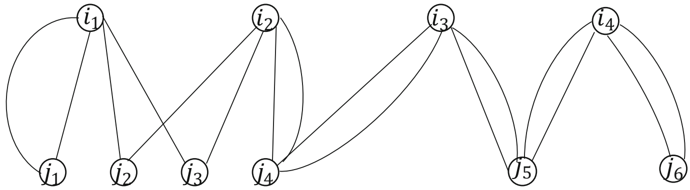

Fig. 5 An example of G_h

with the expansion of

$$\Sigma_{11}^{u_1}\Sigma_{22}^{u_2}\Sigma_{12}^{u_3} = \left(E(T_1 - ET_1)^2\right)^{u_1}\left(E(T_2 - ET_2)^2\right)^{u_2}\left(E(T_1 - ET_1)(T_2 - ET_2)\right)^{u_3}. \tag{2.18}$$

We find that the latter contains more terms than those in (2.17) with more connections among the subgraphs. As examples, see Fig. 4 for subgraphs belonging to both expressions and Fig. 5 for subgraphs belonging to (2.18) but not to (2.17).

and

Therefore, the difference between the two has an order of $o(n^{k/2})$ and thus

$$\sum_G \mathrm{E}\prod_{\ell=1}^{k_1}\left(\pi_{G_{1\ell}} - \mathrm{E}\pi_{G_{1\ell}}\right)\prod_{l=1}^{k_2}\left(\pi_{G_{2l}} - \mathrm{E}\pi_{G_{2l}}\right) = \Sigma_{11}^{u_1}\Sigma_{22}^{u_2}\Sigma_{12}^{u_3} + o\left(n^{k/2}\right). \tag{2.19}$$

Therefore,

$$\begin{aligned} &E(T_1 - ET_1)^{k_1}(T_2 - ET_2)^{k_2} \\ &= \sum_{\substack{2u_1+u_3=k_1 \\ 2u_2+u_3=k_2}} \frac{k_1!k_2!}{2^{u_1+u_2}u_1!u_2!u_3!}\Sigma_{11}^{u_1}\Sigma_{22}^{u_2}\Sigma_{12}^{u_3} + o\left(n^{k/2}\right). \end{aligned} \tag{2.20}$$

Substituting (2.20) into (2.12), we obtain

$$\begin{aligned} &E(\alpha(T_1 - ET_1) + \beta(T_2 - ET_2))^k \\ &= \sum_{u_1+u_2+u_3=k} \frac{k!}{2^{u_1+u_2}u_1!u_2!u_3!}\alpha^{2u_1+u_3}\beta^{2u_2+u_3}\Sigma_{11}^{u_1}\Sigma_{22}^{u_2}\Sigma_{12}^{u_3} + o\left(n^{k/2}\right) \end{aligned}$$

$$= \frac{k!}{2^{k/2}(k/2)!}\left(\alpha^2\Sigma_{11} + 2\alpha\beta\Sigma_{12} + \beta^2\Sigma_{22}\right)^{k/2} + o\left(n^{k/2}\right). \tag{2.21}$$

By (2.2), if $\nu_4 > -2$, and condition (2.1) holds, we know that Σ_{11} has an order of n since

$$\operatorname{tr}\left(\mathbf{A}'\mathbf{A}\right) \circ \left(\mathbf{A}'\mathbf{A}\right) \le \operatorname{tr}\left((\mathbf{A}\mathbf{A}')^2\right).$$

This is also true when $\nu_4 = -2$ under condition (2.5). Since that matrix $\Sigma = (\Sigma_{jk})$ is nonnegative definite, there are α and β such that $\alpha^2\Sigma_{11} + 2\alpha\beta\Sigma_{12} + \beta^2\Sigma_{22}$ has the order as n except at most two pairs of (α, β) with $\alpha^2 + \beta^2 = 1$.

Thus, by the moment convergence theorem, except for the two possible pairs (α, β),

$$(\alpha(T_1 - ET_1) + \beta(T_2 - ET_2)) \overset{\mathcal{D}}{\sim} N\left(0, \alpha^2\Sigma_{11} + 2\alpha\beta\Sigma_{12} + \beta^2\Sigma_{22}\right).$$

By continuity, we conclude that the convergence above is true for all (α, β). Therefore, we conclude that

$$\begin{pmatrix} T_1 - ET_1 \\ T_2 - ET_2 \end{pmatrix} \overset{\mathcal{D}}{\sim} N(\mathbf{0}, \Sigma).$$

The proof of Theorem 2.1 is complete.

