Supplementary Material for "Testing overidentifying restrictions on high-dimensional instruments and covariates"

Hongwei Shi[1], Xinyu Zhang[1], Xu Guo[†1], Baihua He[2], and Chenyang Wang[3]

[1]School of Statistics, Beijing Normal University, Beijing 100875, China

[2]International Institute of Finance, School of Management, University of Science and Technology of China, Hefei 230026, China

[3]School of Mathematics and Statistics, Beijing Technology and Business University, Beijing 102488, China

[†]Corresponding author: Xu Guo. Email address: xustat12@bnu.edu.cn.

Abstract

In the Supplementary material, Appendix A contains the proofs for a proposition and two theorems stated in the main text. Appendix B includes some useful lemmas.

Appendix A Proofs of main results

We denote the following notations in the Appendix. For two sequences of positive constants a_n and b_n, we write $a_n \lesssim b_n$ or $b_n \gtrsim a_n$ if there exists an absolute constant $C > 0$ such that $a_n \leq Cb_n$ when n is large. If $a_n \lesssim b_n$ and $a_n \gtrsim b_n$ hold simultaneously, a_n, b_n are asymptotically equal and it is denoted by $a_n \asymp b_n$. For a d-dimensional vector $\mathbf{U} = (U_1, \ldots, U_d)^\top \in \mathbb{R}^d$, we define $\|\mathbf{U}\|_q = (\sum_{j=1}^d |U_j|^q)^{1/q}$ with $1 \leq q < \infty$ and $\|\mathbf{U}\|_\infty = \max_{1\leq j\leq d} |U_j|$ to denote L_q and L_∞ norms of $\mathbf{U}$. For a random variable $X \in \mathbb{R}$, the L_q norm of X is defined as $\|X\|_q = \{\mathbb{E}(|X|^q)\}^{1/q}$ with $q \geq 1$. X is *sub-Gaussian* if the moment generating function (MGF) of X^2 is bounded at some point, namely $\mathbb{E}\exp(X^2/K^2) \leq 2$, where K is a positive constant. X is *sub-Exponential* if the MGF of $|X|$ is bounded at some point, namely $\mathbb{E}\exp(|X|/K') \leq 2$, where K' is a positive constant. The *sub-Gaussian norm* of a *sub-Gaussian* random variable X is defined as $\|X\|_{\psi_2} = \inf\{t > 0 : \mathbb{E}\exp(X^2/t^2) \leq 2\}$. We denote the unit sphere in $\mathbb{R}^d$ by $\mathbb{S}^{d-1} = \{\mathbf{U} \in \mathbb{R}^d : \|\mathbf{U}\|_2 = 1\}$. The *sub-Gaussian norm* of a *sub-Gaussian* random vector $\mathbf{U} \in \mathbb{R}^d$ is defined as $\|\mathbf{U}\|_{\psi_2} = \sup_{\mathbf{u}\in\mathbb{S}^{d-1}} \|\mathbf{u}^\top\mathbf{U}\|_{\psi_2}$. For a $(q_1 \times q_2)$-dimensional matrix $\mathbf{A}$, denote the L_F and L_2 norms of $\mathbf{A}$ as $\|\mathbf{A}\|_F = \{\mathrm{tr}(\mathbf{A}\mathbf{A}^\top)\}^{1/2}$ and $\|\mathbf{A}\|_2 = \lambda_{\max}^{1/2}(\mathbf{A}^\top\mathbf{A})$, respectively. Here $\lambda_{\max}(\mathbf{A}^\top\mathbf{A})$ is the maximal eigenvalue of $\mathbf{A}^\top\mathbf{A}$. We use the notation $\lceil x \rceil$ to denote the least integer function greater than or equal to $x \in \mathbb{R}$. For simplicity, we denote $\breve{\beta} = \beta - \widehat{\beta}$, $\breve{\boldsymbol{\alpha}}_x = \boldsymbol{\alpha}_x - \widehat{\boldsymbol{\alpha}}_x$, $\boldsymbol{\Phi} = (\boldsymbol{\Phi}_{\mathbf{X}}^\top, \boldsymbol{\Phi}_{\mathbf{Z}}^\top)^\top$ (defined in Assumption 2).

A.1 Proof of Proposition 1

Since the L_2 loss satisfies the restricted strong convexity condition (Negahban et al., 2012), from Theorem 1 in Loh and Wainwright (2015), it suffices to investigate the order of

$$\left\|\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i(Y_i - D_i\widehat{\beta} - \mathbf{X}_i^\top\boldsymbol{\alpha}_x)\right\|_\infty,$$

to obtain the convergence rate of $\widehat{\boldsymbol{\alpha}}_x$. Notice that under $\mathbb{H}_0$,

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i(Y_i - D_i\widehat{\beta} - \mathbf{X}_i^\top\boldsymbol{\alpha}_x)\\
=&\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i(\epsilon_i + D_i\beta - D_i\widehat{\beta})\\
=&\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i\epsilon_i + \breve{\beta}\frac{1}{n}\sum_{i=1}^{n}\{\mathbf{X}_iD_i - \mathbb{E}(\mathbf{X}D)\} + \breve{\beta}\mathbb{E}(\mathbf{X}D).
\end{aligned}$$

Let X_{ij} be the j-th element of $\mathbf{X}_i$, $j = 1, \ldots, p_x$. From Assumptions 2-3, we know that $X_{ij}\epsilon_i$, $X_{ij}D_i - \mathbb{E}(X_jD)$ for $j = 1, \ldots, p_x$ are all *sub-Exponential*.

Applying the Bernstein's inequality, we deduce that

$$\begin{aligned}
\left\|\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i\epsilon_i\right\|_\infty &= O_p\{n^{-1/2}(\log p_x)^{1/2}\},\\
\left\|\frac{1}{n}\sum_{i=1}^{n}\{\mathbf{X}_iD_i - \mathbb{E}(\mathbf{X}D)\}\right\|_\infty &= O_p\{n^{-1/2}(\log p_x)^{1/2}\}.
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\left\|\breve{\beta}\frac{1}{n}\sum_{i=1}^{n}\{\mathbf{X}_iD_i - \mathbb{E}(\mathbf{X}D)\}\right\|_\infty &= o_p\{n^{-1/2}(\log p_x)^{1/2}\},\\
\left\|\breve{\beta}\mathbb{E}(\mathbf{X}D)\right\|_\infty &= O_p(n^{-1/2}) = o_p\{n^{-1/2}(\log p_x)^{1/2}\}.
\end{aligned}$$

Therefore,

$$\left\|\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i(Y_i - D_i\widehat{\beta} - \mathbf{X}_i^\top\boldsymbol{\alpha}_x)\right\|_\infty = O_p\{n^{-1/2}(\log p_x)^{1/2}\}.$$

The proof is completed. □

A.2 Proof of Theorem 1

Under the null hypothesis $\boldsymbol{\alpha}_z = 0$, we decompose T_n into the sum of four terms, that is

$$\begin{aligned}
T_n &= \frac{1}{n}\sum_{i\neq j}(Y_i - D_i\widehat{\beta} - \mathbf{X}_i^\top\widehat{\boldsymbol{\alpha}}_x)(Y_j - D_j\widehat{\beta} - \mathbf{X}_j^\top\widehat{\boldsymbol{\alpha}}_x)\mathbf{Z}_i^\top\mathbf{Z}_j \\
&= \frac{1}{n}\sum_{i\neq j}(\epsilon_i + D_i\breve{\beta} + \mathbf{X}_i^\top\breve{\boldsymbol{\alpha}}_x)(\epsilon_j + D_j\breve{\beta} + \mathbf{X}_j^\top\breve{\boldsymbol{\alpha}}_x)\mathbf{Z}_i^\top\mathbf{Z}_j \\
&= \frac{1}{n}\sum_{i\neq j}\epsilon_i\epsilon_j\mathbf{Z}_i^\top\mathbf{Z}_j + \frac{1}{n}\sum_{i\neq j}(\epsilon_i D_j\breve{\beta} + D_i\breve{\beta}\epsilon_j)\mathbf{Z}_i^\top\mathbf{Z}_j \\
&\quad + \frac{1}{n}\sum_{i\neq j}(\epsilon_i\mathbf{X}_j^\top\breve{\boldsymbol{\alpha}}_x + \mathbf{X}_i^\top\breve{\boldsymbol{\alpha}}_x\epsilon_j)\mathbf{Z}_i^\top\mathbf{Z}_j \\
&\quad + \frac{1}{n}\sum_{i\neq j}(D_i\breve{\beta} + \mathbf{X}_i^\top\breve{\boldsymbol{\alpha}}_x)(D_j\breve{\beta} + \mathbf{X}_j^\top\breve{\boldsymbol{\alpha}}_x)\mathbf{Z}_i^\top\mathbf{Z}_j \\
&=: T_{n1} + T_{n2} + T_{n3} + T_{n4}.
\end{aligned}$$

Following the Martingale Central Limit Theorem (Hall and Heyde, 2014), we have

$$T_{n1}/\Omega^{1/2} \xrightarrow{d} N(0,1) \quad \text{as } (n, p_z) \to \infty. \tag{S1}$$

The detailed proof is given in Lemma S6. In what follows, we aim to prove that T_{n2}, T_{n3} and T_{n4} are $o_p(\Omega^{1/2})$ as $(n, p_x, p_z) \to \infty$.

For the term T_{n2}, we have

$$T_{n2} = \frac{1}{n}\sum_{i\neq j}(\epsilon_i D_j\breve{\beta} + D_i\breve{\beta}\epsilon_j)\mathbf{Z}_i^\top\mathbf{Z}_j = \breve{\beta}\frac{1}{n}\sum_{i\neq j}(\epsilon_i D_j + D_i\epsilon_j)\mathbf{Z}_i^\top\mathbf{Z}_j =: \breve{\beta}U_{n2}. \tag{S2}$$

Denote $u_n^{(2)} = \frac{1}{n-1}U_{n2} = \frac{1}{n(n-1)}\sum_{i\neq j}u_{ij}^{(2)}$ being a U-statisitc with the kernel $u_{ij}^{(2)} = (\epsilon_i D_j + D_i\epsilon_j)\mathbf{Z}_i^\top\mathbf{Z}_j$. We derive that

$$\mathbb{E}(u_n^{(2)}) = \mathbb{E}\{(\epsilon_i D_j + D_i\epsilon_j)\mathbf{Z}_i^\top\mathbf{Z}_j\} = 0,$$

and thus $\mathbb{E}(U_{n2}) = 0$. Further, the projection of $u_{ij}^{(2)}$ to the space $\{D_i, \mathbf{Z}_i, \epsilon_i\}$ is

$$\begin{aligned}
u_{1i}^{(2)} &= \mathbb{E}(u_{ij}^{(2)} \mid D_i, \mathbf{Z}_i, \epsilon_i) \\
&= \mathbb{E}\{(\epsilon_i D_j + D_i\epsilon_j)\mathbf{Z}_i^\top\mathbf{Z}_j \mid D_i, \mathbf{Z}_i, \epsilon_i\} = \epsilon_i\mathbf{Z}_i^\top\boldsymbol{\mu}_{D\mathbf{Z}}.
\end{aligned}$$

By the Hoeffding decomposition, the variance of $u_n^{(2)}$ is

$$\mathrm{Var}(u_n^{(2)}) = \frac{4(n-2)}{n(n-1)}\mathrm{Var}(u_{1i}^{(2)}) + \frac{2}{n(n-1)}\mathrm{Var}(u_{ij}^{(2)}),$$

and then

$$\begin{aligned}\mathrm{Var}(U_{n2}) &= \frac{4(n-1)(n-2)}{n}\mathrm{Var}(u_{1i}^{(2)}) + \frac{2(n-1)}{n}\mathrm{Var}(u_{ij}^{(2)}) \\ &\asymp n\mathrm{Var}(u_{1i}^{(2)}) + \mathrm{Var}(u_{ij}^{(2)}). \end{aligned} \tag{S3}$$

Here, we derive that

$$\begin{aligned} n\mathrm{Var}(u_{1i}^{(2)}) &= n\mathrm{Var}(\epsilon_i \mathbf{Z}_i^\top \boldsymbol{\mu}_{D\mathbf{Z}}) \\ &= n\mathbb{E}\{(\epsilon_i \mathbf{Z}_i^\top \boldsymbol{\mu}_{D\mathbf{Z}})^2\} \\ &= n\mathbb{E}(\epsilon^2)\mathbb{E}\{(\mathbf{Z}^\top \boldsymbol{\mu}_{D\mathbf{Z}})^2\} \\ &\lesssim n\boldsymbol{\mu}_{D\mathbf{Z}}^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\mu}_{D\mathbf{Z}}, \end{aligned} \tag{S4}$$

where the third equality holds based on the random error ϵ is independent of $\mathbf{Z}$, and the first inequality follows from Assumption 3. Moreover, we can deduce that

$$\begin{aligned} \mathrm{Var}(u_{ij}^{(2)}) &= \mathrm{Var}\{(\epsilon_i D_j + D_i \epsilon_j)\mathbf{Z}_i^\top \mathbf{Z}_j\} \\ &= \mathbb{E}\{(\epsilon_i D_j + D_i \epsilon_j)^2 (\mathbf{Z}_i^\top \mathbf{Z}_j)^2\} \\ &\le \mathbb{E}^{1/2}\{(\epsilon_i D_j + D_i \epsilon_j)^4\}\mathbb{E}^{1/2}\{(\mathbf{Z}_i^\top \mathbf{Z}_j)^4\} \lesssim \mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2). \end{aligned} \tag{S5}$$

Here $D_i = \mathbf{X}_i^\top \boldsymbol{\gamma}_x + \mathbf{Z}_i^\top \boldsymbol{\gamma}_z + \delta_i$, the first inequality follows from the Cauchy-Schwarz inequality, and the last inequality holds by Assumptions 2-3 and Lemma S2 with $b = 2$.

Combining the equations (S3), (S4) and (S5), it then follows that

$$\mathrm{Var}(U_{n2}) \lesssim n\boldsymbol{\mu}_{D\mathbf{Z}}^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\mu}_{D\mathbf{Z}} + \mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2) \lesssim \mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2),$$

where the last inequality holds by the equation (S62) in Lemma S7. As a result, we conclude that $U_{n2} = O_p\{\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\}$. Applying Assumption 6 and (S2), we deduce that

$$T_{n2} = O_p\{(s_{\boldsymbol{\gamma}} \log p/n)^{1/2}\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\}.$$

Following Assumption 3, we thus conclude that

$$T_{n2} = o_p(\Omega^{1/2}), \tag{S6}$$

when $n^{-1/2}s_{\boldsymbol{\gamma}}^{1/2}(\log p)^{1/2} = o(1)$ holds.

For the term T_{n3}, by the Hölder's inequality, we have

$$\begin{aligned} T_{n3} &= \frac{1}{n}\sum_{i\neq j}(\epsilon_i \mathbf{X}_j^\top \breve{\boldsymbol{\alpha}}_x + \mathbf{X}_i^\top \breve{\boldsymbol{\alpha}}_x \epsilon_j)\mathbf{Z}_i^\top \mathbf{Z}_j \\ &\le \left\| \frac{1}{n}\sum_{i\neq j}(\epsilon_i \mathbf{X}_j + \mathbf{X}_i \epsilon_j)\mathbf{Z}_i^\top \mathbf{Z}_j \right\|_\infty \|\breve{\boldsymbol{\alpha}}_x\|_1 =: \|\mathbf{U}_{n3}\|_\infty \|\breve{\boldsymbol{\alpha}}_x\|_1. \end{aligned}$$

We denote $\mathbf{u}_n^{(3)} = \frac{1}{n-1}\mathbf{U}_{n3} = \frac{1}{n(n-1)}\sum_{i\neq j}(\epsilon_i \mathbf{X}_j + \mathbf{X}_i\epsilon_j)\mathbf{Z}_i^\top \mathbf{Z}_j$, which is a p_x-dimensional vector with k-th element

$$u_{nk}^{(3)} = \frac{1}{n(n-1)}\sum_{i\neq j}(\epsilon_i X_{jk} + X_{ik}\epsilon_j)\mathbf{Z}_i^\top \mathbf{Z}_j =: \frac{1}{n(n-1)}\sum_{i\neq j} u_{ijk}^{(3)}.$$

Note that $u_{nk}^{(3)}$ is U-statistic with the kernel $u_{ijk}^{(3)} = (\epsilon_i X_{jk} + X_{ik}\epsilon_j)\mathbf{Z}_i^\top \mathbf{Z}_j$ for $k = 1, \ldots, p_x$.

