Supplemental Materials

In this section, we provide all technical proofs in this work, as well as the simulation results for the computing time and the normal case. In section S.1, Theorem 1 is proved. The lemmas used for proving Theorems 2 and 3 are stated in S.2 and technically verified in S.4. Theorems 2 and 3 are systematically proved in S.3. And the additional simulation results are reported in S.5.

S.1 Proof of Theorem 1

Let $\mathbf{u}_i = (y_i, \mathbf{x}_i^T)$, $i = 1, \cdots, n$ and denote $q = p + 1$. Thus, $\mathbf{u}_1, \cdots, \mathbf{u}_n$ be an independent and identically distributed random sample from $EC_q(\boldsymbol{\mu}, \Sigma, \phi)$. To study the asymptotic behaviors of partial correlation of elliptical distribution, we consider the following general partitions of $\mathbf{u}_i$, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$:

$$\mathbf{u}_i = \begin{pmatrix} \mathbf{u}_{1i} \\ \mathbf{u}_{2i} \end{pmatrix}, \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix},$$

where $\mathbf{u}_{1i}$ and $\boldsymbol{\mu}_1$ are q_1-dimensional, while $\mathbf{u}_{2i}$ and $\boldsymbol{\mu}_2$ are q_2-dimensional, $\boldsymbol{\Sigma}_{11}$ is a $q_1 \times q_1$ matrix, and $\boldsymbol{\Sigma}_{22}$ is a $q_2 \times q_2$ matrix. Here $q = q_1 + q_2$. Let $\mathbf{U} = (\mathbf{u}_1, \cdots, \mathbf{u}_n)^T$ and denote

$$\mathbf{A} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{u}_i - \bar{\mathbf{u}})(\mathbf{u}_i - \bar{\mathbf{u}})^T = \frac{1}{n}\mathbf{U}^T(I - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T)\mathbf{U}$$

Partition $\mathbf{A}$ in the same way as Σ. Let $a_{kl.2}$ is the (k, l)-element of $\mathbf{A}_{11.2} \hat{=} \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}$. Then the sample partial correlation of u_{ik} and u_{il} given $\mathbf{u}_{2i}$, $\hat{\rho}(u_{ik}, u_{il}|\mathbf{u}_{2i})$, indeed equals $a_{kl.2}/\sqrt{a_{kk.2}a_{ll.2}}$.

To derive the asymptotic distribution of $\mathbf{A}_{11.2}$, let

$$C = \begin{pmatrix} I & -\mathbf{\Sigma}_{12} \\ 0 & I \end{pmatrix},$$

where I stands for the identity matrix, and $\mathbf{v}_i = C(\mathbf{u}_i - \boldsymbol{\mu})$. Using Theorem 2.16 of Fang, Kotz and Ng (1990), it follows that

$$\mathbf{v}_i \sim EC_q(0, \begin{pmatrix} \mathbf{\Sigma}_{11.2} & 0 \\ 0 & \mathbf{\Sigma}_{22} \end{pmatrix}, \phi) \tag{S.1}$$

where $\mathbf{\Sigma}_{11.2} = \mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}$.

Let $\mathbf{V} = (\mathbf{v}_1, \cdots, \mathbf{v}_n)^T$. By definition of $\mathbf{v}_i$, $\mathbf{V} = (\mathbf{U} - \mathbf{1}_n\boldsymbol{\mu}^T)C^T$, where $\mathbf{1}_n$ is an $n \times 1$ vector with all elements being 1. Define

$$\mathbf{B} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{v}_i - \bar{\mathbf{v}})(\mathbf{v}_i - \bar{\mathbf{v}})^T = \frac{1}{n}\mathbf{V}^T(I - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T)\mathbf{V} = C\mathbf{A}C^T.$$