3 Application to the homoscedasticity test

3.1 A brief review of homoscedasticity test

Consider the classical multivariate linear regression model of p covariates

$$y_i = \mathbf{x}_i\beta + \sigma_i\varepsilon_i, \qquad i = 1, 2, \ldots, n, \tag{3.1}$$

where y_i is the response variable, $\mathbf{x}_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,p})$ is the p-dimensional vector of covariates, $\beta = \left(\beta_1, \beta_2, \ldots, \beta_p\right)'$ is the p-dimensional regression coefficient vector and $\{\varepsilon_i\}$ are the i.i.d random errors with zero mean and unit variance. In most applications of the linear regression models, the homoscedasticity is a very important assumption. Without it, the loss in efficiency in using ordinary least squares (OLS) may be substantial and even worse, the biases in estimated standard errors may lead to invalid inferences. Thus, it is very important to examine the homoscedasticity. Formally, we need to test the hypothesis

$$H_0:\ \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2 = \sigma^2, \tag{3.2}$$

where σ^2 is a positive constant.

In the literature there is a lot of work considering this hypothesis test when the dimension p is fixed. Indeed, many popular tests have been proposed. For example, Breusch and Pagan (1979) and White (1980) proposed statistics to investigate the relationship between the estimated errors and the covariates in economics, while in statistics, Dette and Munk (1998), Glejser (1969), Harrison and McCabe (1979), Cook and Weisberg (1983), Azzalini and Bowman (1993) proposed nonparametric statistics to conduct the hypothesis. One may refer to Li and Yao (2015) for more details in this regard.

The development of computer science makes it possible for people to collect and deal with high-dimensional data. As a consequence, high-dimensional linear regression problems are becoming more and more common due to widely available covariates. Note that the above-mentioned tests are all developed under the low-dimensional framework when the dimension p is fixed and the sample size n tends to infinity.

In Li and Yao's paper, they proposed two test statistics in the high-dimensional setting by using the regression residuals. The first statistic uses the idea of likelihood ratio and the second one uses the idea that "the departure of a sequence of numbers from a constant can be efficiently assessed by its coefficient of variation," which is closely related to John's idea (John 1971). By assuming that the distribution of the covariates is $\mathbf{N}(\mathbf{0}, \mathbf{I}_p)$ and that the error obeys the normal distribution, the "coefficient of variation" statistic turns out to be a function of residuals. But its asymptotic distribution missed some part as indicated from the proof of Lemma 1 in Li and Yao (2015) even in the random design.

3.2 A procedure of homoscedasticity test

Let $\mathbf{X}$ be the design matrix, and $\mathbf{P} = (p_{i,j}) = \mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ the corresponding projection matrix. The n noise terms are represented as $\sigma_j\varepsilon_j$ $(1 \le j \le n)$, or $\mathbf{D}\boldsymbol{\varepsilon}$ in matrix form where $\mathbf{D} = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$. Then the OLS residuals are $\tilde{\boldsymbol{\varepsilon}} = (\tilde{\varepsilon}_1, \ldots, \tilde{\varepsilon}_n)' = \mathbf{PD}\boldsymbol{\varepsilon}$. For the robustness of the procedure, different from Li and Yao's procedure, we will use the *standardized residuals* $\hat{\varepsilon_j} = \tilde{\varepsilon}_j/\sqrt{p_{jj}}$, $1 \le j \le n$. It can be rewritten as $\hat{\boldsymbol{\varepsilon}} = (\hat{\varepsilon_1}, \ldots, \hat{\varepsilon_n})' = \mathbf{GPD}\boldsymbol{\varepsilon}$ where $\mathbf{G} = \text{diag}(p_{jj}^{-1/2})$. Consider the following statistic:

$$\mathbf{T} = \frac{\sum_{i=1}^n \left(\hat{\varepsilon_i}^2 - \frac{1}{n}\sum_{i=1}^n \hat{\varepsilon_i}^2\right)^2}{\frac{1}{n}\left(\sum_{i=1}^n \hat{\varepsilon_i}^2\right)^2}. \tag{3.3}$$