We first calculate that $\mathbb{E}(u_{nk}^{(3)}) = \mathbb{E}\{(\epsilon_i X_{jk} + X_{ik}\epsilon_j)\mathbf{Z}_i^\top \mathbf{Z}_j\} = 0$. By Hoeffding decomposition, we derive that

$$\begin{aligned}
u_{nk}^{(3)} &= 2S_{1nk}^{(3)} + S_{2nk}^{(3)},\\
S_{1nk}^{(3)} &= \frac{1}{n}\sum_{i=1}^{n} g_{1k}^{(3)}(\mathbf{X}_i, \mathbf{Z}_i, \epsilon_i),\\
S_{2nk}^{(3)} &= \frac{1}{n(n-1)}\sum_{i\neq j} g_{2k}^{(3)}(\mathbf{X}_i, \mathbf{Z}_i, \epsilon_i; \mathbf{X}_j, \mathbf{Z}_j, \epsilon_j),
\end{aligned}$$

with

$$\begin{aligned}
g_{1k}^{(3)}(\mathbf{X}_i, \mathbf{Z}_i, \epsilon_i) &= \mathbb{E}(u_{ijk}^{(3)} \mid X_{ik}, \mathbf{Z}_i, \epsilon_i)\\
&= \mathbb{E}\{(\epsilon_i X_{jk} + X_{ik}\epsilon_j)\mathbf{Z}_i^\top \mathbf{Z}_j \mid X_{ik}, \mathbf{Z}_i, \epsilon_i\} = \epsilon_i \mathbf{Z}_i^\top \mathbb{E}(X_k \mathbf{Z}),
\end{aligned}$$

and

$$\begin{aligned}
g_{2k}^{(3)}(\mathbf{X}_i, \mathbf{Z}_i, \epsilon_i; \mathbf{X}_j, \mathbf{Z}_j, \epsilon_j) &= u_{ijk}^{(3)} - g_{1k}^{(3)}(X_{ik}, \mathbf{Z}_i, \epsilon_i) - g_{1k}^{(3)}(X_{jk}, \mathbf{Z}_j, \epsilon_j)\\
&= (\epsilon_i X_{jk} + X_{ik}\epsilon_j)\mathbf{Z}_i^\top \mathbf{Z}_j - \epsilon_i \mathbf{Z}_i^\top \mathbb{E}(X_k \mathbf{Z}) - \epsilon_j \mathbf{Z}_j^\top \mathbb{E}(X_k \mathbf{Z})\\
&= \epsilon_i \mathbf{Z}_i^\top \{X_{jk}\mathbf{Z}_j - \mathbb{E}(X_k \mathbf{Z})\} + \epsilon_j \mathbf{Z}_j^\top \{X_{ik}\mathbf{Z}_i - \mathbb{E}(X_k \mathbf{Z})\}.
\end{aligned}$$

Furthermore, we obtain that

$$\begin{aligned}
&\| \max_{1\le i\le n}\max_{1\le k\le p_x} g_{1k}^{(3)}(\mathbf{X}_i, \mathbf{Z}_i, \epsilon_i)\|_2\\
&= \| \max_{1\le i\le n}\max_{1\le k\le p_x} \epsilon_i \mathbf{Z}_i^\top \mathbb{E}(X_k \mathbf{Z})\|_2\\
&\lesssim \log(np_x) \max_{1\le i\le n}\max_{1\le k\le p_x} \|\epsilon_i \mathbf{Z}_i^\top \mathbb{E}(X_k \mathbf{Z})\|_{\psi_1}\\
&\le \log(np_x) \max_{1\le i\le n}\max_{1\le k\le p_x} \|\epsilon_i\|_{\psi_2} \|\mathbf{Z}_i^\top \mathbb{E}(X_k \mathbf{Z})\|_{\psi_2}\\
&\lesssim \log(np_x) \max_{1\le k\le p_x} \{\mathbb{E}(X_k \mathbf{Z})^\top \mathbf{\Sigma}_{\mathbf{Z}} \mathbb{E}(X_k \mathbf{Z})\}^{1/2}\\
&\lesssim \log(np_x) \max_{1\le k\le p_x} \{\mathbb{E}(X_k \mathbf{Z})^\top \mathbb{E}(X_k \mathbf{Z})\}^{1/2} \lambda_{\max}^{1/2}(\mathbf{\Sigma}_{\mathbf{Z}})\\
&\lesssim \log(np_x)\kappa \lambda_{\max}^{1/2}(\mathbf{\Sigma}_{\mathbf{Z}}),
\end{aligned} \tag{S7}$$

where the first inequality follows from the inequality (S49) in Lemma S4. The second and third inequalities hold by the fact that the product of two *sub-Guassian* random variables is *sub-Exponential* variable, ϵ is a *sub-Gaussian* random variable with bounded *sub-Guassian* norm in Assumption 3 and $\mathbf{Z}^\top\mathbb{E}(X_k\mathbf{Z})$ is a *sub-Gaussian* random variable with

$$\begin{aligned}\|\mathbf{Z}^\top\mathbb{E}(X_k\mathbf{Z})\|_{\psi_2} &= \|\mathbb{E}(X_k\mathbf{Z})^\top\boldsymbol{\Phi}_{\mathbf{Z}}\boldsymbol{\nu}\|_{\psi_2}\\ &\lesssim \|\mathbb{E}(X_k\mathbf{Z})^\top\boldsymbol{\Phi}_{\mathbf{Z}}\|_2\|\boldsymbol{\nu}\|_{\psi_2} \lesssim \{\mathbb{E}(X_k\mathbf{Z})^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\mathbb{E}(X_k\mathbf{Z})\}^{1/2},\end{aligned}$$

which is obtained by Lemma S5. Similarly, we derive that

$$\begin{aligned}\max_{1\le k\le p_x}\|g_{1k}^{(3)}(\mathbf{X}_i,\mathbf{Z}_i,\epsilon_i)\|_2 &= \max_{1\le k\le p_x}\|\epsilon_i\mathbf{Z}_i^\top\mathbb{E}(X_k\mathbf{Z})\|_2\\ &\lesssim \max_{1\le k\le p_x}\|\epsilon_i\mathbf{Z}_i^\top\mathbb{E}(X_k\mathbf{Z})\|_{\psi_1} \lesssim \kappa\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}).\end{aligned}\tag{S8}$$

Here, the first inequality holds by the equation (S48) in Lemma S4, and the second inequality has been derived in the equation (S7).

Taking $q=\lceil 4/(1-3a)\rceil$ for some constant $0<a<1/3$, we calculate that

$$\begin{aligned}&\|\max_{1\le i\ne j\le n}\max_{1\le k\le p_x} g_{2k}^{(3)}(\mathbf{X}_i,\mathbf{Z}_i,\epsilon_i;\mathbf{X}_j,\mathbf{Z}_j,\epsilon_j)\|_4\\ \le{}& \|\max_{1\le i\ne j\le n}\max_{1\le k\le p_x} g_{2k}^{(3)}(\mathbf{X}_i,\mathbf{Z}_i,\epsilon_i;\mathbf{X}_j,\mathbf{Z}_j,\epsilon_j)\|_q\\ \lesssim{}& \|\max_{1\le i\ne j\le n}\max_{1\le k\le p_x}\epsilon_i\mathbf{Z}_i^\top X_{jk}\mathbf{Z}_j\|_q\\ \lesssim{}& \|\max_{1\le i\ne j\le n}\max_{1\le k\le p_x}\epsilon_i X_{jk}\|_{2q}\cdot\|\max_{1\le i\ne j\le n}\mathbf{Z}_i^\top\mathbf{Z}_j\|_{2q}\\ \lesssim{}& \log(np_x)\max_{1\le i\ne j\le n}\max_{1\le k\le p_x}\|\epsilon_i X_{jk}\|_{\psi_1}\cdot\{n(n-1)\}^{1/2q}\max_{1\le i\ne j\le n}\|\mathbf{Z}_i^\top\mathbf{Z}_j\|_{2q}\\ \lesssim{}& \log(np_x)n^{1/q}\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2),\end{aligned}\tag{S9}$$

where the first inequality follows from the Lyapunov inequality with $q\ge 4$, the third inequality is obtained by the Cauchy-Schwartz inequality, the fourth inequality is due to Lemma S4, and the last inequality holds by Assumptions 2-3 and Lemma S2. Similarly, it follows that

$$\begin{aligned}\max_{1\le k\le p_x}\|g_{2k}^{(3)}(\mathbf{X}_i,\mathbf{Z}_i,\epsilon_i;\mathbf{X}_j,\mathbf{Z}_j,\epsilon_j)\|_2 &\le \max_{1\le k\le p_x}\|g_{2k}^{(3)}(\mathbf{X}_i,\mathbf{Z}_i,\epsilon_i;\mathbf{X}_j,\mathbf{Z}_j,\epsilon_j)\|_4\\ &\lesssim \max_{1\le k\le p_x}\|\epsilon_i\mathbf{Z}_i^\top X_{jk}\mathbf{Z}_j\|_4\\ &\lesssim \max_{1\le k\le p_x}\|\epsilon_i X_{jk}\|_8\cdot\|\mathbf{Z}_i^\top\mathbf{Z}_j\|_8\\ &\lesssim \mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2).\end{aligned}\tag{S10}$$

Combining the equations (S7)-(S10) and according to Lemma S3, we have that

$$\begin{aligned}\mathbb{E}(\|\mathbf{U}_{n3}\|_\infty) &\asymp n\mathbb{E}(\|\mathbf{u}_n^{(3)}\|_\infty)\\ &\lesssim \kappa\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}})\{n^{1/2}(\log p_x)^{1/2}+\log p_x\log(np_x)\}\\ &\quad+\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\{\log p_x+n^{-1/2+1/q}(\log p_x)^{3/2}\log(np_x)\}\\ &\lesssim n^{1/2}(\log p_x)^{1/2}\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}})\kappa+\log p_x\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2),\end{aligned}\tag{S11}$$

where we calculate that $n^{1/2}(\log p_x)^{1/2}$ and $\log p_x$ dominate the bound under Assumption 4. Recall that $T_{n3} \le \|\mathbf{U}_{n3}\|_\infty \|\breve{\boldsymbol{\alpha}}_x\|_1$. By the equation (S11) and Proposition 1, we obtain that

$$\begin{aligned} T_{n3} &= O_p\{s_{\boldsymbol{\alpha}_x} n^{-1/2}(\log p_x)^{1/2}\} \cdot O_p\{n^{1/2}(\log p_x)^{1/2}\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}})\kappa + \log p_x \mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\} \\ &= O_p\{s_{\boldsymbol{\alpha}_x} \log p_x \lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}})\kappa + n^{-1/2} s_{\boldsymbol{\alpha}_x}(\log p_x)^{3/2}\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\} \\ &= o_p(\Omega^{1/2}), \end{aligned} \tag{S12}$$

where the first equality holds by Proposition 1, and we utilize our conditions

$$\begin{aligned} s_{\boldsymbol{\alpha}_x} \log p_x \lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}})\kappa &= o\{\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\}, \\ n^{-1/2} s_{\boldsymbol{\alpha}_x}(\log p_x)^{3/2} &= o(1), \end{aligned}$$

to obtain the last equality.

For the term T_{n4}, we can proceed T_{n4} as

$$\begin{aligned} T_{n4} &= \frac{1}{n}\sum_{i\neq j}(D_i\breve{\beta} + \mathbf{X}_i^\top \breve{\boldsymbol{\alpha}}_x)(D_j\breve{\beta} + \mathbf{X}_j^\top \breve{\boldsymbol{\alpha}}_x)\mathbf{Z}_i^\top \mathbf{Z}_j \\ &= \frac{1}{n}\sum_{i\neq j} D_i\breve{\beta} D_j\breve{\beta}\mathbf{Z}_i^\top \mathbf{Z}_j + \frac{1}{n}\sum_{i\neq j}\mathbf{X}_i^\top \breve{\boldsymbol{\alpha}}_x \mathbf{X}_j^\top \breve{\boldsymbol{\alpha}}_x \mathbf{Z}_i^\top \mathbf{Z}_j \\ &\quad + \frac{1}{n}\sum_{i\neq j}(D_i\breve{\beta}\mathbf{X}_j^\top \breve{\boldsymbol{\alpha}}_x + \mathbf{X}_i^\top \breve{\boldsymbol{\alpha}}_x D_j\breve{\beta})\mathbf{Z}_i^\top \mathbf{Z}_j \\ &=: T_{n4a} + T_{n4b} + T_{n4c}. \end{aligned}$$

For the term T_{n4a}, we know that

$$T_{n4a} = \frac{1}{n}\sum_{i\neq j} D_i\breve{\beta} D_j\breve{\beta}\mathbf{Z}_i^\top \mathbf{Z}_j = \breve{\beta}^2 \frac{1}{n}\sum_{i\neq j} D_i D_j \mathbf{Z}_i^\top \mathbf{Z}_j =: \breve{\beta}^2 U_{n4a}.$$

Similar to the derivation of the term T_{n2}, we denote $u_n^{(4a)} = \frac{1}{n-1}U_{n4a} = \frac{1}{n(n-1)}\sum_{i\neq j} u_{ij}^{(4a)}$ being a U-statistic with the kernel $u_{ij}^{(4a)} = D_i D_j \mathbf{Z}_i^\top \mathbf{Z}_j$. By Assumption 5, We yield that

$$\begin{aligned} \mathbb{E}(U_{n4a}) \asymp n\mathbb{E}(u_n^{(4a)}) &= n\mathbb{E}(D_i D_j \mathbf{Z}_i^\top \mathbf{Z}_j) \\ &= n\mathbb{E}(D_i\mathbf{Z}_i)^\top \mathbb{E}(D_j \mathbf{Z}_j) \\ &= n\boldsymbol{\mu}_{D\mathbf{Z}}^\top \boldsymbol{\mu}_{D\mathbf{Z}} = O\{\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\}. \end{aligned} \tag{S13}$$

We also yield that the projection of $u_{ij}^{(4a)}$ to the space $\{D_i, \mathbf{Z}_i\}$ is

$$u_{1i}^{(4a)} = \mathbb{E}(u_{ij}^{(4a)} \mid D_i, \mathbf{Z}_i) = \mathbb{E}(D_i D_j \mathbf{Z}_i^\top \mathbf{Z}_j \mid D_i, \mathbf{Z}_i) = D_i \mathbf{Z}_i^\top \boldsymbol{\mu}_{D\mathbf{Z}}.$$

By the Hoeffding decomposition and similar to the equation (S3), the variance of U_{n4a} is

$$\begin{aligned} \mathrm{Var}(U_{n4a}) &\asymp n\mathrm{Var}(u_{1i}^{(4a)}) + \mathrm{Var}(u_{ij}^{(4a)}) \\ &\lesssim n\mathbb{E}\{(D_i\mathbf{Z}_i^\top \boldsymbol{\mu}_{D\mathbf{Z}})^2\} + \mathbb{E}\{(D_i D_j \mathbf{Z}_i^\top \mathbf{Z}_j)^2\} \\ &\lesssim n\mathbb{E}^{1/2}(D_i^4)\mathbb{E}^{1/2}\{(\mathbf{Z}_i^\top \boldsymbol{\mu}_{D\mathbf{Z}})^4\} + \mathbb{E}^{1/2}(D_i^4)\mathbb{E}^{1/2}(D_j^4)\mathbb{E}^{1/2}\{(\mathbf{Z}_i^\top \mathbf{Z}_j)^4\} \\ &\lesssim n\mathbb{E}^{1/2}\{(\mathbf{Z}_i^\top \boldsymbol{\mu}_{D\mathbf{Z}})^4\} + \mathbb{E}^{1/2}\{(\mathbf{Z}_i^\top \mathbf{Z}_j)^4\} \\ &\lesssim n\boldsymbol{\mu}_{D\mathbf{Z}}^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\mu}_{D\mathbf{Z}} + \mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2) \lesssim \mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2), \end{aligned} \tag{S14}$$

where the second inequality is due to the Cauchy-Schwarz inequality, the third inequality holds by Assumption 2-3 with $D = \mathbf{X}^\top \boldsymbol{\gamma}_x + \mathbf{Z}^\top \boldsymbol{\gamma}_z + \delta$, the fourth inequality follows from the equation (S63) in Lemma S7 together with Lemma S2, and the last inequality is obtained by the equation (S62) in Lemma S7.

Therefore, we have $U_{n4a} = O_p\{\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\}$ combining the equations (S13) and (S14). Further, with Assumption 6, we conclude that

$$T_{n4a} = \breve{\beta}^2 U_{n4a} = O_p(s_{\boldsymbol{\gamma}} \log p/n) \cdot O_p\{\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\} = o_p(\Omega^{1/2}), \tag{S15}$$

where the last equality holds by the condition $n^{-1/2} s_{\boldsymbol{\gamma}}^{1/2} (\log p)^{1/2} = o(1)$.