Partition $\mathbf{B}$ in the same way as $\mathbf{A}$, then $\mathbf{B}_{11} = \mathbf{A}_{11} - \mathbf{\Sigma}_{12}\mathbf{A}_{21} - \mathbf{A}_{12}\mathbf{\Sigma}_{21} + \mathbf{\Sigma}_{12}\mathbf{A}_{22}\mathbf{\Sigma}_{21}$, $\mathbf{B}_{12} = \mathbf{A}_{12} - \mathbf{\Sigma}_{12}\mathbf{A}_{22}$, $\mathbf{B}_{21} = \mathbf{A}_{21} - \mathbf{A}_{22}\mathbf{\Sigma}_{21}$ and $\mathbf{B}_{22} = \mathbf{A}_{22}$. By direct calculation, it follows that $\mathbf{B}_{11.2} \hat{=} \mathbf{B}_{11} - \mathbf{B}_{12}\mathbf{B}_{22}^{-1}\mathbf{B}_{21} = \mathbf{A}_{11.2}$. This enables us to derive the asymptotic distribution of $\mathbf{A}_{11.2}$ through $\mathbf{B}_{11.2}$.

Define $\mathbf{W}_{11} = \sqrt{n}(\mathbf{B}_{11} - \mathbf{\Sigma}_{11.2})$, $\mathbf{W}_{12} = \sqrt{n}\mathbf{B}_{12}$, $\mathbf{W}_{21} = \sqrt{n}\mathbf{B}_{21}$, and $\mathbf{W}_{22} = \sqrt{n}(\mathbf{B}_{22} - \mathbf{\Sigma}_{22})$. The assumption that all fourth moments of $\mathbf{u}_i$ are finite implies that all fourth moments of $\mathbf{v}_i$ are finite. Thus, it follows by the central limit theorem that $\mathbf{W}_{kl}$ for $k = 1, 2$ and $l = 1, 2$, has an asymptotic normal distribution with mean zero and a finite covariance matrix. Then

$$\mathbf{B}_{11.2} = \frac{1}{\sqrt{n}}\mathbf{W}_{11} + \mathbf{\Sigma}_{11.2} - \frac{1}{n}\mathbf{W}_{12}(\mathbf{\Sigma}_{22} + \frac{1}{\sqrt{n}}\mathbf{W}_{22})^{-1}\mathbf{W}_{21}.$$

By the assumptions of Theorem 1, the largest eigenvalue of $(\mathbf{\Sigma}_{22} + \frac{1}{\sqrt{n}}\mathbf{W}_{22})^{-1}$ is positive and finite. Therefore it follows that if $q_2 = o_(\sqrt{n})$,

$$\begin{aligned}\sqrt{n}(\mathbf{A}_{11.2} - \mathbf{\Sigma}_{11.2}) &= \sqrt{n}(\mathbf{B}_{11.2} - \mathbf{\Sigma}_{11.2}) = \mathbf{W}_{11} + o_P(1) \\ &= \sqrt{n}(\mathbf{B}_{11} - \Sigma_{11.2}) + o_P(1).\end{aligned}$$

This implies that $\mathbf{A}_{11.2}$ and $\mathbf{B}_{11}$ have the same asymptotic normal distribution, and hence $a_{kl.2}/\sqrt{a_{kk.2}a_{ll.2}}$ and $b_{kl}/\sqrt{b_{kk}b_{ll}}$ have the same asymptotic distribution, where b_{kl} is the (k,l)-element of $\mathbf{B}_{11}$. Further notice that $\mathbf{v}_{1i} \sim EC_{q_1}(0, \mathbf{\Sigma}_{11.2}), \phi)$ by (S.1), where $\mathbf{v}_{1i}$ consists of the first q_1 elements of $\mathbf{v}_i$. Therefore, the asymptotic normal distribution of the sample correlation coefficient $\hat{\rho}(v_{ik}, v_{il})$, which indeed equals to $b_{kl}/\sqrt{b_{kk}b_{ll}}$, can be derived from (2.2) with replacing Σ by $\Sigma_{11.2}$. Thus, Theorem 1 holds by setting $\mathbf{u}_{1i} = (y_i, x_{i\mathcal{S}^c}^T)^T$ and $\mathbf{u}_{2i} = \mathbf{x}_{i\mathcal{S}}$.