Set $T_1 = \sum_{i=1}^n \hat{\varepsilon_i}^2$ and $T_2 = \sum_{i=1}^n \hat{\varepsilon_i}^4$. By the δ-method and applying Theorem 2.1 with $\mathbf{A} = (a_{ij}) = \mathbf{GPD}$, or precisely $a_{ij} = \sigma_j p_{ij}/\sqrt{p_{ii}}$, we have

$$\frac{\mathbf{T} - a}{\sqrt{b}} \xrightarrow{\mathcal{D}} N(0, 1), \tag{3.4}$$

where the asymptotic mean is

$$a = \frac{nET_2}{(ET_1)^2} - 1 = \frac{n\left(3\mathrm{tr}\left(\mathbf{AA}' \circ \mathbf{AA}'\right) + \nu_4 \mathrm{tr}\left(\mathbf{A}^{\circ 2}\mathbf{A}'^{\circ 2}\right)\right)}{(\mathrm{tr}\,(\mathbf{AA}'))^2} - 1,$$

and the asymptotic variance is

$$b = (\Delta_1, \Delta_2)\Sigma \begin{pmatrix} \Delta_1 \\ \Delta_2 \end{pmatrix} = \Delta_1^2 \Sigma_{11} + 2\Delta_1\Delta_2\Sigma_{12} + \Delta_2^2\Sigma_{22},$$

with

$$\Delta_1 = -\frac{2n\left(3\mathrm{tr}\left(\mathbf{AA}' \circ \mathbf{AA}'\right)\right) + \nu_4 \mathrm{tr}\left(\mathbf{A}^{\circ 2}\mathbf{A}'^{\circ 2}\right)\big)}{(\mathrm{tr}\,(\mathbf{AA}'))^3}, \quad \Delta_2 = \frac{n}{(\mathrm{tr}\,(\mathbf{AA}'))^2},$$

$$\Sigma_{11} = 2\mathrm{tr}\left(\left(\mathbf{AA}'\right)^2\right) + \nu_4 \mathrm{tr}\left(\mathbf{A}'\mathbf{A}\right) \circ \left(\mathbf{A}'\mathbf{A}\right), \tag{3.5}$$

$$\begin{aligned}\Sigma_{12} = \Sigma_{21} = &\Big(12\mathrm{tr}\left(\left(\mathbf{AA}'\right)^2 \circ \left(\mathbf{AA}'\right)\right) + 8\nu_4 \mathrm{tr}\left(\left(\mathbf{AA}'\mathbf{A}\right)\mathbf{A}'^{\circ 3}\right) \\ &+ 6\,\nu_4 \mathrm{tr}\left(\left(\mathbf{A}'\mathbf{D}_{\mathbf{AA}'}\mathbf{A}\right) \circ \left(\mathbf{A}'\mathbf{A}\right)\right) + \nu_6 \left[\mathrm{Diag}'\left(\mathbf{A}'\mathbf{A}\right)\left(\mathbf{A}'^{\circ 4}\right)\mathbf{1}\right]\Big)\end{aligned} \tag{3.6}$$