For the term T_{n4b}, we utilize the Hölder's inequality to obtain that

$$\begin{aligned} T_{n4b} &= \frac{1}{n} \sum_{i \neq j} \mathbf{X}_i^\top \breve{\boldsymbol{\alpha}}_x \mathbf{X}_j^\top \breve{\boldsymbol{\alpha}}_x \mathbf{Z}_i^\top \mathbf{Z}_j \\ &\leq \left\| \frac{1}{n} \sum_{i \neq j} \mathbf{X}_i \mathbf{X}_j^\top \mathbf{Z}_i^\top \mathbf{Z}_j \right\|_\infty \|\breve{\boldsymbol{\alpha}}_x\|_1^2 =: \|\mathbf{U}_{n4b}\|_\infty \|\breve{\boldsymbol{\alpha}}_x\|_1^2. \end{aligned}$$

We further denote $\mathbf{u}_n^{(4b)} = \frac{1}{n-1} \mathbf{U}_{n4b} = \frac{1}{n(n-1)} \sum_{i \neq j} \mathbf{X}_i \mathbf{X}_j^\top \mathbf{Z}_i^\top \mathbf{Z}_j$, which is a $p_x \times p_x$-dimensional matrix with (k_1, k_2)-th element

$$u_{nk_1k_2}^{(4b)} = \frac{1}{n(n-1)} \sum_{i \neq j} \frac{1}{2} (X_{ik_1} X_{jk_2} + X_{ik_2} X_{jk_1}) \mathbf{Z}_i^\top \mathbf{Z}_j =: \frac{1}{n(n-1)} \sum_{i \neq j} u_{ijk_1k_2}^{(4b)}.$$

Notice that $u_{nk_1k_2}^{(4b)}$ is U-statistic with the kernel $u_{ijk_1k_2}^{(4b)} = (X_{ik_1} X_{jk_2} + X_{ik_2} X_{jk_1}) \mathbf{Z}_i^\top \mathbf{Z}_j / 2$ for $k_1, k_2 = 1, \ldots, p_x$.

Similar to the derivation of the term T_{n3}, we first deduce that

$$\mathbb{E}(u_{nk_1k_2}^{(4b)}) = \mathbb{E}\{(X_{ik_1} X_{jk_2} + X_{ik_2} X_{jk_1}) \mathbf{Z}_i^\top \mathbf{Z}_j / 2\} = \mathbb{E}(X_{k_1} \mathbf{Z})^\top \mathbb{E}(X_{k_2} \mathbf{Z}).$$

By Hoeffding decomposition, we then yield that

$$\begin{aligned} u_{nk_1k_2}^{(4b)} &= \mathbb{E}(X_{k_1} \mathbf{Z})^\top \mathbb{E}(X_{k_2} \mathbf{Z}) + 2 S_{1nk_1k_2}^{(4b)} + S_{2nk_1k_2}^{(4b)}, \\ S_{1nk_1k_2}^{(4b)} &= \frac{1}{n} \sum_{i=1}^n g_{1k_1k_2}^{(4b)}(\mathbf{X}_i, \mathbf{Z}_i), \\ S_{2nk_1k_2}^{(4b)} &= \frac{1}{n(n-1)} \sum_{i \neq j} g_{2k_1k_2}^{(4b)}(\mathbf{X}_i, \mathbf{Z}_i; \mathbf{X}_j, \mathbf{Z}_j), \end{aligned}$$

with

$$\begin{aligned} & g_{1k_1k_2}^{(4b)}(\mathbf{X}_i, \mathbf{Z}_i) \\ &= \mathbb{E}(u_{ijk_1k_2}^{(4b)} \mid X_{ik}, \mathbf{Z}_i) - \mathbb{E}(X_{k_1} \mathbf{Z})^\top \mathbb{E}(X_{k_2} \mathbf{Z}) \\ &= \mathbb{E}\{(X_{ik_1} X_{jk_2} + X_{ik_2} X_{jk_1}) \mathbf{Z}_i^\top \mathbf{Z}_j / 2 \mid X_{ik}, \mathbf{Z}_i\} - \mathbb{E}(X_{k_1} \mathbf{Z})^\top \mathbb{E}(X_{k_2} \mathbf{Z}) \\ &= \{X_{ik_1} \mathbf{Z}_i^\top \mathbb{E}(X_{k_2} \mathbf{Z}) + X_{ik_2} \mathbf{Z}_i^\top \mathbb{E}(X_{k_1} \mathbf{Z})\}/2 - \mathbb{E}(X_{k_1} \mathbf{Z})^\top \mathbb{E}(X_{k_2} \mathbf{Z}) \\ &= \{X_{ik_1} \mathbf{Z}_i - \mathbb{E}(X_{k_1} \mathbf{Z})\}^\top \mathbb{E}(X_{k_2} \mathbf{Z})/2 + \{X_{ik_2} \mathbf{Z}_i - \mathbb{E}(X_{k_2} \mathbf{Z})\}^\top \mathbb{E}(X_{k_1} \mathbf{Z})/2, \end{aligned}$$

and

$$
\begin{aligned}
& g^{(4b)}_{2k_1k_2}(\mathbf{X}_i,\mathbf{Z}_i;\mathbf{X}_j,\mathbf{Z}_j) \\
&= u^{(4b)}_{ijk_1k_2} - \mathbb{E}(X_{k_1}\mathbf{Z})^\top\mathbb{E}(X_{k_2}\mathbf{Z}) - g^{(4b)}_{1k_1k_2}(\mathbf{X}_i,\mathbf{Z}_i) - g^{(4b)}_{1k_1k_2}(\mathbf{X}_j,\mathbf{Z}_j) \\
&= (X_{ik_1}X_{jk_2} + X_{ik_2}X_{jk_1})\mathbf{Z}_i^\top\mathbf{Z}_j/2 - \{X_{ik_1}\mathbf{Z}_i^\top\mathbb{E}(X_{k_2}\mathbf{Z}) + X_{ik_2}\mathbf{Z}_i^\top\mathbb{E}(X_{k_1}\mathbf{Z})\}/2 \\
&\quad - \{X_{jk_1}\mathbf{Z}_j^\top\mathbb{E}(X_{k_2}\mathbf{Z}) + X_{jk_2}\mathbf{Z}_j^\top\mathbb{E}(X_{k_1}\mathbf{Z})\}/2 + \mathbb{E}(X_{k_1}\mathbf{Z})^\top\mathbb{E}(X_{k_2}\mathbf{Z}) \\
&= \{X_{jk_2}\mathbf{Z}_j - \mathbb{E}(X_{k_2}\mathbf{Z})\}^\top\{X_{ik_1}\mathbf{Z}_j - \mathbb{E}(X_{k_1}\mathbf{Z})\}/2 \\
&\quad + \{X_{jk_1}\mathbf{Z}_j - \mathbb{E}(X_{k_1}\mathbf{Z})\}^\top\{X_{ik_2}\mathbf{Z}_i - \mathbb{E}(X_{k_2}\mathbf{Z})\}/2.
\end{aligned}
$$

Moreover, we calculate that

$$
\begin{aligned}
& \| \max_{1\le i\le n}\max_{1\le k_1,k_2\le p_x} g^{(4b)}_{1k_1k_2}(\mathbf{X}_i,\mathbf{Z}_i)\|_2 \\
&\lesssim \log(np_x)\max_{1\le i\le n}\max_{1\le k_1,k_2\le p_x}\|g^{(4b)}_{1k_1k_2}(\mathbf{X}_i,\mathbf{Z}_i)\|_{\psi_1} \\
&\lesssim \log(np_x)\max_{1\le i\le n}\max_{1\le k_1,k_2\le p_x}\|X_{ik_1}\mathbf{Z}_i^\top\mathbb{E}(X_{k_2}\mathbf{Z})\|_{\psi_1} \\
&\le \log(np_x)\max_{1\le i\le n}\max_{1\le k_1,k_2\le p_x}\|X_{ik_1}\|_{\psi_2}\|\mathbf{Z}_i^\top\mathbb{E}(X_{k_2}\mathbf{Z})\|_{\psi_2} \\
&\lesssim \log(np_x)\max_{1\le k_1,k_2\le p_x}\{\mathbb{E}(X_{k_2}\mathbf{Z})^\top\mathbf{\Sigma}_{\mathbf{Z}}\mathbb{E}(X_{k_2}\mathbf{Z})\}^{1/2} \\
&\lesssim \log(np_x)\max_{1\le k_1,k_2\le p_x}\{\mathbb{E}(X_{k_2}\mathbf{Z})^\top\mathbb{E}(X_{k_2}\mathbf{Z})\}^{1/2}\lambda^{1/2}_{\max}(\mathbf{\Sigma}_{\mathbf{Z}}) \\
&\lesssim \log(np_x)\kappa\lambda^{1/2}_{\max}(\mathbf{\Sigma}_{\mathbf{Z}}),
\end{aligned}
\tag{S16}
$$

where the first inequality follows from the equation (S49) in Lemma S4, and we deduce the remaining inequalities by the similar techniques used in the equation (S7) together with Assumption 2. Similarly, we derive that

$$
\begin{aligned}
\max_{1\le k_1,k_2\le p_x}\|g^{(4b)}_{1k_1k_2}(\mathbf{X}_i,\mathbf{Z}_i)\|_2 &\lesssim \max_{1\le k_1,k_2\le p_x}\|X_{ik_1}\mathbf{Z}_i^\top\mathbb{E}(X_{k_2}\mathbf{Z})\|_2 \\
&\lesssim \max_{1\le k_1,k_2\le p_x}\|X_{ik_1}\mathbf{Z}_i^\top\mathbb{E}(X_{k_2}\mathbf{Z})\|_{\psi_1} \lesssim \kappa\lambda^{1/2}_{\max}(\mathbf{\Sigma}_{\mathbf{Z}}).
\end{aligned}
\tag{S17}
$$

Here, the second inequality holds by the equation (S48) in Lemma S4, and the last inequality has been derived in the equation (S16).

Taking $q = \lceil 4/(1-3a)\rceil$ with $0 < a < 1/3$, we apply the similar derivation used in the equation (S9) to yield that

$$
\begin{aligned}
& \| \max_{1\le i\ne j\le n}\max_{1\le k_1,k_2\le p_x} g^{(4b)}_{2k_1k_2}(\mathbf{X}_i,\mathbf{Z}_i;\mathbf{X}_j,\mathbf{Z}_j)\|_4 \\
&\le \| \max_{1\le i\ne j\le n}\max_{1\le k_1,k_2\le p_x} g^{(4b)}_{2k_1k_2}(\mathbf{X}_i,\mathbf{Z}_i;\mathbf{X}_j,\mathbf{Z}_j)\|_q \\
&\lesssim \| \max_{1\le i\ne j\le n}\max_{1\le k_1,k_2\le p_x} X_{ik_1}X_{jk_2}\mathbf{Z}_i^\top\mathbf{Z}_j\|_q \\
&\lesssim \| \max_{1\le i\ne j\le n}\max_{1\le k_1,k_2\le p_x} X_{ik_1}X_{jk_2}\|_{2q}\cdot\| \max_{1\le i\ne j\le n}\mathbf{Z}_i^\top\mathbf{Z}_j\|_{2q}
\end{aligned}
\tag{S18}
$$

$$\lesssim \log(np_x) \max_{1\le i\neq j\le n} \max_{1\le k\le p_x} \|X_{ik_1}X_{jk_2}\|_{\psi_1} \cdot \{n(n-1)\}^{1/2q} \max_{1\le i\neq j\le n} \|\mathbf{Z}_i^\top \mathbf{Z}_j\|_{2q}$$
$$\lesssim \log(np_x) n^{1/q} \mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2),$$

where the first inequality follows from the Lyapunov inequality with $q \ge 4$, the third inequality is obtained by the Cauchy-Schwartz inequality, the fourth inequality holds based on Lemma S4, and the last inequality is due to Assumption 2 and Lemma S2. Similarly, it follows that

$$\begin{aligned}\max_{1\le k\le p_x} \|g^{(4b)}_{2k_1k_2}(\mathbf{X}_i,\mathbf{Z}_i;\mathbf{X}_j,\mathbf{Z}_j)\|_2 &\le \max_{1\le k\le p_x} \|g^{(4b)}_{2k_1k_2}(\mathbf{X}_i,\mathbf{Z}_i;\mathbf{X}_j,\mathbf{Z}_j)\|_4 \\ &\lesssim \max_{1\le k\le p_x} \|X_{ik_1}X_{jk_2}\mathbf{Z}_i^\top\mathbf{Z}_j\|_4 \\ &\lesssim \max_{1\le k\le p_x} \|X_{ik_1}X_{jk_2}\|_8 \cdot \|\mathbf{Z}_i^\top\mathbf{Z}_j\|_8 \\ &\lesssim \mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2).\end{aligned} \tag{S19}$$

Under Lemma S3 with the equations (S16)-(S19), we derive that

$$\begin{aligned}\mathbb{E}(\|\mathbf{U}_{n4b}\|_\infty) &\asymp n\mathbb{E}(\|\mathbf{u}_n^{(4b)}\|_\infty) \\ &\lesssim n\kappa^2 + \kappa\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}})\{n^{1/2}(\log p_x)^{1/2} + \log p_x \log(np_x)\} \\ &\quad + \mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\{\log p_x + n^{-1/2+1/q}(\log p_x)^{3/2}\log(np_x)\} \\ &\lesssim n\kappa^2 + n^{1/2}(\log p_x)^{1/2}\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}})\kappa + \log p_x \mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2).\end{aligned} \tag{S20}$$

Here, we obtain that $n^{1/2}(\log p_x)^{1/2}$ and $\log p_x$ dominate the bound by Assumption 4. Recall that $T_{n4b} \le \|\mathbf{U}_{n4b}\|_\infty \|\breve{\boldsymbol{\alpha}}_x\|_1^2$. By the equation (S20) and Proposition 1, we obtain that

$$\begin{aligned}&T_{n4b} \\ &= O_p\{n^{-1}s^2_{\boldsymbol{\alpha}_x}\log p_x\} \cdot O_p\{n\kappa^2 + n^{1/2}(\log p_x)^{1/2}\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}})\kappa + \log p_x \mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\} \\ &= O_p\{s^2_{\boldsymbol{\alpha}_x}\log p_x\kappa^2 + n^{-1/2}s^2_{\boldsymbol{\alpha}_x}(\log p_x)^{3/2}\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}})\kappa + n^{-1}s^2_{\boldsymbol{\alpha}_x}(\log p_x)^2\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\} \\ &= o_p(\Omega^{1/2}),\end{aligned} \tag{S21}$$

where the first equality holds by Proposition 1, and we utilize our conditions

$$\begin{aligned}s^2_{\boldsymbol{\alpha}_x}\log p_x\kappa^2 &= o\{\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\}, \\ n^{-1/2}s_{\boldsymbol{\alpha}_x}(\log p_x)^{3/2} &= o(1), \\ s_{\boldsymbol{\alpha}_x}\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}})\kappa &= o\{\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\}, \\ n^{-1/2}s_{\boldsymbol{\alpha}_x}\log p_x &= o(1),\end{aligned}$$

to obtain the last equality.

For the term T_{n4c}, we can proceed it as

$$\begin{aligned}
T_{n4c} &= \frac{1}{n}\sum_{i\neq j}(D_i\breve{\beta}\mathbf{X}_j^\top\breve{\boldsymbol{\alpha}}_x + \mathbf{X}_i^\top\breve{\boldsymbol{\alpha}}_x D_j\breve{\beta})\mathbf{Z}_i^\top\mathbf{Z}_j \\
&= \breve{\beta}\frac{1}{n}\sum_{i\neq j}(D_i\mathbf{X}_j^\top\breve{\boldsymbol{\alpha}}_x + \mathbf{X}_i^\top\breve{\boldsymbol{\alpha}}_x D_j)\mathbf{Z}_i^\top\mathbf{Z}_j \\
&\leq \breve{\beta}\left\|\frac{1}{n}\sum_{i\neq j}(D_i\mathbf{X}_j + \mathbf{X}_i D_j)\mathbf{Z}_i^\top\mathbf{Z}_j\right\|_\infty \|\breve{\boldsymbol{\alpha}}_x\|_1 =: \breve{\beta}\|\mathbf{U}_{n4c}\|_\infty\|\breve{\boldsymbol{\alpha}}_x\|_1,
\end{aligned}$$

where the first inequality follows from the Hölder's inequality. Denote the p_x-dimensional vector $\mathbf{u}_n^{(4c)} = \frac{1}{n-1}\mathbf{U}_{n4c} = \frac{1}{n(n-1)}\sum_{i\neq j}(D_i\mathbf{X}_j + \mathbf{X}_i D_j)\mathbf{Z}_i^\top\mathbf{Z}_j$ with the k-th element

$$u_{nk}^{(4c)} = \frac{1}{n(n-1)}\sum_{i\neq j}(D_i X_{jk} + X_{ik}D_j)\mathbf{Z}_i^\top\mathbf{Z}_j =: \frac{1}{n(n-1)}\sum_{i\neq j}u_{ijk}^{(4c)}.$$

Note that $u_{nk}^{(4c)}$ is U-statistic with the kernel $u_{ijk}^{(4c)} = (D_i X_{jk} + X_{ik}D_j)\mathbf{Z}_i^\top\mathbf{Z}_j$ for $k = 1, \ldots, p_x$.