S.2 Lemmas

In this section, we introduce the following lemmas which are used repeatedly in the proofs of Theorems 2 and 3.

Lemma 1. *(Hoeffding's Inequality) Assume the independent random sample* $\{X_i : i = 1, \cdots, n\}$ *satisfies* $P(X_i \in [a_i, b_i]) = 1$ *for some* a_i *and* b_i, $\forall i =$

$1, \cdots, n$. *Then, for any* $\epsilon > 0$*, the sample mean* $\bar{X}$ *satisfies*

$$P(|\bar{X} - E(\bar{X})| > \epsilon) \leq 2\exp\left(-\frac{2\epsilon^2 n^2}{\sum_{i=1}^n (b_i - a_i)^2}\right). \tag{S.2}$$

Lemma 2. *Suppose* X *is a random variable with* $E(e^{a|X|}) < \infty$ *for some* $a > 0$*. Then for any* $M > 0$*, there exist positive constants* b *and* c *such that*

$$P(|X| \geq M) \leq be^{-cM}. \tag{S.3}$$

Lemma 3. *Suppose* $\hat{\gamma}_1$ *and* $\hat{\gamma}_2$ *are estimates of the finite parameters* γ_1 *and* γ_2 *based on a size-n sample, respectively. Assume there exist positive constants* b_1*,* b_2 *and* ν *such that for any* $0 < \epsilon < 1$*,*

$$P\{|\hat{\gamma}_j - \gamma_j| > \epsilon\} \quad \leq \quad b_j \exp(-n^\nu / b_j), \quad j = 1, 2. \tag{S.4}$$

Then

$$\begin{aligned} P\left\{|(\hat{\gamma}_1 - \hat{\gamma}_2) - (\gamma_1 - \gamma_2)| > \epsilon\right\} &\leq b_3 \exp(-n^\nu / b_3), \\ P\left\{|\hat{\gamma}_1\hat{\gamma}_2 - \gamma_1\gamma_2| > \epsilon\right\} &\leq b_4 \exp(-n^\nu / b_4), \end{aligned}$$

where $b_3 = b_1 + b_2$*, and* $b_4 = 2b_1 + b_2$*. If* $\gamma_2 \neq 0$*,*

$$P\left\{\left|\frac{\hat{\gamma}_1}{\hat{\gamma}_2} - \frac{\gamma_1}{\gamma_2}\right| > \epsilon\right\} \quad \leq \quad b_5 \exp(-n^\nu / b_5),$$

where $b_5 = b_1 + 3b_2$. *If we further assume* $\gamma_2 > 0$, *then*

$$\begin{aligned} P\left\{|\sqrt{\hat{\gamma}_2} - \sqrt{\gamma_2}| > \epsilon\right\} &\leq b_6 \exp(-n^{\nu}/b_6), \\ P\left\{|\log\hat{\gamma}_2 - \log\gamma_2| > \epsilon\right\} &\leq b_2 \exp(-n^{\nu}/b_2), \end{aligned}$$

where $b_6 = 2b_2$.

The proof of Lemma 3 can be found at section S.4 of the supplemental materials.

S.3 Proof of Theorems 2 and 3

For ease of notation, denote $\overline{x}_j = \frac{1}{n}\sum_{i=1}^n x_{ij}$, $\overline{y} = \frac{1}{n}\sum_{i=1}^n y_i$, $\overline{x_j y} = \frac{1}{n}\sum_{i=1}^n x_{ij}y_i$, $\overline{x_j^2} = \frac{1}{n}\sum_{i=1}^n x_{ij}^2$, and $\overline{y^2} = \frac{1}{n}\sum_{i=1}^n y_i^2$. Then

$$\hat{\rho}(y, x_j) = \frac{\overline{x_j y} - \overline{x}_j\overline{y}}{\sqrt{(\overline{x_j^2} - \overline{x}_j^2)(\overline{y^2} - \overline{y}^2)}} \tag{S.5}$$

The proof of Theorem 2. We divide the proof into three parts.