and

$$\begin{aligned}\Sigma_{22} = &\,72\mathrm{Diag}'\left(\mathbf{AA}'\right)\left(\left(\mathbf{AA}'\right) \circ \left(\mathbf{AA}'\right)\right)\mathrm{Diag}\left(\mathbf{AA}'\right) + 24\mathrm{tr}\left(\left(\mathbf{AA}'\right) \circ \left(\mathbf{AA}'\right)\right)^2 \\ &+ \nu_4 \Big(96\mathrm{tr}(\mathbf{AA}')\mathbf{D}_{(\mathbf{AA}')}\mathbf{A}(\mathbf{A}')^{\circ 3} + 72\mathrm{tr}\left(\left(\left(\mathbf{AA}'\right) \circ \left(\mathbf{AA}'\right)\right)\left(\mathbf{A} \circ \mathbf{A}\right)\left(\mathbf{A}' \circ \mathbf{A}'\right)\right) \\ &+ 36\mathrm{Diag}'\left(\mathbf{AA}'\right)\left(\mathbf{A}^{\circ 2}(\mathbf{A}')^{\circ 2}\right)\mathrm{Diag}\left(\mathbf{AA}'\right)\Big) \\ &+ \nu_4^2\left(18\mathrm{tr}\left(\mathbf{A}^{\circ 2}(\mathbf{A}')^{\circ 2}\right)^2 + 16\mathrm{tr}\left(\mathbf{A}^{\circ 3}\mathbf{A}'\mathbf{A}(\mathbf{A}')^{\circ 3}\right)\right) \\ &+ \nu_6\left(12\mathrm{tr}\left(\left(\mathbf{A}'\mathbf{D}_{\mathbf{AA}'}\mathbf{A}\right) \circ \left(\mathbf{A}'^{\circ 2}\mathbf{A}^{\circ 2}\right)\right) + 16\mathrm{tr}(\mathbf{AA}')\mathbf{A}^{\circ 3}(\mathbf{A}')^{\circ 3}\right) \\ &+ \nu_8 \mathbf{1}'\left(\mathbf{A}^{\circ 4}(\mathbf{A}')^{\circ 4}\right)\mathbf{1}.\end{aligned} \tag{3.7}$$

Here $\nu_4 = M_4 - 3$, $\nu_6 = M_6 - 15M_4 + 30$ and $\nu_8 = M_8 - 28M_6 - 35M_4^2 + 420M_4 - 630$ are the corresponding cumulants of random variable ε_1.

Under the null hypothesis, by definition T is scale-invariant, so we below assume for simplicity and without loss of generality that $\sigma^2 = 1$ when evaluating the asymptotic parameters of T.

3.3 Some simulation results

We conduct some simulation results to investigate the performance of our test statistics. Firstly, we consider the condition when the random error obeys the normal distribution. Table 1 shows the empirical size compared with Li and Yao's result in Li and

Table 1 Empirical size under different distributions

p	$N(0,1)$		$t(1)$		$F(3,2)$		$e^{(N(5,3))}$	
	FCVT	CVT	FCVT	CVT	FCVT	CVT	FCVT	CVT
4	0.0577	0.0525	0.0611	0.0603	0.0596	0.0598	0.0575	0.0602
16	0.0591	0.0551	0.0574	0.0585	0.0588	0.0865	0.0604	0.0792
64	0.0604	0.0554	0.0593	0.2547	0.0543	0.2546	0.0576	0.2535
128	0.0587	0.0558	0.0587	0.5672	0.0587	0.5758	0.0580	0.6758
256	0.0554	0.0525	0.0602	0.9816	0.0586	0.9988	0.0604	0.9947
384	0.0571	0.0556	0.0591	1.0000	0.0588	1.0000	0.0600	1.0000

Table 2 Empirical power under model 1

p	$N(0,1)$		$t(1)$		$F(3,2)$		$e^{(N(5,3))}$	
	FCVT	CVT	FCVT	CVT	FCVT	CVT	FCVT	CVT
4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
64	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
128	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
256	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
384	0.8532	0.8614	0.9827	1.0000	0.9789	1.0000	0.9875	1.0000

Yao (2015) under four different design distributions. We use "CVT″ and "FCVT″ to represent their test and our test, respectively. The entries of design matrices are $i.i.d$ random samples generated from $N(0,1)$, $t(1)$ (t distribution with freedom degree 1), $F(3,2)$ (F distribution with parameters 3 and 2) and logarithmic normal distribution, respectively. The sample size n is 512 and the dimension of covariates varies from 4 to 384. We also follow Dette and Munk (1998) and consider the following two models:

Model 1 $\quad y_i = \mathbf{x}_i\beta + \varepsilon_i(1+\mathbf{x}_i\mathbf{h}), \qquad i = 1, 2, \ldots, n,$
where $\mathbf{h} = (1, \mathbf{0}_{(p-1)})$,

Model 2 $\quad y_i = \mathbf{x}_i\beta + \varepsilon_i(1+\mathbf{x}_i\mathbf{h}), \qquad i = 1, 2, \ldots, n$
where $\mathbf{h} = (\mathbf{1}_{(p/2)}, \mathbf{0}_{(p/2)})$.

Tables 2 and 3 show the empirical power compared with Li and Yao's results under the four different regressors distributions mentioned above.