We know that

$$\mathbb{E}(u_{nk}^{(4c)}) = \mathbb{E}\{(D_i X_{jk} + X_{ik}D_j)\mathbf{Z}_i^\top\mathbf{Z}_j\} = 2\mathbb{E}(D\mathbf{Z})^\top\mathbb{E}(X_k\mathbf{Z}) = 2\boldsymbol{\mu}_{D\mathbf{Z}}^\top\mathbb{E}(X_k\mathbf{Z}).$$

Using Hoeffding decomposition, we derive that

$$\begin{aligned}
u_{nk}^{(4c)} &= 2\boldsymbol{\mu}_{D\mathbf{Z}}^\top\mathbb{E}(X_k\mathbf{Z}) + 2S_{1nk}^{(4c)} + S_{2nk}^{(4c)}, \\
S_{1nk}^{(4c)} &= \frac{1}{n}\sum_{i=1}^n g_{1k}^{(4c)}(D_i, \mathbf{X}_i, \mathbf{Z}_i), \\
S_{2nk}^{(4c)} &= \frac{1}{n(n-1)}\sum_{i\neq j}g_{2k}^{(4c)}(D_i, \mathbf{X}_i, \mathbf{Z}_i; D_j, \mathbf{X}_j, \mathbf{Z}_j),
\end{aligned}$$

with

$$\begin{aligned}
&g_{1k}^{(4c)}(D_i, \mathbf{X}_i, \mathbf{Z}_i) \\
&= \mathbb{E}(u_{ijk}^{(4c)} \mid D_i, \mathbf{X}_i, \mathbf{Z}_i) - 2\boldsymbol{\mu}_{D\mathbf{Z}}^\top\mathbb{E}(X_k\mathbf{Z}) \\
&= \mathbb{E}\{(D_i X_{jk} + X_{ik}D_j)\mathbf{Z}_i^\top\mathbf{Z}_j \mid D_i, \mathbf{X}_i, \mathbf{Z}_i\} - 2\boldsymbol{\mu}_{D\mathbf{Z}}^\top\mathbb{E}(X_k\mathbf{Z}) \\
&= D_i\mathbf{Z}_i^\top\mathbb{E}(X_k\mathbf{Z}) + X_{ik}\mathbf{Z}_i^\top\boldsymbol{\mu}_{D\mathbf{Z}} - 2\boldsymbol{\mu}_{D\mathbf{Z}}^\top\mathbb{E}(X_k\mathbf{Z}) \\
&= \{D_i\mathbf{Z}_i - \boldsymbol{\mu}_{D\mathbf{Z}}\}^\top\mathbb{E}(X_k\mathbf{Z}) + \{X_{ik}\mathbf{Z}_i - \mathbb{E}(X_k\mathbf{Z})\}^\top\boldsymbol{\mu}_{D\mathbf{Z}},
\end{aligned}$$

and

$$\begin{aligned}
&g_{2k}^{(4c)}(D_i, \mathbf{X}_i, \mathbf{Z}_i; D_j, \mathbf{X}_j, \mathbf{Z}_j) \\
&= u_{ijk}^{(4c)} - 2\boldsymbol{\mu}_{D\mathbf{Z}}^\top\mathbb{E}(X_k\mathbf{Z}) - g_{1k}^{(4c)}(D_i, \mathbf{X}_i, \mathbf{Z}_i) - g_{1k}^{(4c)}(D_j, \mathbf{X}_j, \mathbf{Z}_j) \\
&= (D_i X_{jk} + X_{ik}D_j)\mathbf{Z}_i^\top\mathbf{Z}_j - D_i\mathbf{Z}_i^\top\mathbb{E}(X_k\mathbf{Z}) - X_{ik}\mathbf{Z}_i^\top\boldsymbol{\mu}_{D\mathbf{Z}} \\
&\quad - D_j\mathbf{Z}_j^\top\mathbb{E}(X_k\mathbf{Z}) - X_{jk}\mathbf{Z}_j^\top\boldsymbol{\mu}_{D\mathbf{Z}} + 2\boldsymbol{\mu}_{D\mathbf{Z}}^\top\mathbb{E}(X_k\mathbf{Z}) \\
&= \{D_i\mathbf{Z}_i - \boldsymbol{\mu}_{D\mathbf{Z}}\}^\top\{X_{jk}\mathbf{Z}_j - \mathbb{E}(X_k\mathbf{Z})\} + \{D_j\mathbf{Z}_j - \boldsymbol{\mu}_{D\mathbf{Z}}\}^\top\{X_{ik}\mathbf{Z}_i - \mathbb{E}(X_k\mathbf{Z})\}.
\end{aligned}$$

To proceed, we know that

$$
\begin{aligned}
&\| \max_{1\le i\le n}\max_{1\le k\le p_x} g_{1k}^{(4c)}(D_i,\mathbf{X}_i,\mathbf{Z}_i)\|_2 \\
&= \| \max_{1\le i\le n}\max_{1\le k\le p_x} \{D_i\mathbf{Z}_i - \boldsymbol{\mu}_{D\mathbf{Z}}\}^\top \mathbb{E}(X_k\mathbf{Z}) + \{X_{ik}\mathbf{Z}_i - \mathbb{E}(X_k\mathbf{Z})\}^\top \boldsymbol{\mu}_{D\mathbf{Z}}\|_2 \\
&\le \| \max_{1\le i\le n}\max_{1\le k\le p_x} \{D_i\mathbf{Z}_i - \boldsymbol{\mu}_{D\mathbf{Z}}\}^\top \mathbb{E}(X_k\mathbf{Z})\|_2 \\
&\quad + \| \max_{1\le i\le n}\max_{1\le k\le p_x} \{X_{ik}\mathbf{Z}_i - \mathbb{E}(X_k\mathbf{Z})\}^\top \boldsymbol{\mu}_{D\mathbf{Z}}\|_2 \\
&\lesssim \log(np_x)\{\max_{1\le i\le n}\max_{1\le k\le p_x} \|D_i\mathbf{Z}_i^\top \mathbb{E}(X_k\mathbf{Z})\|_{\psi_1} + \max_{1\le i\le n}\max_{1\le k\le p_x}\|X_{ik}\mathbf{Z}_i^\top \boldsymbol{\mu}_{D\mathbf{Z}}\|_{\psi_1}\} \qquad \text{(S22)} \\
&\le \log(np_x)\{\max_{1\le i\le n}\max_{1\le k\le p_x} \|D_i\|_{\psi_2}\|\mathbf{Z}_i^\top \mathbb{E}(X_k\mathbf{Z})\|_{\psi_2} + \max_{1\le i\le n}\max_{1\le k\le p_x}\|X_{ik}\|_{\psi_2}\|\mathbf{Z}_i^\top \boldsymbol{\mu}_{D\mathbf{Z}}\|_{\psi_2}\} \\
&\lesssim \log(np_x)\max_{1\le k\le p_x}\{\mathbb{E}(X_k\mathbf{Z})^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\mathbb{E}(X_k\mathbf{Z})\}^{1/2} + \log(np_x)(\boldsymbol{\mu}_{D\mathbf{Z}}^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\mu}_{D\mathbf{Z}})^{1/2} \\
&\lesssim \log(np_x)\kappa\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}) + \log(np_x)(\boldsymbol{\mu}_{D\mathbf{Z}}^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\mu}_{D\mathbf{Z}})^{1/2},
\end{aligned}
$$

where the second inequality is due to the equation (S49) in Lemma S4, and we use the similar techniques used in the equation (S7) and Assumption 2-3 with $D = \mathbf{X}^\top\boldsymbol{\gamma}_x + \mathbf{Z}^\top\boldsymbol{\gamma}_z + \delta$ to obtain the remaining inequalities. Analogously, we yield that

$$
\begin{aligned}
\max_{1\le k\le p_x}\|g_{1k}^{(4c)}(D_i,\mathbf{X}_i,\mathbf{Z}_i)\|_2 &\le \max_{1\le k\le p_x}\|D_i\mathbf{Z}_i^\top \mathbb{E}(X_k\mathbf{Z})\|_2 + \max_{1\le k\le p_x}\|X_{ik}\mathbf{Z}_i^\top \boldsymbol{\mu}_{D\mathbf{Z}}\|_2 \\
&\lesssim \max_{1\le k\le p_x}\|D_i\mathbf{Z}_i^\top \mathbb{E}(X_k\mathbf{Z})\|_{\psi_1} + \max_{1\le k\le p_x}\|X_{ik}\mathbf{Z}_i^\top \boldsymbol{\mu}_{D\mathbf{Z}}\|_{\psi_1} \qquad \text{(S23)} \\
&\lesssim \kappa\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}) + (\boldsymbol{\mu}_{D\mathbf{Z}}^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\mu}_{D\mathbf{Z}})^{1/2}.
\end{aligned}
$$

Here, the second inequality holds by the equation (S48) in Lemma S4, and the last inequality has been derived in the equation (S22).

Considering $q = \lceil 4/(1-3a)\rceil$ with $0 < a < 1/3$, we employ a derivation analogous to that in the equations (S9) and (S10) to obtain that

$$
\begin{aligned}
&\| \max_{1\le i\ne j\le n}\max_{1\le k\le p_x} g_{2k}^{(4c)}(D_i,\mathbf{X}_i,\mathbf{Z}_i;D_j,\mathbf{X}_j,\mathbf{Z}_j)\|_4 \\
&\le \| \max_{1\le i\ne j\le n}\max_{1\le k\le p_x} g_{2k}^{(4c)}(D_i,\mathbf{X}_i,\mathbf{Z}_i;D_j,\mathbf{X}_j,\mathbf{Z}_j)\|_q \\
&\lesssim \| \max_{1\le i\ne j\le n}\max_{1\le k\le p_x} D_i\mathbf{Z}_i^\top X_{jk}\mathbf{Z}_j\|_q \\
&\lesssim \| \max_{1\le i\ne j\le n}\max_{1\le k\le p_x} D_iX_{jk}\|_{2q}\cdot\| \max_{1\le i\ne j\le n}\mathbf{Z}_i^\top\mathbf{Z}_j\|_{2q} \qquad \text{(S24)} \\
&\lesssim \log(np_x)\max_{1\le i\ne j\le n}\max_{1\le k\le p_x}\|D_iX_{jk}\|_{\psi_1}\cdot\{n(n-1)\}^{1/2q}\max_{1\le i\ne j\le n}\|\mathbf{Z}_i^\top\mathbf{Z}_j\|_{2q} \\
&\lesssim \log(np_x)n^{1/q}\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2),
\end{aligned}
$$

and

$$
\max_{1\le k\le p_x}\|g_{2k}^{(4c)}(D_i,\mathbf{X}_i,\mathbf{Z}_i;D_j,\mathbf{X}_j,\mathbf{Z}_j)\|_2 \le \max_{1\le k\le p_x}\|g_{2k}^{(4c)}(D_i,\mathbf{X}_i,\mathbf{Z}_i;D_j,\mathbf{X}_j,\mathbf{Z}_j)\|_4
$$

$$
\begin{aligned}
&\lesssim \max_{1\le k\le p_x} \|D_i \mathbf{Z}_i^\top X_{jk}\mathbf{Z}_j\|_4 \\
&\lesssim \max_{1\le k\le p_x} \|D_i X_{jk}\|_8 \cdot \|\mathbf{Z}_i^\top \mathbf{Z}_j\|_8 \qquad \text{(S25)}\\
&\lesssim \mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2).
\end{aligned}
$$

Combining the equations (S22)-(S25) and under Lemma S3, we deduce that

$$
\begin{aligned}
&\mathbb{E}(\|\mathbf{U}_{n4c}\|_\infty)\\
&\asymp n\mathbb{E}(\|\mathbf{u}_n^{(4c)}\|_\infty)\\
&\lesssim n \max_{1\le k\le p_x} \boldsymbol{\mu}_{D\mathbf{Z}}^\top \mathbb{E}(X_k\mathbf{Z}) + \kappa\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}})\{n^{1/2}(\log p_x)^{1/2} + \log p_x \log(np_x)\}\\
&\quad + (\boldsymbol{\mu}_{D\mathbf{Z}}^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\mu}_{D\mathbf{Z}})^{1/2}\{n^{1/2}(\log p_x)^{1/2} + \log p_x\log(np_x)\}\\
&\quad + \mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\{\log p_x + n^{-1/2+1/q}(\log p_x)^{3/2}\log(np_x)\} \qquad \text{(S26)}\\
&\lesssim n(\boldsymbol{\mu}_{D\mathbf{Z}}^\top\boldsymbol{\mu}_{D\mathbf{Z}})^{1/2} \max_{1\le k\le p_x}\{\mathbb{E}(X_k\mathbf{Z})^\top\mathbb{E}(X_k\mathbf{Z})\}^{1/2}\\
&\quad + n^{1/2}(\log p_x)^{1/2}\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}})\kappa + n^{1/2}(\log p_x)^{1/2}(\boldsymbol{\mu}_{D\mathbf{Z}}^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\mu}_{D\mathbf{Z}})^{1/2}\\
&\quad + \log p_x \mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\\
&\lesssim n^{1/2}\mathrm{tr}^{1/4}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\kappa + n^{1/2}(\log p_x)^{1/2}\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}})\kappa + \log p_x\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2),
\end{aligned}
$$

where the second inequality follows from the Cauchy-Schwarz inequality and Assumption 4 (implying that $n^{1/2}(\log p_x)^{1/2}$ and $\log p_x$ dominate the bound). The last inequality is due to Assumption 5 ($n\boldsymbol{\mu}_{D\mathbf{Z}}^\top\boldsymbol{\mu}_{D\mathbf{Z}} = O\{\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\}$) and the equation (S62) in Lemma S7. Recall that $T_{n4c} \le \breve{\beta}\|\mathbf{U}_{n4c}\|_\infty\|\breve{\boldsymbol{\alpha}}_x\|_1$. Referring to the equation (S26), Assumption 6 and Proposition 1, it follows that

$$
\begin{aligned}
T_{n4c} &= O_p\{n^{-1/2}(s_{\boldsymbol{\gamma}}\log p)^{1/2}\}\cdot O_p\{n^{-1/2}s_{\boldsymbol{\alpha}_x}(\log p_x)^{1/2}\}\\
&\quad \cdot O_p\{n^{1/2}\mathrm{tr}^{1/4}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\kappa + n^{1/2}(\log p_x)^{1/2}\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}})\kappa + \log p_x\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\}\\
&= O_p\{n^{-1/2}(s_{\boldsymbol{\gamma}}\log p)^{1/2}\}\cdot O_p\{s_{\boldsymbol{\alpha}_x}(\log p_x)^{1/2}\mathrm{tr}^{1/4}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\kappa + s_{\boldsymbol{\alpha}_x}\log p_x\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}})\kappa \qquad \text{(S27)}\\
&\qquad + n^{-1/2}s_{\boldsymbol{\alpha}_x}(\log p_x)^{3/2}\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\}\\
&= o_p(\Omega^{1/2}),
\end{aligned}
$$

where the last equality holds by <u>our conditions</u>

$$
\begin{aligned}
(s_{\boldsymbol{\gamma}}\log p/n)^{1/2} &= o(1),\\
s_{\boldsymbol{\alpha}_x}^2\log p_x\kappa^2 &= o\{\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\},\\
s_{\boldsymbol{\alpha}_x}\log p_x\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}})\kappa &= o\{\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\},\\
n^{-1/2}s_{\boldsymbol{\alpha}_x}(\log p_x)^{3/2} &= o(1).
\end{aligned}
$$

Combining the equations (S15), (S21) and (S27), we conclude that

$$
T_{n4} = o_p(\Omega^{1/2}). \qquad \text{(S28)}
$$

In sum, following the results (S1), (S6), (S12) and (S28), we verify that

$$T_n/\Omega^{1/2} \xrightarrow{d} N(0,1) \qquad \text{as } (n, p_x, p_z) \to \infty.$$