Step 1: Study the consistency of $\hat{Z}_n(y, x_j|x_{\mathcal{S}})/\sqrt{1+\hat{\kappa}}$. First consider $\bar{x}_j$. For any $0 < \epsilon < 1$ and any $M > 0$,

$$\begin{aligned} P(|\bar{x}_j - \mathrm{E}x_j| > \epsilon) &\leq P(|\bar{x}_j - \mathrm{E}x_j| > \epsilon, \max_{1\leq i\leq n}|x_{ij}| \leq M) + P(\max_{1\leq i\leq n}|x_{ij}| > M) \\ &\leq P(|\bar{x}_j - \mathrm{E}x_j| > \epsilon, \max_{1\leq i\leq n}|x_{ij}| \leq M) + nP(|x_{ij}| > M) \\ &\leq 2\exp(-\frac{n\epsilon^2}{2M^4}) + nC_2\exp(-C_1 M) \end{aligned} \tag{S.6}$$

for some positive constants C_1 and C_2. The first term above is obtained by Hoeffding's inequality in Lemma 1, and the second term is by condition (D2) and Lemma 2. Take $M = O(n^{1/5})$, then for large n, (S.6) is simplified as $P(|\bar{x}_j - \mathrm{E}x_j| > \epsilon) \le C_3 \exp(-n^\nu/C_3)$, where $0 < \nu < 1/5$ and $C_3 > 0$. In the same fashion, there exist some positive constants C_4, C_5, C_6 and C_7, such that for large n,

$$\begin{aligned}
&P(|\bar{y} - \mathrm{E}y| > \epsilon) \le C_4 \exp(-n^\nu/C_4),\\
&P(|\overline{x_j^2} - \mathrm{E}(x_j^2)| > \epsilon) \le C_5 \exp(-n^\nu/C_5),\\
&P(|\overline{y^2} - \mathrm{E}(y^2)| > \epsilon) \le C_6 \exp(-n^\nu/C_6),\\
&P(|\overline{x_j y} - \mathrm{E}(x_j y)| > \epsilon) \le C_7 \exp(-n^\nu/C_7).
\end{aligned}$$

Therefore by (S.5) and Lemma 3,

$$P\{|\hat{\rho}(y, x_j) - \rho(y, x_j)| > \epsilon\} \le C_8 \exp(-n^\nu/C_8),$$

where the positive constant C_8 is determined by $C_3, \ldots, C_7$.

Note that

$$\rho(y, x_j|x_{\mathcal{S}}) = \frac{\rho(y, x_j|x_{\mathcal{S}\setminus\{k\}}) - \rho(y, x_k|x_{\mathcal{S}\setminus\{k\}})\rho(x_j, x_k|x_{\mathcal{S}\setminus\{k\}})}{[\{1 - \rho^2(y, x_k|x_{\mathcal{S}\setminus\{k\}})\}\{1 - \rho^2(x_j, x_k|x_{\mathcal{S}\setminus\{k\}})\}]^{1/2}}, \tag{S.7}$$

for any $k \in \mathcal{S}$.

Under the bounded condition (D5), applying Lemma 3 to the sample version of (S.7) and the Z-transformation (2.4) recursively, we conclude that for some $C_9 > 0$ and $C_{10} > 0$,

$$P\left\{\left|\hat{\rho}(y, x_j|x_{\mathcal{S}}) - \rho(y, x_j|x_{\mathcal{S}})\right| > \epsilon\right\} \leq C_9 \exp(-n^{\nu}/C_9), \quad \text{and}$$
$$P\left\{\left|\hat{Z}(y, x_j|x_{\mathcal{S}}) - Z(y, x_j|x_{\mathcal{S}})\right| > \epsilon\right\} \leq C_{10} \exp(-n^{\nu}/C_{10}).$$