According to the simulation result, it is shown that when $p/n \to [0, 1)$ as $n \to \infty$, our test always has good size and power under all regressors distributions.

3.4 Parameter estimation under null

If the underlying variables are not Gaussian, one needs to estimate the cumulants of the underlying variables to perform the test.

At first, the variance σ_0^2 can be estimated by $\hat{\sigma}_n^2 = \frac{T_1}{\mathrm{tr}(\mathbf{GPG})} = \frac{T_1}{n}$.

Table 3 Empirical power under model 2

p	$N(0,1)$		$t(1)$		$F(3,2)$		$e^{(N(5,3))}$	
	FCVT	CVT	FCVT	CVT	FCVT	CVT	FCVT	CVT
4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
64	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
128	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
256	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
384	0.9042	0.9047	0.9788	1.0000	0.9443	1.0000	0.9011	1.0000

Next, ν_4 can be estimated by

$$\hat{\nu}_4 = \frac{T_2 - 3\hat{\sigma_n^4}\mathrm{tr}(\mathbf{GPG} \circ \mathbf{GPG})}{\hat{\sigma_n^4}\mathrm{tr}(\mathbf{GP})^{\circ 2}(\mathbf{PG})^{\circ 2}} = \frac{T_2 - 3n\hat{\sigma_n^4}}{\hat{\sigma_n^4}\mathrm{tr}(\mathbf{GP})^{\circ 2}(\mathbf{PG})^{\circ 2}}.$$

To estimate ν_6 and ν_8, define $T_3 = \sum_{i=1}^n \hat{e}_i^6$ and $T_4 = \sum_{i=1}^n \hat{e}_i^8$. Note that

$$\begin{aligned}
ET_3 =& \sigma_0^6 \left(\nu_6 \sum_{i,j=1}^n (\mathbf{e_i GPe_j})^6 + 15\nu_4 \sum_{i,j_1 j_2} (\mathbf{e_i GPe}_{j_1})^4 (\mathbf{e_i GPe}_{j_2})^2 \right.\\
&\left. +15 \sum_{i,j_1 j_2 j_3 = 1}^n (\mathbf{e_i GPe}_{j_1})^2 (\mathbf{e_i GPe}_{j_2})^2 (\mathbf{e_i GPe}_{j_3})^2 \right)\\
=& \sigma_0^6 \left(\nu_6 \mathrm{tr}\, (\mathbf{GP})^{\circ 3} (\mathbf{PG})^{\circ 3} + 15\nu_4 \mathbf{1}' (\mathbf{GP})^{\circ 4} \mathbf{1} + 15n \right)\\
ET_4 =& \sigma_0^8 \left(\nu_8 \sum_{i,j=1}^n (\mathbf{e_i GPe_j})^8 + 28\nu_6 \sum_{i,j_1 j_2} (\mathbf{e_i GPe}_{j_1})^6 (\mathbf{e_i GPe}_{j_2})^2 \right.\\
& + 35\nu_4^2 \sum_{i,j_1 j_2} (\mathbf{e_i GPe}_{j_1})^4 (\mathbf{e_i GPe}_{j_2})^4\\
& + 210\nu_4 \sum_{i,j_1 j_2 j_3 = 1}^n (\mathbf{e_i GPe}_{j_1})^4 (\mathbf{e_i GPe}_{j_2})^2 (\mathbf{e_i GPe}_{j_3})^2\\
&\left. +105 \sum_{i,j_1 j_2 j_3 j_4 = 1}^n (\mathbf{e_i GPe}_{j_1})^2 (\mathbf{e_i GPe}_{j_2})^2 (\mathbf{e_i GPe}_{j_3})^2 (\mathbf{e_i GPe}_{j_4})^2 \right)\\
=& \sigma_0^8 \left(\nu_8 \mathrm{tr}\, (\mathbf{PG})^{\circ 4} (\mathbf{GP})^{\circ 4} + 28\nu_6 \mathbf{1}' (\mathbf{GP})^{\circ 4} \mathbf{1} + 35\nu_4^2 \mathbf{1}' (\mathbf{GP})^{\circ 4} (\mathbf{PG})^{\circ 4} \mathbf{1} \right.\\
&\left. +210\nu_4 \mathbf{1}' (\mathbf{GP})^{\circ 4} \mathbf{1} + 105n \right).
\end{aligned}$$