The proof is completed. $\square$

A.3 Proof of Theorem 2

Under the local alternatives with $Y_i = D_i\beta + \mathbf{X}_i^\top \boldsymbol{\alpha}_x + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z + \epsilon_i$, we decompose T_n as

$$\begin{aligned}
T_n &= \frac{1}{n}\sum_{i\neq j}(Y_i - D_i\widehat{\beta} - \mathbf{X}_i^\top \widehat{\boldsymbol{\alpha}}_x)(Y_j - D_j\widehat{\beta} - \mathbf{X}_j^\top \widehat{\boldsymbol{\alpha}}_x)\mathbf{Z}_i^\top \mathbf{Z}_j \\
&= \frac{1}{n}\sum_{i\neq j}(\epsilon_i + D_i\breve{\beta} + \mathbf{X}_i^\top \breve{\boldsymbol{\alpha}}_x + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z)(\epsilon_j + D_j\breve{\beta} + \mathbf{X}_j^\top \breve{\boldsymbol{\alpha}}_x + \mathbf{Z}_j^\top \boldsymbol{\alpha}_z)\mathbf{Z}_i^\top \mathbf{Z}_j \\
&= \frac{1}{n}\sum_{i\neq j}(\epsilon_i + D_i\breve{\beta} + \mathbf{X}_i^\top \breve{\boldsymbol{\alpha}}_x)(\epsilon_j + D_j\breve{\beta} + \mathbf{X}_j^\top \breve{\boldsymbol{\alpha}}_x)\mathbf{Z}_i^\top \mathbf{Z}_j \\
&\quad + \frac{1}{n}\sum_{i\neq j}(\epsilon_i \mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z \epsilon_j)\mathbf{Z}_i^\top \mathbf{Z}_j + \frac{1}{n}\sum_{i\neq j}(D_i\breve{\beta}\mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z D_j\breve{\beta})\mathbf{Z}_i^\top \mathbf{Z}_j \\
&\quad + \frac{1}{n}\sum_{i\neq j}(\mathbf{X}_i^\top \breve{\boldsymbol{\alpha}}_x \mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z \mathbf{X}_j^\top \breve{\boldsymbol{\alpha}}_x)\mathbf{Z}_i^\top \mathbf{Z}_j + \frac{1}{n}\sum_{i\neq j}\mathbf{Z}_i^\top \boldsymbol{\alpha}_z \mathbf{Z}_j^\top \boldsymbol{\alpha}_z \mathbf{Z}_i^\top \mathbf{Z}_j \\
&=: \sum_{l=1}^{4} T_{nl} + T_{n5} + T_{n6} + T_{n7} + T_{n8}.
\end{aligned}$$

From the proof of Theorem 1 in Appendix A.2, we know that

$$\sum_{l=1}^{4} T_{nl}/\Omega^{1/2} \xrightarrow{d} N(0,1) \qquad \text{as } (n, p_x, p_z) \to \infty. \tag{S29}$$

<u>**For the term T_{n5}**</u>, we denote $u_n^{(5)} = \frac{1}{n-1}T_{n5} = \frac{1}{n(n-1)}\sum_{i\neq j} u_{ij}^{(5)}$ being a U-statisitc with the kernel $u_{ij}^{(5)} = (\epsilon_i \mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z \epsilon_j)\mathbf{Z}_i^\top \mathbf{Z}_j$. Firstly, we derive that

$$\mathbb{E}(T_{n5}) \asymp n\mathbb{E}(u_n^{(5)}) = n\mathbb{E}\{(\epsilon_i \mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z \epsilon_j)\mathbf{Z}_i^\top \mathbf{Z}_j\} = 0.$$

Secondly, the projection of $u_{ij}^{(5)}$ to the space $\{\mathbf{Z}_i, \epsilon_i\}$ is

$$\begin{aligned}
u_{1i}^{(5)} &= \mathbb{E}(u_{ij}^{(5)} \mid \mathbf{Z}_i, \epsilon_i) \\
&= \mathbb{E}\{(\epsilon_i \mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z \epsilon_j)\mathbf{Z}_i^\top \mathbf{Z}_j \mid \mathbf{Z}_i, \epsilon_i\} \\
&= \epsilon_i \mathbf{Z}_i^\top \mathbb{E}(\mathbf{Z}_j \mathbf{Z}_j^\top \boldsymbol{\alpha}_z) = \epsilon_i \boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \mathbf{Z}_i.
\end{aligned}$$

Similar to the previous derivation with the Hoeffding decomposition, the variance of T_{n5} is

$$\mathrm{Var}(T_{n5}) \asymp n\mathrm{Var}(u_{1i}^{(5)}) + \mathrm{Var}(u_{ij}^{(5)}). \tag{S30}$$

Here, we calculate that

$$
\begin{aligned}
n\mathrm{Var}(u_{1i}^{(5)}) &= n\mathrm{Var}(\epsilon_i \boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \mathbf{Z}_i) \\
&= n\mathbb{E}\{(\epsilon_i \boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \mathbf{Z}_i)^2\} \\
&= n\mathbb{E}(\epsilon^2)\mathbb{E}(\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \mathbf{Z}_i \mathbf{Z}_i^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\alpha}_z) \\
&\lesssim n\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}^3 \boldsymbol{\alpha}_z,
\end{aligned} \tag{S31}
$$

where the third equality holds based on the random error ϵ is independent of $\mathbf{Z}$, and the first inequality follows from Assumption 3. Further, we yield that

$$
\begin{aligned}
\mathrm{Var}(u_{ij}^{(5)}) &= \mathrm{Var}\{(\epsilon_i \mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z \epsilon_j)\mathbf{Z}_i^\top \mathbf{Z}_j\} \\
&= \mathbb{E}[\{(\epsilon_i \mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z \epsilon_j)\mathbf{Z}_i^\top \mathbf{Z}_j\}^2] \\
&\lesssim \mathbb{E}\{(\epsilon_i \mathbf{Z}_j^\top \boldsymbol{\alpha}_z \mathbf{Z}_i^\top \mathbf{Z}_j)^2\} \\
&= \mathbb{E}(\epsilon^2)\mathbb{E}\{(\mathbf{Z}_j^\top \boldsymbol{\alpha}_z)^2(\mathbf{Z}_i^\top \mathbf{Z}_j)^2\} \\
&\lesssim \mathbb{E}^{1/2}\{(\mathbf{Z}_j^\top \boldsymbol{\alpha}_z)^4\}\mathbb{E}^{1/2}\{(\mathbf{Z}_i^\top \mathbf{Z}_j)^4\} \lesssim (\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\alpha}_z)\mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2),
\end{aligned} \tag{S32}
$$

where the third equality is due to the random error ϵ is independent of $\mathbf{Z}$, the second inequality follows from Assumption 3 together with the Cauchy-Schwarz inequality, and the last inequality holds by the equation (S64) in Lemma S8 and Lemma S2 with $b = 2$.

Combining the equations (S30), (S31) and (S32), it then follows that

$$
\mathrm{Var}(T_{n5}) \lesssim n\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}^3 \boldsymbol{\alpha}_z + (\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\alpha}_z)\mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2) = o\{\mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\},
$$

where the last equality holds by the conditions $\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\alpha}_z = o(1)$ and $\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}^3 \boldsymbol{\alpha}_z = o\{\mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)/n\}$ in $\mathscr{L}(\boldsymbol{\alpha}_z)$. As a result, we conclude that

$$
T_{n5} = o_p(\Omega^{1/2}). \tag{S33}
$$

For the term T_{n6}, we have

$$
\begin{aligned}
T_{n6} &= \frac{1}{n}\sum_{i \neq j}(D_i \breve{\beta} \mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z D_j \breve{\beta})\mathbf{Z}_i^\top \mathbf{Z}_j \\
&= \breve{\beta}\frac{1}{n}\sum_{i \neq j}(D_i \mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z D_j)\mathbf{Z}_i^\top \mathbf{Z}_j \\
&=: \breve{\beta} U_{n6}.
\end{aligned}
$$

Denote $u_n^{(6)} = \frac{1}{n-1}U_{n6} = \frac{1}{n(n-1)}\sum_{i \neq j} u_{ij}^{(6)}$ being a U-statistic with the kernel $u_{ij}^{(6)} = (D_i \mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z D_j)\mathbf{Z}_i^\top \mathbf{Z}_j$. We first calculate that

$$
\mathbb{E}(U_{n6}) \asymp n\mathbb{E}(u_n^{(6)}) = n\mathbb{E}\{(D_i \mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z D_j)\mathbf{Z}_i^\top \mathbf{Z}_j\} = 2n\boldsymbol{\mu}_{DZ}^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\mu}_{DZ}.
$$

Secondly, the projection of $u_{ij}^{(6)}$ to the space $\{D_i, \mathbf{Z}_i\}$ is

$$\begin{aligned} u_{1i}^{(6)} &= \mathbb{E}(u_{ij}^{(6)} \mid D_i, \mathbf{Z}_i) \\ &= \mathbb{E}\{(D_i\mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z D_j)\mathbf{Z}_i^\top \mathbf{Z}_j \mid D_i, \mathbf{Z}_i\} \\ &= D_i \mathbf{Z}_i^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\alpha}_z + \boldsymbol{\mu}_{D\mathbf{Z}}^\top \mathbf{Z}_i \mathbf{Z}_i^\top \boldsymbol{\alpha}_z. \end{aligned}$$

Using the Hoeffding decomposition, we deduce that

$$\mathrm{Var}(U_{n6}) \asymp n\mathrm{Var}(u_{1i}^{(6)}) + \mathrm{Var}(u_{ij}^{(6)}), \tag{S34}$$

with

$$\begin{aligned} n\mathrm{Var}(u_{1i}^{(6)}) &= n\mathrm{Var}(D_i\mathbf{Z}_i^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z + \boldsymbol{\mu}_{D\mathbf{Z}}^\top \mathbf{Z}_i\mathbf{Z}_i^\top \boldsymbol{\alpha}_z) \\ &\lesssim n\mathrm{Var}(D_i\mathbf{Z}_i^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z) + n\mathrm{Var}(\boldsymbol{\mu}_{D\mathbf{Z}}^\top \mathbf{Z}_i\mathbf{Z}_i^\top \boldsymbol{\alpha}_z) \\ &\le n\mathbb{E}\{(D_i\mathbf{Z}_i^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z))^2\} + n\mathbb{E}\{(\boldsymbol{\mu}_{D\mathbf{Z}}^\top \mathbf{Z}_i\mathbf{Z}_i^\top \boldsymbol{\alpha}_z)^2\} \\ &\lesssim n\mathbb{E}^{1/2}(D_i^4)\mathbb{E}^{1/2}\{(\mathbf{Z}^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)^4\} + n\mathbb{E}^{1/2}\{(\boldsymbol{\mu}_{D\mathbf{Z}}^\top \mathbf{Z})^4\}\mathbb{E}^{1/2}\{(\mathbf{Z}^\top \boldsymbol{\alpha}_z)^4\} \\ &\lesssim n\mathbb{E}^{1/2}\{(\mathbf{Z}^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)^4\} + n\mathbb{E}^{1/2}\{(\boldsymbol{\mu}_{D\mathbf{Z}}^\top \mathbf{Z})^4\}\mathbb{E}^{1/2}\{(\mathbf{Z}^\top \boldsymbol{\alpha}_z)^4\} \\ &\lesssim n\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}^3\boldsymbol{\alpha}_z + n(\boldsymbol{\mu}_{D\mathbf{Z}}^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\mu}_{D\mathbf{Z}})(\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z), \end{aligned} \tag{S35}$$

where the third equality follows from the Cauchy-Schwarz inequality, the fourth inequality holds by Assumption 2-3 with $D = \mathbf{X}^\top \boldsymbol{\gamma}_x + \mathbf{Z}^\top \boldsymbol{\gamma}_z + \delta$, and the last inequality is due to the equations (S64)-(S65) in Lemma S8 as well as the equation (S63) in Lemma S7. Further, we know that

$$\begin{aligned} \mathrm{Var}(u_{ij}^{(6)}) &= \mathrm{Var}\{(D_i\mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z D_j)\mathbf{Z}_i^\top \mathbf{Z}_j\} \\ &\le \mathbb{E}[\{(D_i\mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z D_j)\mathbf{Z}_i^\top \mathbf{Z}_j\}^2] \\ &\lesssim \mathbb{E}\{(D_i\mathbf{Z}_j^\top \boldsymbol{\alpha}_z\mathbf{Z}_i^\top \mathbf{Z}_j)^2\} \\ &\lesssim \mathbb{E}^{1/2}(D_i^4)\mathbb{E}^{1/2}\{(\mathbf{Z}_j^\top \boldsymbol{\alpha}_z)^4\}\mathbb{E}^{1/2}\{(\mathbf{Z}_i^\top \mathbf{Z}_j)^4\} \\ &\lesssim \mathbb{E}^{1/2}\{(\mathbf{Z}_j^\top \boldsymbol{\alpha}_z)^4\}\mathbb{E}^{1/2}\{(\mathbf{Z}_i^\top \mathbf{Z}_j)^4\} \lesssim (\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)\mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2), \end{aligned} \tag{S36}$$

where we apply the Cauchy-Schwarz inequality to obtain the third inequality, the fourth inequality follows from Assumption 2-3 with $D = \mathbf{X}^\top \boldsymbol{\gamma}_x + \mathbf{Z}^\top \boldsymbol{\gamma}_z + \delta$, and the last inequality holds by the equation (S64) in Lemma S8 and Lemma S2 with $b = 2$.

Combining the equations (S34), (S35) and (S36), we derive that

$$\begin{aligned} \mathrm{Var}(U_{n6}) &\lesssim n\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}^3\boldsymbol{\alpha}_z + n(\boldsymbol{\mu}_{D\mathbf{Z}}^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\mu}_{D\mathbf{Z}})(\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z) + (\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)\mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2) \\ &= o\{\mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\}, \end{aligned}$$

where the last inequality holds by the conditions $\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z = o(1)$, $\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}^3\boldsymbol{\alpha}_z = o\{\mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)/n\}$ in $\mathscr{L}(\boldsymbol{\alpha}_z)$ together with $n\boldsymbol{\mu}_{D\mathbf{Z}}^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\mu}_{D\mathbf{Z}} \lesssim \mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)$ (the equation (S62) in Lemma S7). Recall that $\mathbb{E}(U_{n6}) = 2n\boldsymbol{\mu}_{DZ}^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\mu}_{DZ} \lesssim \mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)$ and $\breve{\beta} = O_p\{(s_{\boldsymbol{\gamma}} \log p/n)^{1/2}\}$ in Assumption 6, we thus have

$$T_{n6} = \breve{\beta} U_{n6} = O_p\{(s_{\boldsymbol{\gamma}} \log p/n)^{1/2}\mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\} = o_p(\Omega^{1/2}), \tag{S37}$$

where the last equality follows from the condition $\underline{\underline{n^{-1/2}s_{\gamma}^{1/2}(\log p)^{1/2} = o(1)}}$.

For the term T_{n7}, under the Hölder's inequality, we deduce that

$$\begin{aligned} T_{n7} &= \frac{1}{n}\sum_{i\neq j}(\mathbf{X}_i^\top \breve{\boldsymbol{\alpha}}_x \mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z \mathbf{X}_j^\top \breve{\boldsymbol{\alpha}}_x)\mathbf{Z}_i^\top \mathbf{Z}_j \\ &\le \left\| \frac{1}{n}\sum_{i\neq j}(\mathbf{X}_i \mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z \mathbf{X}_j)\mathbf{Z}_i^\top \mathbf{Z}_j \right\|_\infty \|\breve{\boldsymbol{\alpha}}_x\|_1 =: \|\mathbf{U}_{n7}\|_\infty \|\breve{\boldsymbol{\alpha}}_x\|_1. \end{aligned}$$

We define the p_x-dimensional vector $\mathbf{u}_n^{(7)} = \frac{1}{n-1}\mathbf{U}_{n7} = \frac{1}{n(n-1)}\sum_{i\neq j}(\mathbf{X}_i \mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z \mathbf{X}_j)\mathbf{Z}_i^\top \mathbf{Z}_j$ with the k-th element

$$u_{nk}^{(7)} = \frac{1}{n(n-1)}\sum_{i\neq j}(X_{ik}\mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z X_{jk})\mathbf{Z}_i^\top \mathbf{Z}_j =: \frac{1}{n(n-1)}\sum_{i\neq j} u_{ijk}^{(7)},$$

where $u_{nk}^{(7)}$ is U-statistic with the kernel $u_{ijk}^{(7)} = (X_{ik}\mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z X_{jk})\mathbf{Z}_i^\top \mathbf{Z}_j$ for $k = 1, \ldots, p_x$.