Furthermore, by the same argument, the sample kurtosis is consistent to the population version with the same rate, that is, there exists $C_{11} > 0$ such that $P\{|\hat{\kappa} - \kappa| > \epsilon\} \leq C_{11} \exp(-n^{\nu}/C_{11})$, and hence for some $C_{12} > 0$,

$$P\left\{\left|\frac{\hat{Z}(y, x_j|x_{\mathcal{S}})}{\sqrt{1+\hat{\kappa}}} - \frac{Z(y, x_j|x_{\mathcal{S}})}{\sqrt{1+\kappa}}\right| > \epsilon\right\} \leq C_{12} \exp(-n^{\nu}/C_{12}).$$

Step 2: Compute $P(E_{j|\mathcal{S}}) = P\{$an error occurs when testing $\rho(y, x_j|x_{\mathcal{S}}) = 0\}$. Denote $E_{j|\mathcal{S}} = E^I_{j|\mathcal{S}} \cup E^{II}_{j|\mathcal{S}}$, where $E^I_{j|\mathcal{S}}$ is the event that the type I error occurs and $E^{II}_{j|\mathcal{S}}$ is the event that the type II error occurs. Then by choosing $\alpha_n = 2\{1 - \Phi(\frac{c_n}{2}\sqrt{\frac{n}{1+\kappa}})\}$,

$$\begin{aligned}
P(E^I_{j|\mathcal{S}}) &= P\left\{\left|\frac{\hat{Z}(y, x_j|x_{\mathcal{S}})}{\sqrt{1+\hat{\kappa}}}\right| > \frac{\Phi^{-1}(1-\alpha_n/2)}{(n-|\mathcal{S}|-1)^{1/2}} \text{ when } Z(y, x_j|x_{\mathcal{S}}) = 0\right\} \\
&\leq P\left\{\left|\frac{\hat{Z}(y, x_j|x_{\mathcal{S}})}{\sqrt{1+\hat{\kappa}}} - \frac{Z(y, x_j|x_{\mathcal{S}})}{\sqrt{1+\kappa}}\right| > \frac{\Phi^{-1}(1-\alpha_n/2)}{(n-|\mathcal{S}|-1)^{1/2}}\right\} \\
&= P\left\{\left|\frac{\hat{Z}(y, x_j|x_{\mathcal{S}})}{\sqrt{1+\hat{\kappa}}} - \frac{Z(y, x_j|x_{\mathcal{S}})}{\sqrt{1+\kappa}}\right| > \frac{c_n\sqrt{n}}{2}\{(n-|\mathcal{S}|-1)(1+\kappa)\}^{-1/2}\right\} \\
&\leq P\left\{\left|\frac{\hat{Z}(y, x_j|x_{\mathcal{S}})}{\sqrt{1+\hat{\kappa}}} - \frac{Z(y, x_j|x_{\mathcal{S}})}{\sqrt{1+\kappa}}\right| > \frac{c_n}{2\sqrt{1+\kappa}}\right\} \\
&\leq C_{12} \exp(-n^{\nu}/C_{12}),
\end{aligned}$$

and

$$
\begin{aligned}
P(E^{II}_{j|\mathcal{S}}) &= P\left\{\left|\frac{\hat{Z}(y,x_j|x_{\mathcal{S}})}{\sqrt{1+\hat{\kappa}}}\right| \le \frac{\Phi^{-1}(1-\alpha_n/2)}{(n-|\mathcal{S}|-1)^{1/2}} \text{ when } Z(y,x_j|x_{\mathcal{S}}) \ne 0\right\} \\
&\le P\left\{\left|\frac{Z(y,x_j|x_{\mathcal{S}})}{\sqrt{1+\kappa}}\right| - \left|\frac{\hat{Z}(y,x_j|x_{\mathcal{S}})}{\sqrt{1+\hat{\kappa}}} - \frac{Z(y,x_j|x_{\mathcal{S}})}{\sqrt{1+\kappa}}\right| \le \frac{\Phi^{-1}(1-\alpha_n/2)}{(n-|\mathcal{S}|-1)^{1/2}}\right\} \\
&\le P\left\{\left|\frac{\hat{Z}(y,x_j|x_{\mathcal{S}})}{\sqrt{1+\hat{\kappa}}} - \frac{Z(y,x_j|x_{\mathcal{S}})}{\sqrt{1+\kappa}}\right| \ge \left|\frac{Z(y,x_j|x_{\mathcal{S}})}{\sqrt{1+\kappa}}\right| - \frac{c_n}{2\sqrt{1+\kappa}}\right\}
\end{aligned}
$$