Therefore, we can estimate ν_6 and ν_8 by

$$\hat{\nu_6} = \frac{T_3 - \hat{\sigma_n^6}\left(15n + 15\hat{\nu_4}\mathbf{1}'(\mathbf{GP})^{\circ 4}\mathbf{1}\right)}{\hat{\sigma_n^6}\mathrm{tr}(\mathbf{GP})^{\circ 3}(\mathbf{PG})^{\circ 3}}$$

$$\hat{\nu_8} = \frac{T_4 - \hat{\sigma_n^8}\left(105n + 210\hat{\nu_4}\mathbf{1}'(\mathbf{GP})^{\circ 4}\mathbf{1} + 35\hat{\nu_4^2}\mathbf{1}'(\mathbf{GP})^{\circ 4}(\mathbf{PG})^{\circ 4}\mathbf{1} + 28\hat{\nu_6}\mathbf{1}'(\mathbf{GP})^{\circ 4}\mathbf{1}\right)}{\hat{\sigma_n^8}\mathrm{tr}(\mathbf{PG})^{\circ 4}(\mathbf{GP})^{\circ 4}}$$

It is not difficult to verify that all the estimators above are consistent.

4 Conclusion

In this paper, we established the CLTs as well as the joint CLTs for sums of linearly transformed random variables. We then give a modified test statistic to detect the existence of heteroscedasticity for both low- and high-dimensional linear regression models without assuming randomness of covariates. That is, this test is applicable to both fixed and random design matrices. Some simulations are also performed to investigate the advantage of the tests procedure we proposed.

References

Amari SV, Misra RB (1997) Closed-form expressions for distribution of sum of exponential random variables. IEEE Trans Reliab 46(4):519–522

Azzalini A, Bowman A (1993) On the use of nonparametric regression for checking linear relationships. J R Stat Soc Ser B Methodol 55(2):549–557

Bai Z, Silverstein JW (2010) Spectral analysis of large dimensional random matrices. Springer, Berlin

Breusch TS, Pagan AR (1979) A simple test for heteroscedasticity and random coefficient variation. Econom J Econom Soc 47(5):1287–1294

Cook RD, Weisberg S (1983) Diagnostics for heteroscedasticity in regression. Biometrika 70(1):1–10

de Jong P (1987) A central limit theorem for generalized quadratic forms. Probab Theory Relat Fields 75(2):261–277

Dette H, Munk A (1998) Testing heteroscedasticity in nonparametric regression. J R Stat Soc Ser B Stat Methodol 60(4):693–708

Deya A, Nourdin I (2014) Invariance principles for homogeneous sums of free random variables. Bernoulli 20(2):586–603

Glejser H (1969) A new test for heteroskedasticity. J Am Stat Assoc 64(325):316–323

Gotze F, Tikhomirov AN (1999) Asymptotic distribution of quadratic forms. Ann Probab 27(2):1072–1098

Harrison MJ, McCabe BPM (1979) A test for heteroscedasticity based on ordinary least squares residuals. J Am Stat Assoc 74(366a):494–499

Jensen DR, Solomon H (1972) A Gaussian approximation to the distribution of a definite quadratic form. J Am Stat Assoc 67(340):898–902

John S (1971) Some optimal multivariate tests. Biometrika 58(1):123–127

Li Z, Yao J (2015) Homoscedasticity tests valid in both low and high-dimensional regressions. arXiv preprint arXiv:1510.00097

Liu H, Tang Y, Zhang HH (2009) A new chi-square approximation to the distribution of non-negative definite quadratic forms in non-central normal variables. Comput Stat Data Anal 53(4):853–856

Nourdin I, Peccati G, Reinert G et al (2010) Invariance principles for homogeneous sums: universality of Gaussian Wiener chaos. Ann Probab 38(5):1947–1985

Nourdin I, Peccati G, Poly G, Simone R (2016) Multidimensional limit theorems for homogeneous sums: a survey and a general transfer principle. ESAIM Probab Stat 20:293–308

White H (1980) A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econom J Econom Soc 48(4):817–838

Whittle P (1964) On the convergence to normality of quadratic forms in independent variables. Theory Probab Appl 9(1):103–108