We first yield that

$$\mathbb{E}(u_{nk}^{(7)}) = \mathbb{E}\{(X_{ik}\mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z X_{jk})\mathbf{Z}_i^\top \mathbf{Z}_j\} = 2\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\mathbb{E}(X_k \mathbf{Z}).$$

Using Hoeffding decomposition, we know that

$$\begin{aligned} u_{nk}^{(7)} &= 2\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\mathbb{E}(X_k \mathbf{Z}) + 2S_{1nk}^{(7)} + S_{2nk}^{(7)}, \\ S_{1nk}^{(7)} &= \frac{1}{n}\sum_{i=1}^n g_{1k}^{(7)}(\mathbf{X}_i, \mathbf{Z}_i), \\ S_{2nk}^{(7)} &= \frac{1}{n(n-1)}\sum_{i\neq j} g_{2k}^{(7)}(\mathbf{X}_i, \mathbf{Z}_i; \mathbf{X}_j, \mathbf{Z}_j), \end{aligned}$$

with

$$\begin{aligned} g_{1k}^{(7)}(\mathbf{X}_i, \mathbf{Z}_i) &= \mathbb{E}(u_{ijk}^{(7)} \mid \mathbf{X}_i, \mathbf{Z}_i) - 2\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\mathbb{E}(X_k \mathbf{Z}) \\ &= \mathbb{E}\{(X_{ik}\mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z X_{jk})\mathbf{Z}_i^\top \mathbf{Z}_j \mid \mathbf{X}_i, \mathbf{Z}_i\} - 2\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\mathbb{E}(X_k \mathbf{Z}) \\ &= X_{ik}\mathbf{Z}_i^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z + \boldsymbol{\alpha}_z^\top \mathbf{Z}_i \mathbf{Z}_i^\top \mathbb{E}(X_k \mathbf{Z}) - 2\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\mathbb{E}(X_k \mathbf{Z}) \\ &= \boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\{X_{ik}\mathbf{Z}_i - \mathbb{E}(X_k \mathbf{Z})\} + \mathbb{E}(X_k \mathbf{Z})^\top (\mathbf{Z}_i \mathbf{Z}_i^\top \boldsymbol{\alpha}_z - \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z), \end{aligned}$$

and

$$\begin{aligned} & g_{2k}^{(7)}(\mathbf{X}_i, \mathbf{Z}_i; \mathbf{X}_j, \mathbf{Z}_j) \\ &= u_{ijk}^{(7)} - 2\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\mathbb{E}(X_k \mathbf{Z}) - g_{1k}^{(7)}(\mathbf{X}_i, \mathbf{Z}_i) - g_{1k}^{(7)}(\mathbf{X}_j, \mathbf{Z}_j) \\ &= (X_{ik}\mathbf{Z}_j^\top \boldsymbol{\alpha}_z + \mathbf{Z}_i^\top \boldsymbol{\alpha}_z X_{jk})\mathbf{Z}_i^\top \mathbf{Z}_j - X_{ik}\mathbf{Z}_i^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z + \boldsymbol{\alpha}_z^\top \mathbf{Z}_i \mathbf{Z}_i^\top \mathbb{E}(X_k \mathbf{Z}) \\ &\quad - X_{jk}\mathbf{Z}_j^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z + \boldsymbol{\alpha}_z^\top \mathbf{Z}_j \mathbf{Z}_j^\top \mathbb{E}(X_k \mathbf{Z}) + 2\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\mathbb{E}(X_k \mathbf{Z}) \\ &= (\mathbf{Z}_i \mathbf{Z}_i^\top \boldsymbol{\alpha}_z - \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)^\top \{X_{jk}\mathbf{Z}_j - \mathbb{E}(X_k \mathbf{Z})\} \\ &\quad + (\mathbf{Z}_j \mathbf{Z}_j^\top \boldsymbol{\alpha}_z - \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)^\top \{X_{ik}\mathbf{Z}_i - \mathbb{E}(X_k \mathbf{Z})\}. \end{aligned}$$

To proceed, using a similar line of (S22), we derive that

$$\begin{aligned}
&\|\max_{1\le i\le n}\max_{1\le k\le p_x} g_{1k}^{(7)}(\mathbf{X}_i,\mathbf{Z}_i)\|_2 \\
&=\|\max_{1\le i\le n}\max_{1\le k\le p_x} \boldsymbol{\alpha}_z^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\{X_{ik}\mathbf{Z}_i-\mathbb{E}(X_k\mathbf{Z})\}+\mathbb{E}(X_k\mathbf{Z})^\top(\mathbf{Z}_i\mathbf{Z}_i^\top\boldsymbol{\alpha}_z-\boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)\|_2 \\
&\le\|\max_{1\le i\le n}\max_{1\le k\le p_x} \boldsymbol{\alpha}_z^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\{X_{ik}\mathbf{Z}_i-\mathbb{E}(X_k\mathbf{Z})\}\|_2 \\
&\quad+\|\max_{1\le i\le n}\max_{1\le k\le p_x}\mathbb{E}(X_k\mathbf{Z})^\top(\mathbf{Z}_i\mathbf{Z}_i^\top\boldsymbol{\alpha}_z-\boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)\|_2 \\
&\lesssim\log(np_x)\{\max_{1\le i\le n}\max_{1\le k\le p_x}\|\boldsymbol{\alpha}_z^\top\boldsymbol{\Sigma}_{\mathbf{Z}}X_{ik}\mathbf{Z}_i\|_{\psi_1}+\max_{1\le i\le n}\max_{1\le k\le p_x}\|\mathbb{E}(X_k\mathbf{Z})^\top\mathbf{Z}_i\mathbf{Z}_i^\top\boldsymbol{\alpha}_z\|_{\psi_1}\} \qquad \text{(S38)}\\
&\le\log(np_x)\{\max_{1\le i\le n}\max_{1\le k\le p_x}\|X_{ik}\|_{\psi_2}\|\boldsymbol{\alpha}_z^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\mathbf{Z}_i\|_{\psi_2} \\
&\qquad\qquad+\max_{1\le i\le n}\max_{1\le k\le p_x}\|\mathbb{E}(X_k\mathbf{Z})^\top\mathbf{Z}_i\|_{\psi_2}\|\mathbf{Z}_i^\top\boldsymbol{\alpha}_z\|_{\psi_2}\} \\
&\lesssim\log(np_x)[(\boldsymbol{\alpha}_z^\top\boldsymbol{\Sigma}_{\mathbf{Z}}^3\boldsymbol{\alpha}_z)^{1/2}+\{(\boldsymbol{\alpha}_z^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)\mathbb{E}(X_k\mathbf{Z})^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\mathbb{E}(X_k\mathbf{Z})\}^{1/2}] \\
&\lesssim\log(np_x)(\boldsymbol{\alpha}_z^\top\boldsymbol{\Sigma}_{\mathbf{Z}}^3\boldsymbol{\alpha}_z)^{1/2}+\log(np_x)(\boldsymbol{\alpha}_z^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)^{1/2}\kappa\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}),
\end{aligned}$$

where the second inequality is due to the equation (S49) in Lemma S4, and we apply the similar techniques used in the equation (S7) and Assumption 2 to establish the remaining inequalities. Specifically, we apply Lemma S5 to obtain that $\boldsymbol{\alpha}_z^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\mathbf{Z}$, $\mathbb{E}(X_k\mathbf{Z})^\top\mathbf{Z}$ and $\boldsymbol{\alpha}_z^\top\mathbf{Z}$ are *sub-Gaussian* random variables with

$$\begin{aligned}
\|\boldsymbol{\alpha}_z^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\mathbf{Z}\|_{\psi_2}&\lesssim(\boldsymbol{\alpha}_z^\top\boldsymbol{\Sigma}_{\mathbf{Z}}^3\boldsymbol{\alpha}_z)^{1/2},\\
\|\mathbb{E}(X_k\mathbf{Z})^\top\mathbf{Z}\|_{\psi_2}&\lesssim\{\mathbb{E}(X_k\mathbf{Z})^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\mathbb{E}(X_k\mathbf{Z})\}^{1/2},\\
\|\boldsymbol{\alpha}_z^\top\mathbf{Z}\|_{\psi_2}&\lesssim(\boldsymbol{\alpha}_z^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)^{1/2},
\end{aligned}$$

respectively. Similarly, we yield that

$$\begin{aligned}
\max_{1\le k\le p_x}\|g_{1k}^{(7)}(\mathbf{X}_i,\mathbf{Z}_i)\|_2&\le\max_{1\le k\le p_x}\|\boldsymbol{\alpha}_z^\top\boldsymbol{\Sigma}_{\mathbf{Z}}X_{ik}\mathbf{Z}_i\|_2+\max_{1\le k\le p_x}\|\mathbb{E}(X_k\mathbf{Z})^\top\mathbf{Z}_i\mathbf{Z}_i^\top\boldsymbol{\alpha}_z\|_2\\
&\lesssim\max_{1\le k\le p_x}\|\boldsymbol{\alpha}_z^\top\boldsymbol{\Sigma}_{\mathbf{Z}}X_{ik}\mathbf{Z}_i\|_{\psi_1}+\max_{1\le k\le p_x}\|\mathbb{E}(X_k\mathbf{Z})^\top\mathbf{Z}_i\mathbf{Z}_i^\top\boldsymbol{\alpha}_z\|_{\psi_1} \qquad \text{(S39)}\\
&\lesssim(\boldsymbol{\alpha}_z^\top\boldsymbol{\Sigma}_{\mathbf{Z}}^3\boldsymbol{\alpha}_z)^{1/2}+(\boldsymbol{\alpha}_z^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)^{1/2}\kappa\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}).
\end{aligned}$$

Here, the second inequality holds by the equation (S48) in Lemma S4, and the last inequality has been derived in the equation (S38).

Considering $q=\lceil 4/(1-3a)\rceil$ with $0<a<1/3$, we use the similar lines of the equations (S9), (S10) and (S38) to deduce that

$$\begin{aligned}
&\|\max_{1\le i\ne j\le n}\max_{1\le k\le p_x} g_{2k}^{(7)}(\mathbf{X}_i,\mathbf{Z}_i;\mathbf{X}_j,\mathbf{Z}_j)\|_4\\
&\le\|\max_{1\le i\ne j\le n}\max_{1\le k\le p_x} g_{2k}^{(7)}(\mathbf{X}_i,\mathbf{Z}_i;\mathbf{X}_j,\mathbf{Z}_j)\|_q\\
&\lesssim\|\max_{1\le i\ne j\le n}\max_{1\le k\le p_x} X_{jk}\boldsymbol{\alpha}_z^\top\mathbf{Z}_i\mathbf{Z}_i^\top\mathbf{Z}_j\|_q
\end{aligned}$$

$$
\begin{aligned}
&\lesssim \| \max_{1\le i\ne j\le n} \max_{1\le k\le p_x} X_{jk}\boldsymbol{\alpha}_z^\top \mathbf{Z}_i\|_{2q} \cdot \| \max_{1\le i\ne j\le n} \mathbf{Z}_i^\top \mathbf{Z}_j\|_{2q} \qquad \text{(S40)}\\
&\lesssim \log(np_x) \max_{1\le i\ne j\le n} \max_{1\le k\le p_x} \|X_{jk}\boldsymbol{\alpha}_z^\top \mathbf{Z}_i\|_{\psi_1} \cdot \{n(n-1)\}^{1/2q} \max_{1\le i\ne j\le n} \|\mathbf{Z}_i^\top \mathbf{Z}_j\|_{2q}\\
&\lesssim \log(np_x) \max_{1\le i\ne j\le n} \max_{1\le k\le p_x} \|X_{jk}\|_{\psi_2}\|\boldsymbol{\alpha}_z^\top \mathbf{Z}_i\|_{\psi_2} \cdot n^{1/q}\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\\
&\lesssim \log(np_x)(\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)^{1/2} n^{1/q}\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2),
\end{aligned}
$$

and

$$
\begin{aligned}
\max_{1\le k\le p_x} \|g_{2k}^{(7)}(\mathbf{X}_i,\mathbf{Z}_i;\mathbf{X}_j,\mathbf{Z}_j)\|_2 &\le \max_{1\le k\le p_x} \|g_{2k}^{(7)}(\mathbf{X}_i,\mathbf{Z}_i;\mathbf{X}_j,\mathbf{Z}_j)\|_4\\
&\lesssim \max_{1\le k\le p_x} \|X_{jk}\boldsymbol{\alpha}_z^\top \mathbf{Z}_i\mathbf{Z}_i^\top \mathbf{Z}_j\|_4\\
&\lesssim \max_{1\le k\le p_x} \|X_{jk}\boldsymbol{\alpha}_z^\top \mathbf{Z}_i\|_8 \cdot \|\mathbf{Z}_i^\top \mathbf{Z}_j\|_8 \qquad \text{(S41)}\\
&\lesssim \max_{1\le k\le p_x} \|X_{jk}\boldsymbol{\alpha}_z^\top \mathbf{Z}_i\|_{\psi_1} \cdot \|\mathbf{Z}_i^\top \mathbf{Z}_j\|_8\\
&\lesssim (\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)^{1/2}\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2).
\end{aligned}
$$

Combining the equations (S38)-(S41) and under Lemma S3, we deduce that

$$
\begin{aligned}
&\mathbb{E}(\|\mathbf{U}_{n7}\|_\infty)\\
&\asymp n\mathbb{E}(\|\mathbf{u}_n^{(7)}\|_\infty)\\
&\lesssim n \max_{1\le k\le p_x} \boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\mathbb{E}(X_k\mathbf{Z})\\
&\quad + \{n^{1/2}(\log p_x)^{1/2} + \log p_x \log(np_x)\}\{(\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}^3\boldsymbol{\alpha}_z)^{1/2} + (\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)^{1/2}\kappa\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}})\} \qquad \text{(S42)}\\
&\quad + (\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)^{1/2}\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\{\log p_x + n^{-1/2+1/q}(\log p_x)^{3/2}\log(np_x)\}\\
&\lesssim n(\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}^2\boldsymbol{\alpha}_z)^{1/2}\kappa + n^{1/2}(\log p_x)^{1/2}\{(\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}^3\boldsymbol{\alpha}_z)^{1/2} + (\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)^{1/2}\kappa\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}})\}\\
&\quad + \log p_x(\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)^{1/2}\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2),
\end{aligned}
$$

where the second inequality holds by the Cauchy-Schwarz inequality and Assumption 4 (implying that $n^{1/2}(\log p_x)^{1/2}$ and $\log p_x$ dominate the bound). Referring to $T_{n7} \le \|\mathbf{U}_{n7}\|_\infty \|\breve{\boldsymbol{\alpha}}_x\|_1$, the equation (S42) and Proposition 1, we yield that

$$
\begin{aligned}
T_{n7} &= O_p\{n^{-1/2}s_{\boldsymbol{\alpha}_x}(\log p_x)^{1/2}\}\\
&\quad \cdot O_p\{n(\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}^2\boldsymbol{\alpha}_z)^{1/2}\kappa\\
&\qquad + n^{1/2}(\log p_x)^{1/2}\{(\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}^3\boldsymbol{\alpha}_z)^{1/2} + (\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)^{1/2}\kappa\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}})\}\\
&\qquad + \log p_x(\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)^{1/2}\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\}\\
&= O_p\{n^{1/2}s_{\boldsymbol{\alpha}_x}(\log p_x)^{1/2}\kappa(\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}^2\boldsymbol{\alpha}_z)^{1/2} + s_{\boldsymbol{\alpha}_x}\log p_x(\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}^3\boldsymbol{\alpha}_z)^{1/2} \qquad \text{(S43)}\\
&\qquad + s_{\boldsymbol{\alpha}_x}\log p_x\lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}})\kappa(\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)^{1/2} +\\
&\qquad + n^{-1/2}s_{\boldsymbol{\alpha}_x}(\log p_x)^{3/2}(\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\alpha}_z)^{1/2}\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\}\\
&= o_p(\Omega^{1/2}),
\end{aligned}
$$

where the last equality is due to $\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\alpha}_z = o(1)$, $\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}^2 \boldsymbol{\alpha}_z = o\{\mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)/(n s_{\boldsymbol{\alpha}_x}^2 \log p_x \kappa^2)\}$ and $\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}^3 \boldsymbol{\alpha}_z = o\{\mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)/n\}$ in $\mathscr{L}(\boldsymbol{\alpha}_z)$ and other conditions

$$\begin{aligned} n^{-1/2} s_{\boldsymbol{\alpha}_x} \log p_x &= o(1), \\ s_{\boldsymbol{\alpha}_x} \log p_x \lambda_{\max}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}) \kappa &= o\{\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\}, \\ n^{-1/2} s_{\boldsymbol{\alpha}_x} (\log p_x)^{3/2} &= o(1). \end{aligned}$$

For the term T_{n8}, we denote $u_n^{(8)} = \frac{1}{n-1} T_{n8} = \frac{1}{n(n-1)} \sum_{i \neq j} u_{ij}^{(8)}$ being a U-statisitc with the kernel $u_{ij}^{(8)} = \mathbf{Z}_i^\top \boldsymbol{\alpha}_z \mathbf{Z}_j^\top \boldsymbol{\alpha}_z \mathbf{Z}_i^\top \mathbf{Z}_j$. Firstly, we calculate that

$$\mathbb{E}(T_{n8}) \asymp n \mathbb{E}(u_n^{(8)}) = n \mathbb{E}(\mathbf{Z}_i^\top \boldsymbol{\alpha}_z \mathbf{Z}_j^\top \boldsymbol{\alpha}_z \mathbf{Z}_i^\top \mathbf{Z}_j) = \boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}^2 \boldsymbol{\alpha}_z.$$

Secondly, the projection of $u_{ij}^{(8)}$ to the space $\{\mathbf{Z}_i\}$ is

$$u_{1i}^{(8)} = \mathbb{E}(u_{ij}^{(8)} \mid \mathbf{Z}_i) = \mathbb{E}(\mathbf{Z}_i^\top \boldsymbol{\alpha}_z \mathbf{Z}_j^\top \boldsymbol{\alpha}_z \mathbf{Z}_i^\top \mathbf{Z}_j \mid \mathbf{Z}_i) = \boldsymbol{\alpha}_z^\top \mathbf{Z}_i \mathbf{Z}_i^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\alpha}_z.$$

Applying the Hoeffding decomposition, the variance of T_{n5} is

$$\mathrm{Var}(T_{n8}) \asymp n \mathrm{Var}(u_{1i}^{(8)}) + \mathrm{Var}(u_{ij}^{(8)}), \tag{S44}$$

where we derive that

$$\begin{aligned} n \mathrm{Var}(u_{1i}^{(8)}) &= n \mathrm{Var}(\boldsymbol{\alpha}_z^\top \mathbf{Z}_i \mathbf{Z}_i^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\alpha}_z) \\ &\le n \mathbb{E}\{(\boldsymbol{\alpha}_z^\top \mathbf{Z}_i \mathbf{Z}_i^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\alpha}_z)^2\} \\ &\lesssim n \mathbb{E}^{1/2}\{(\boldsymbol{\alpha}_z^\top \mathbf{Z}_i)^4\} \mathbb{E}^{1/2}\{(\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \mathbf{Z}_i)^4\} \\ &\lesssim n (\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\alpha}_z)(\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}^3 \boldsymbol{\alpha}_z), \end{aligned} \tag{S45}$$

where the second inequality holds by the Cauchy-Schwarz inequality, and the last inequality follows from the equations (S64)-(S65) in Lemma S8. Further, we yield that

$$\begin{aligned} \mathrm{Var}(u_{ij}^{(8)}) &= \mathrm{Var}(\mathbf{Z}_i^\top \boldsymbol{\alpha}_z \mathbf{Z}_j^\top \boldsymbol{\alpha}_z \mathbf{Z}_i^\top \mathbf{Z}_j) \\ &\le \mathbb{E}\{(\mathbf{Z}_i^\top \boldsymbol{\alpha}_z \mathbf{Z}_j^\top \boldsymbol{\alpha}_z \mathbf{Z}_i^\top \mathbf{Z}_j)^2\} \\ &\lesssim \mathbb{E}^{1/2}\{(\mathbf{Z}_i^\top \boldsymbol{\alpha}_z)^4\} \mathbb{E}^{1/2}\{(\mathbf{Z}_j^\top \boldsymbol{\alpha}_z)^4\} \mathbb{E}^{1/2}\{(\mathbf{Z}_i^\top \mathbf{Z}_j)^4\} \\ &\lesssim (\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\alpha}_z)^2 \mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2), \end{aligned} \tag{S46}$$

where the second equality is due to the Cauchy-Schwarz inequality, and the last inequality holds by the equation (S64) in Lemma S8 and Lemma S2 with $b = 2$.

Combining the equations (S44), (S45) and (S46), it then follows that

$$\mathrm{Var}(T_{n8}) \lesssim n (\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\alpha}_z)(\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}^3 \boldsymbol{\alpha}_z) + (\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\alpha}_z)^2 \mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2) = o\{\mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\},$$

where the last equality holds by the conditions $\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\alpha}_z = o(1)$ and $\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}^3 \boldsymbol{\alpha}_z = o\{\mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)/n\}$ in $\mathscr{L}(\boldsymbol{\alpha}_z)$. As a result, we conclude that

$$T_{n8} = n \boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}^2 \boldsymbol{\alpha}_z + o_p(\Omega^{1/2}). \tag{S47}$$

In sum, following the results (S29), (S33), (S37), (S43) and (S47), we verify that

$$(T_n - n \boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}}^2 \boldsymbol{\alpha}_z)/\Omega^{1/2} \xrightarrow{d} N(0,1) \qquad \text{as } (n, p_x, p_z) \to \infty.$$

The proof is completed. □

Appendix B Auxiliary lemmas

Lemma S1. *Let* $\{\mathbf{A}_i\}_{i=1}^k$ *be a sequence of* $p \times p$*-dimensional semi-positive matrices and* k *is a fixed positive integer. Suppose that* $\boldsymbol{\nu}$ *is a* p*-dimensional sub-Gaussian random vector with sub-Gaussian norm* $\|\boldsymbol{\nu}\|_{\psi_2}$*, we have*

$$\mathbb{E}(\prod_{i=1}^k \boldsymbol{\nu}^\top \mathbf{A}_i \boldsymbol{\nu}) \lesssim \|\boldsymbol{\nu}\|_{\psi_2}^{2k} \prod_{i=1}^k \operatorname{tr}(\mathbf{A}_i).$$

Proof. See Yang et al. (2022) for detailed proof. □

Lemma S2. *Under Assumption 2, for a fixed positive integer* b*, we have*

$$\mathbb{E}\{(\mathbf{V}_1^\top \mathbf{V}_2)^{2b}\} \lesssim \{\operatorname{tr}(\boldsymbol{\Sigma}_{\mathbf{V}}^2)\}^b,$$

where $\boldsymbol{\Sigma}_{\mathbf{V}} = \mathbb{E}(\mathbf{V}\mathbf{V}^\top) = \mathbb{E}(\boldsymbol{\Phi}\boldsymbol{\nu}\boldsymbol{\nu}^\top\boldsymbol{\Phi}^\top) = \boldsymbol{\Phi}\boldsymbol{\Phi}^\top$.

Proof. This is a corollary of Lemma S1, it can be shown that

$$\begin{aligned}
\mathbb{E}\{(\mathbf{V}_1^\top \mathbf{V}_2)^{2b}\} &= \mathbb{E}\{(\mathbf{V}_1^\top \mathbf{V}_2 \mathbf{V}_2^\top \mathbf{V}_1)^b\} \\
&= \mathbb{E}[\mathbb{E}\{(\mathbf{V}_2^\top \mathbf{V}_1 \mathbf{V}_1^\top \mathbf{V}_2)^b \mid \mathbf{V}_1\}] \\
&= \mathbb{E}[\mathbb{E}\{(\boldsymbol{\nu}_2^\top \underline{\boldsymbol{\Phi}^\top \mathbf{V}_1 \mathbf{V}_1^\top \boldsymbol{\Phi}} \boldsymbol{\nu}_2)^b \mid \mathbf{V}_1\}] \\
&\lesssim \mathbb{E}[\{\operatorname{tr}(\boldsymbol{\Phi}^\top \mathbf{V}_1 \mathbf{V}_1^\top \boldsymbol{\Phi})\}^b] = \mathbb{E}\{(\mathbf{V}_1^\top \boldsymbol{\Phi}\boldsymbol{\Phi}^\top \mathbf{V}_1)^b\} \\
&= \mathbb{E}\{(\boldsymbol{\nu}_1^\top \underline{\boldsymbol{\Phi}^\top \boldsymbol{\Sigma}_{\mathbf{V}} \boldsymbol{\Phi}} \boldsymbol{\nu}_1)^b\} \\
&\lesssim \{\operatorname{tr}(\boldsymbol{\Phi}^\top \boldsymbol{\Sigma}_{\mathbf{V}} \boldsymbol{\Phi})\}^b = \{\operatorname{tr}(\boldsymbol{\Sigma}_{\mathbf{V}}^2)\}^b.
\end{aligned}$$

Here, the first inequality holds by Lemma S1 and $\mathbf{V}_1$ is independent of $\mathbf{V}_2$, and the second inequality follows from Lemma S1. □

Lemma S3 (A maximal inequality for U-statistics of order two)**.** *Let* $\{X_i\}_{i=1}^n$ *be a sample of i.i.d. random variables in a separable and measurable space* $(S, \mathcal{S})$*. Let* $\boldsymbol{g} : S \times S \to \mathbb{R}^p$ *be an* $\mathcal{S} \otimes \mathcal{S}$*-measurable, symmetric kernel such that* $\mathbb{E}\{|g_k(X_1, X_2)|\} < \infty$ *for all* $k = 1, \ldots, p$*. Denote* $\mathbf{U}_n = \frac{1}{n(n-1)} \sum_{1 \le i \ne j \le n} \boldsymbol{g}(X_i, X_j)$ *with*

$$\begin{aligned}
\boldsymbol{g}_1(X) &= \mathbb{E}\{\boldsymbol{g}(X_1, X_2) \mid X_1\} - \mathbb{E}\{\boldsymbol{g}(X_1, X_2)\}, \\
\boldsymbol{g}_2(X_1, X_2) &= \boldsymbol{g}(X_1, X_2) - \mathbb{E}\{\boldsymbol{g}(X_1, X_2)\} - \boldsymbol{g}_1(X_1) - \boldsymbol{g}_1(X_2).
\end{aligned}$$

We assume $2 \le p \le \exp(cn)$ *for some constant* $c > 0$*, it then follows that*

$$\begin{aligned}
\mathbb{E}(\|\mathbf{U}_n\|_\infty) \lesssim\ & \|\mathbb{E}\{\boldsymbol{g}(X_1, X_2)\}\|_\infty + (\log p/n)^{1/2} A_{12} + (\log p/n) B_{12} \\
& + (\log p/n) A_{22} + (\log p/n)^{5/4} A_{24} + (\log p/n)^{3/2} B_{24},
\end{aligned}$$

where

$$A_{12} = \max_{1\le k\le p} \|g_{1k}(X_1)\|_2, \qquad B_{12} = \|\max_{1\le i\le n}\max_{1\le k\le p} |g_{1k}(X_i)|\|_2,$$
$$A_{22} = \max_{1\le k\le p} \|g_{2k}(X_1, X_2)\|_2, \quad B_{24} = \|\max_{1\le i\neq j\le n}\max_{1\le k\le p} |g_{2k}(X_i, X_j)|\|_4,$$
$$A_{24} = \max_{1\le k\le p} \|g_{2k}(X_1, X_2)\|_4.$$

Note that g_k, g_{1k} and g_{2k} are the k-th components of $\boldsymbol{g}$, $\boldsymbol{g}_1$ and $\boldsymbol{g}_2$, respectively.

Proof. See Yang et al. (2022) for detailed proof. □

Lemma S4 (The properties of Orlicz norm)**.** *We define the ψ_α-Orlicz norm for any random variable X and $\alpha \in (0, 2]$ by*

$$\|X\|_{\psi_\alpha} = \inf\{t > 0 : \mathrm{E}\{\exp(|X|^\alpha/t^\alpha)\} \le 2\},$$

for $\psi_\alpha(x) = \exp(x^\alpha) - 1$. Then, for any random variables $X_1, \dots, X_m$ (without independence assumption) and any fixed integer $q \ge 1$,

$$\|X\|_q \le q!\|X\|_{\psi_1}, \quad \|X\|_{\psi_p} \le \|X\|_{\psi_q}(\log 2)^{p/q}, \quad p \le q, \tag{S48}$$
$$\|\max_{1\le i\le m} X_i\|_q \lesssim \log(m) \max_{1\le i\le m} \|X_i\|_{\psi_1}, \tag{S49}$$
$$\|\max_{1\le i\le m} X_i\|_q \le m^{1/q} \max_{1\le i\le m} \|X_i\|_q. \tag{S50}$$

Proof. For the properties of Orlicz norm in (S48), see Van der Vaart and Wellner (1996) on page 95. We further apply Lemma 2.2.2 in Van der Vaart and Wellner (1996) on page 96 to obtain that

$$\|\max_{1\le i\le m} X_i\|_q \lesssim \|\max_{1\le i\le m} X_i\|_{\psi_1} \lesssim \log(m) \max_{1\le i\le m} \|X_i\|_{\psi_1}.$$

Using the fact that $\max_{1\le i\le m} |X_i|^q \le \sum_{i=1}^m |X_i|^q$, one easily obtains for the L_q norms

$$\|\max_{1\le i\le m} X_i\|_q = \{\mathbb{E}(\max_{1\le i\le m} |X_i|^q)\}^{1/q} \le m^{1/q} \max_{1\le i\le m} \|X_i\|_q.$$

□

Lemma S5. *Suppose $\boldsymbol{\nu}$ is a p-dimensional sub-Gaussian random vector. For any vector $\boldsymbol{u} \in \mathbb{R}^p$ such that $\|\boldsymbol{u}\|_2 > 0$, it can be shown that*

$$\|\boldsymbol{u}^\top \boldsymbol{\nu}\|_{\psi_2} \le \|\boldsymbol{u}\|_2 \|\boldsymbol{\nu}\|_{\psi_2}.$$

Proof. Applying the definition of *sub-Gaussian* norm of *sub-Gaussian* vector, we yield that

$$\|\boldsymbol{u}^\top \boldsymbol{\nu}\|_{\psi_2} = \|\boldsymbol{u}\|_2 \left\| \frac{1}{\|\boldsymbol{u}\|_2} \boldsymbol{u}^\top \boldsymbol{\nu} \right\|_{\psi_2} \le \|\boldsymbol{u}\|_2 \|\boldsymbol{\nu}\|_{\psi_2}.$$

□

Lemma S6. *Suppose that Assumptions 1-3 hold, then the term T_{n1} in the proof of Theorem 1 asymptotically follows*

$$T_{n1}/\Omega^{1/2} = \frac{\sum_{i\neq j}\epsilon_i\epsilon_j\mathbf{Z}_i^\top\mathbf{Z}_j}{n\Omega^{1/2}} \xrightarrow{d} N(0,1) \quad as\ (n,p_z)\to\infty,$$

where $\Omega = 2\sigma^4\mathrm{tr}(\mathbf{\Sigma}_{\mathbf{Z}}^2)$.

Proof. Denote $\eta_{ni} = 2n^{-1}\sum_{j=1}^{i-1}\epsilon_i\epsilon_j\mathbf{Z}_i^\top\mathbf{Z}_j$, and let $S_{nk} = \sum_{i=2}^k \eta_{ni}$ with $S_{nk} - S_{n(k-1)} = \eta_{nk}$ defined as martingale differences, and $\mathscr{F}_k = \sigma\{(\mathbf{Z}_i,\epsilon_i), i=1,\dots,k\}$. Also, $\mathbb{E}(\eta_{nk} \mid \mathscr{F}_{k-1}) = 0$ due to $(S_{nk},\mathscr{F}_k)$ is a zero-mean martingale sequence. We define $v_{ni} = \mathrm{Var}(\eta_{ni} \mid \mathscr{F}_{i-1})$ and $V_n = \sum_{i=2}^n v_{ni}$. Note that

$$S_{nn} = \frac{2}{n}\sum_{i=2}^n\sum_{j=1}^{i-1}\epsilon_i\epsilon_j\mathbf{Z}_i^\top\mathbf{Z}_j = \frac{1}{n}\sum_{i\neq j}\epsilon_i\epsilon_j\mathbf{Z}_i^\top\mathbf{Z}_j = T_{n1}.$$

Therefore, by the Martingale Central Limit Theorem (Hall and Heyde, 2014), it suffices to show that the following two conditions hold.

$$\frac{V_n}{\mathrm{Var}(S_{nn})} \xrightarrow{p} 1, \tag{S51}$$

and for all $\zeta > 0$,

$$\frac{\sum_{i=2}^n \mathbb{E}\left[\eta_{ni}^2 I\{|\eta_{ni}|/\sqrt{\mathrm{tr}(\mathbf{\Sigma}_{\mathbf{Z}}^2)} > \zeta\} \mid \mathscr{F}_{i-1}\right]}{\mathrm{tr}(\mathbf{\Sigma}_{\mathbf{Z}}^2)} \xrightarrow{p} 0. \tag{S52}$$

We first establish the equation (S51). Observe that

$$\begin{aligned} v_{ni} &= \mathrm{Var}(\eta_{ni}|\mathscr{F}_{i-1}) \\ &= \mathrm{Var}\left(\frac{2}{n}\sum_{j=1}^{i-1}\epsilon_i\epsilon_j\mathbf{Z}_i^\top\mathbf{Z}_j \,\middle|\, \mathscr{F}_{i-1}\right) \\ &= \frac{4\sigma^2}{n^2}\left\{\sum_{j=1}^{i-1}\epsilon_j^2\mathbf{Z}_j^\top\mathbf{\Sigma}_{\mathbf{Z}}\mathbf{Z}_j + \sum_{1\le j_1\neq j_2\le i-1}\epsilon_{j_1}\epsilon_{j_2}\mathbf{Z}_{j_1}^\top\mathbf{\Sigma}_{\mathbf{Z}}\mathbf{Z}_{j_2}\right\}, \end{aligned}$$

and

$$\begin{aligned} V_n &= \sum_{i=2}^n v_{ni} \\ &= \frac{4\sigma^2}{n^2}\sum_{i=2}^n\sum_{j=1}^{i-1}\epsilon_j^2\mathbf{Z}_j^\top\mathbf{\Sigma}_{\mathbf{Z}}\mathbf{Z}_j + \frac{4\sigma^2}{n^2}\sum_{i=2}^n\sum_{1\le j_1\neq j_2\le i-1}\epsilon_{j_1}\epsilon_{j_2}\mathbf{Z}_{j_1}^\top\mathbf{\Sigma}_{\mathbf{Z}}\mathbf{Z}_{j_2}. \end{aligned} \tag{S53}$$

Notice that $S_{nn} = \frac{1}{n}\sum_{i\neq j}\epsilon_i\epsilon_j\mathbf{Z}_i^\top\mathbf{Z}_j$ and denote $u_n^{(s)} := \frac{1}{n-1}S_{nn}$ being a U-statistic with the kernel $u_{ij}^{(s)} = \epsilon_i\epsilon_j\mathbf{Z}_i^\top\mathbf{Z}_j$. The projection of $u_{ij}^{(s)}$ to the space $\{\mathbf{Z}_i,\epsilon_i\}$ is

$$u_{1i}^{(s)} = \mathbb{E}(u_{ij}^{(s)} \mid \mathbf{Z}_i,\epsilon_i) = \mathbb{E}(\epsilon_i\epsilon_j\mathbf{Z}_i^\top\mathbf{Z}_j \mid \mathbf{Z}_i,\epsilon_i) = 0.$$

Furthermore, by the Hoeffding decomposition,

$$
\begin{aligned}
\operatorname{Var}(S_{nn}) &= \frac{4(n-1)(n-2)}{n}\operatorname{Var}(u_{1i}^{(s)}) + \frac{2(n-1)}{n}\operatorname{Var}(u_{ij}^{(s)}) \\
&= \frac{2(n-1)}{n}\operatorname{Var}(u_{ij}^{(s)}) \\
&= \frac{2(n-1)}{n}\mathbb{E}\{\epsilon_i^2\epsilon_j^2(\mathbf{Z}_i^\top\mathbf{Z}_j)^2\} \\
&= \frac{2(n-1)}{n}\{\mathbb{E}(\epsilon^2)\}^2\operatorname{tr}\{\mathbb{E}(\mathbf{Z}_i\mathbf{Z}_i^\top\mathbf{Z}_j\mathbf{Z}_j^\top)\} \\
&= \frac{2(n-1)}{n}\sigma^4\operatorname{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2).
\end{aligned} \tag{S54}
$$

Then combining the equations (S53) and (S54), we write

$$
\begin{aligned}
&\frac{V_n}{\operatorname{Var}(S_{nn})} \\
&= \frac{2}{n(n-1)\sigma^2\operatorname{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)}\left(\sum_{i=2}^{n}\sum_{j=1}^{i-1}\epsilon_j^2\mathbf{Z}_j^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\mathbf{Z}_j + \sum_{i=2}^{n}\sum_{1\le j_1\ne j_2\le i-1}\epsilon_{j_1}\epsilon_{j_2}\mathbf{Z}_{j_1}^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\mathbf{Z}_{j_2}\right) \\
&=: R_1 + R_2.
\end{aligned}
$$

Now we need to show that $R_1 \xrightarrow{p} 1$ and $R_2 \xrightarrow{p} 0$. It can be derived that

$$
\begin{aligned}
\mathbb{E}(R_1) &= \mathbb{E}\left\{\frac{2}{n(n-1)\sigma^2\operatorname{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)}\sum_{i=2}^{n}\sum_{j=1}^{i-1}\epsilon_j^2\mathbf{Z}_j^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\mathbf{Z}_j\right\} \\
&= \mathbb{E}\left\{\frac{1}{n(n-1)\sigma^2\operatorname{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)}\sum_{i\ne j}\epsilon_j^2\mathbf{Z}_j^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\mathbf{Z}_j\right\} \\
&= \frac{1}{\sigma^2\operatorname{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)}\mathbb{E}(\epsilon^2)\mathbb{E}(\mathbf{Z}^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\mathbf{Z}) = 1.
\end{aligned} \tag{S55}
$$

Here, the last equality holds by $\mathbb{E}(\mathbf{Z}^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\mathbf{Z}) = \operatorname{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)$.

We know that

$$
\begin{aligned}
\operatorname{Var}(R_1) &= \operatorname{Var}\left(\frac{2}{n(n-1)\sigma^2\operatorname{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)}\sum_{i=2}^{n}\sum_{j=1}^{i-1}\epsilon_j^2\mathbf{Z}_j^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\mathbf{Z}_j\right) \\
&= \operatorname{Var}\left(\frac{2}{n(n-1)\sigma^2\operatorname{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)}\sum_{j=1}^{n-1}(n-j)\epsilon_j^2\mathbf{Z}_j^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\mathbf{Z}_j\right) \\
&\lesssim \frac{1}{n^4\operatorname{tr}^2(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)}\sum_{j=1}^{n-1}(n-j)^2\operatorname{Var}\left(\mathbf{Z}_j^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\mathbf{Z}_j\right) \\
&= \frac{1}{n^4\operatorname{tr}^2(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)}\sum_{j=1}^{n-1}(n-j)^2[\mathbb{E}\{(\mathbf{Z}_j^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\mathbf{Z}_j)^2\} - \mathbb{E}^2(\mathbf{Z}_j^\top\boldsymbol{\Sigma}_{\mathbf{Z}}\mathbf{Z}_j)]
\end{aligned} \tag{S56}
$$

$$\lesssim \frac{1}{n^4 \mathrm{tr}^2(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)} \sum_{j=1}^{n-1} (n-j)^2 \mathrm{tr}^2(\boldsymbol{\Sigma}_{\mathbf{Z}}^2) \lesssim \frac{1}{n} = o(1),$$

where the first inequality follows from Assumption 3, and the second inequality is obtained by $\mathbb{E}\{(\mathbf{Z}^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \mathbf{Z})^2\} \lesssim \mathrm{tr}^2(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)$ with Lemma S1 as well as $\mathbb{E}(\mathbf{Z}^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \mathbf{Z}) = \mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)$.

Similar to the derivation of $\mathrm{Var}(R_1)$, we obtain that

$$\begin{aligned}
\mathrm{Var}(R_2) &= \mathrm{Var}\left(\frac{2}{n(n-1)\sigma^2 \mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)} \sum_{i=2}^{n} \sum_{1 \le j_1 \ne j_2 \le i-1} \epsilon_{j_1} \epsilon_{j_2} \mathbf{Z}_{j_1}^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \mathbf{Z}_{j_2} \right) \\
&\lesssim \frac{1}{n^4 \mathrm{tr}^2(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)} \sum_{j_1 < k_1} \sum_{j_2 < k_2} (n-k_1)(n-k_2) \mathbb{E}(\epsilon_{j_1} \epsilon_{j_2} \epsilon_{k_1} \epsilon_{k_2} \mathbf{Z}_{j_1}^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \mathbf{Z}_{k_1} \mathbf{Z}_{j_2}^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \mathbf{Z}_{k_2}) \qquad \text{(S57)} \\
&\lesssim \frac{\sum_{k=1}^{n} (n-k)^2 (k-1)}{n^4} \cdot \frac{\mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^4)}{\mathrm{tr}^2(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)} = o(1).
\end{aligned}$$

Here, the last equality holds by the condition $\mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^4) = o\{\mathrm{tr}^2(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)\}$ in Assumption 1. Observe that $\mathbb{E}(R_2) = 0$, combining the equations (S55), (S56) and (S57), Chebyshev inequality yields that $R_1 \xrightarrow{p} 1$ and $R_2 \xrightarrow{p} 0$. Up to now, the equation (S51) is verified.

Next, we establish the equation (S52). For all $\zeta > 0$, we have

$$\frac{\sum_{i=2}^{n} \mathbb{E}[\eta_{ni}^2 I\{|\eta_{ni}| / \sqrt{\mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)} > \zeta\} | \mathscr{F}_{i-1}]}{\mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)} \le \frac{1}{\zeta^2} \cdot \frac{\sum_{i=2}^{n} \mathbb{E}(\eta_{ni}^4 | \mathscr{F}_{i-1})}{\mathrm{tr}^2(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)}, \qquad \text{(S58)}$$

which is due to the Markov inequality. Furthermore, we obtain that

$$\begin{aligned}
&\mathbb{E}\left\{ \frac{\sum_{i=2}^{n} \mathbb{E}\left(\eta_{ni}^4 | \mathscr{F}_{i-1}\right)}{\mathrm{tr}^2(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)} \right\} = \frac{\sum_{i=2}^{n} \mathbb{E}\left(\eta_{ni}^4\right)}{\mathrm{tr}^2(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)} \\
&= \frac{1}{\mathrm{tr}^2(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)} \sum_{i=2}^{n} \mathbb{E}\left\{ \frac{2^4}{n^4} \left(\sum_{j=1}^{i-1} \epsilon_i \epsilon_j \mathbf{Z}_i^\top \mathbf{Z}_j \right)^4 \right\} \\
&\lesssim \frac{1}{n^4 \mathrm{tr}^2(\boldsymbol{\Sigma}_{\mathbf{Z}}^2)} \left[\sum_{i=2}^{n} \sum_{s \ne t} \mathbb{E}\{(\mathbf{Z}_i^\top \mathbf{Z}_s)^2 (\mathbf{Z}_i^\top \mathbf{Z}_t)^2\} + \sum_{i=2}^{n} \sum_{j=1}^{i-1} \mathbb{E}(\mathbf{Z}_i^\top \mathbf{Z}_j)^4 \right] \qquad \text{(S59)} \\
&\lesssim \frac{1}{n} = o(1),
\end{aligned}$$

where the first inequality holds by Assumption 3, and the last inequality holds by the equations (S60) and (S61) as follows,

$$\sum_{i=2}^{n} \sum_{s \ne t} \mathbb{E}\{(\mathbf{Z}_i^\top \mathbf{Z}_s)^2 (\mathbf{Z}_i^\top \mathbf{Z}_t)^2\} \le n^3 \mathbb{E}^{1/2}\{(\mathbf{Z}_i^\top \mathbf{Z}_s)^4\} \mathbb{E}^{1/2}\{(\mathbf{Z}_i^\top \mathbf{Z}_t)^4\} \lesssim n^3 \mathrm{tr}^2(\boldsymbol{\Sigma}_{\mathbf{Z}}^2), \qquad \text{(S60)}$$

where the first inequality is obtained by the Cauchy-Schwarz inequality, and the last inequality holds by Lemma S2. Similarly, it follows that

$$\sum_{i=2}^{n} \sum_{j=1}^{i-1} \mathbb{E}(\mathbf{Z}_i^\top \mathbf{Z}_j)^4 \lesssim n^2 \mathbb{E}(\mathbf{Z}_i^\top \mathbf{Z}_j)^4 \lesssim n^2 \mathrm{tr}^2(\boldsymbol{\Sigma}_{\mathbf{Z}}^2). \qquad \text{(S61)}$$

The equation (S52) is thus proved by combining the equations (S58) and (S59). □

Lemma S7 (Some technical results for the proof of Theorem 1)**.** *Under conditions in Theorem 1, it can be shown that*

$$n\boldsymbol{\mu}_{D\mathbf{Z}}^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\mu}_{D\mathbf{Z}} \lesssim \mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2), \tag{S62}$$

$$\mathbb{E}\{(\mathbf{Z}^\top \boldsymbol{\mu}_{D\mathbf{Z}})^4\} \lesssim (\boldsymbol{\mu}_{D\mathbf{Z}}^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\mu}_{D\mathbf{Z}})^2. \tag{S63}$$

Proof. Firstly, we can proceed $n\boldsymbol{\mu}_{D\mathbf{Z}}^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\mu}_{D\mathbf{Z}}$ in the equation (S62) as

$$n\boldsymbol{\mu}_{D\mathbf{Z}}^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\mu}_{D\mathbf{Z}} \leq n\boldsymbol{\mu}_{D\mathbf{Z}}^\top \boldsymbol{\mu}_{D\mathbf{Z}} \lambda_{\max}(\boldsymbol{\Sigma}_{\mathbf{Z}}) \leq n\boldsymbol{\mu}_{D\mathbf{Z}}^\top \boldsymbol{\mu}_{D\mathbf{Z}} \mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2) \lesssim \mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}^2),$$

where the second inequality holds by the fact that the L_F norm is an upper bound on the L_2 norm, and the last inequality follows from Assumption 5. Secondly, we prove the equation (S63),

$$\begin{aligned}
\mathbb{E}\{(\mathbf{Z}^\top \boldsymbol{\mu}_{D\mathbf{Z}})^4\} &= \mathbb{E}\{(\mathbf{Z}^\top \boldsymbol{\mu}_{D\mathbf{Z}} \boldsymbol{\mu}_{D\mathbf{Z}}^\top \mathbf{Z})^2\} \\
&= \mathbb{E}\{(\boldsymbol{\nu}^\top \underline{\boldsymbol{\Phi}_{\mathbf{Z}}^\top \boldsymbol{\mu}_{D\mathbf{Z}} \boldsymbol{\mu}_{D\mathbf{Z}}^\top \boldsymbol{\Phi}_{\mathbf{Z}}} \boldsymbol{\nu})\} \\
&\lesssim \{\mathrm{tr}(\boldsymbol{\Phi}_{\mathbf{Z}}^\top \boldsymbol{\mu}_{D\mathbf{Z}} \boldsymbol{\mu}_{D\mathbf{Z}}^\top \boldsymbol{\Phi}_{\mathbf{Z}})\}^2 = (\boldsymbol{\mu}_{D\mathbf{Z}}^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\mu}_{D\mathbf{Z}})^2,
\end{aligned}$$

where we apply Lemma S1 to deduce the first inequality. □

Lemma S8 (Some technical results for the proof of Theorem 2)**.** *Under conditions in Theorem 2, it can be shown that*

$$\mathbb{E}\{(\mathbf{Z}^\top \boldsymbol{\alpha}_z)^4\} \lesssim (\boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\alpha}_z)^2, \tag{S64}$$

$$\mathbb{E}\{(\mathbf{Z}^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\alpha}_z)^4\} \lesssim (\boldsymbol{\alpha}_z \boldsymbol{\Sigma}_{\mathbf{Z}}^3 \boldsymbol{\alpha}_z)^2. \tag{S65}$$

Proof. The equation (S64) can be similarly verified as the equation (S63). Next, we prove the equation (S65),

$$\begin{aligned}
\mathbb{E}\{(\mathbf{Z}^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\alpha}_z)^4\} &= \mathbb{E}\{(\mathbf{Z}^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\alpha}_z \boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \mathbf{Z})^2\} \\
&= \mathbb{E}\{(\boldsymbol{\nu}^\top \underline{\boldsymbol{\Phi}_{\mathbf{Z}}^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\alpha}_z \boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\Phi}_{\mathbf{Z}}} \boldsymbol{\nu})\} \\
&\lesssim \{\mathrm{tr}(\boldsymbol{\Phi}_{\mathbf{Z}}^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\alpha}_z \boldsymbol{\alpha}_z^\top \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{\Phi}_{\mathbf{Z}})\}^2 = (\boldsymbol{\alpha}_z \boldsymbol{\Sigma}_{\mathbf{Z}}^3 \boldsymbol{\alpha}_z)^2,
\end{aligned}$$

where the first inequality is due to Lemma S1.

□

References

Hall, P. and Heyde, C. C. (2014). *Martingale Limit Theory and Its Application*. Academic Press, New York.

Loh, P.-L. and Wainwright, M. J. (2015). Regularized M-estimators with nonconvexity: Statistical and algorithmic theory for local optima. *Journal of Machine Learning Research*, 16(1):559–616.

Negahban, S. N., Ravikumar, P., Wainwright, M. J., and Yu, B. (2012). A unified framework for high-dimensional analysis of M-estimators with decomposable regularizers. *Statistical Science*, 27(4):538–557.

Van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer Science & Business Media, New York.

Yang, W., Guo, X., and Zhu, L. (2022). Score function-based tests for ultrahigh-dimensional linear models. *arXiv preprint arXiv:2212.08446*.