Note that $|g(u)| = |\frac{1}{2}\log\{(1+u)/(1-u)\}| \ge |u|$ for all $u \in (-1,1)$, then $|Z(y,x_j|x_{\mathcal{S}})| \ge |\rho_n(y,x_j|x_{\mathcal{S}})| \ge c_n$ under condition (D4). Thus,

$$
\begin{aligned}
P(E^{II}_{j|\mathcal{S}}) &\le P\left\{\left|\frac{\hat{Z}(y,x_j|x_{\mathcal{S}})}{\sqrt{1+\hat{\kappa}}} - \frac{Z(y,x_j|x_{\mathcal{S}})}{\sqrt{1+\kappa}}\right| \ge \frac{c_n}{\sqrt{1+\kappa}} - \frac{c_n}{2\sqrt{1+\kappa}}\right\} \\
&= P\left\{\left|\frac{\hat{Z}(y,x_j|x_{\mathcal{S}})}{\sqrt{1+\hat{\kappa}}} - \frac{Z(y,x_j|x_{\mathcal{S}})}{\sqrt{1+\kappa}}\right| \ge \frac{c_n}{2\sqrt{1+\kappa}}\right\} \\
&\le C_{12}\exp(-n^{\nu}/C_{12}).
\end{aligned}
$$

Therefore, $P(E_{j|\mathcal{S}}) = P(E^{I}_{j|\mathcal{S}}) + P(E^{II}_{j|\mathcal{S}}) \le 2C_{12}\exp(-n^{\nu}/C_{12})$.

Step 3: Study $P\{\hat{\mathcal{A}}_n(\alpha_n) = \mathcal{A}\}$. Now consider all $j = 1, \cdots, p$ and all $\mathcal{S} \subseteq \{j\}^c$ subject to $|\mathcal{S}| \le m_n$, where $m_n \le \hat{m}_{reach}$. Define $K_j^{m_n} = \{\mathcal{S} \subseteq$

$\{j\}^c, |\mathcal{S}| \leq m_n\}$, $j = 1, \cdots, p$.

$$\begin{aligned} P\{\hat{\mathcal{A}}_n(\alpha_n) \neq \mathcal{A}\} &= P\{\text{an error occurs for some } j \text{ and some } \mathcal{S}\} \\ &= P\left\{ \bigcup_{j=1,\cdots,p_n; \mathcal{S} \in K_j^{m_n}} E_{j|\mathcal{S}} \right\} \leq \sum_{j=1,\cdots,p_n; \mathcal{S} \in K_j^{m_n}} P(E_{j|\mathcal{S}}) \\ &\leq p^{m_n+1} \cdot \sup_{j=1,\cdots,p_n; \mathcal{S} \in K_j^{m_n}} P(E_{j|\mathcal{S}}) \leq 2p^{m_n+1} C_{12} \exp(-n^\nu / C_{12}) \\ &\leq 2p^{d_0+1} C_{12} \exp(-n^\nu / C_{12}). \end{aligned} \tag{S.8}$$

The second inequality holds since the number of possible choices of j is p and there are p^{m_n} possible choices for $\mathcal{S}$. The last inequality in (S.8) is obtained because $P(\hat{m}_{reach} = m_{reach}) \to 1$ and $m_{reach} \leq d_0$ by the same technique as Lemma 3. Thus for large n, $m_n \leq \hat{m}_{reach} \leq d_0$.

Moreover, recall that ν can be chosen arbitrarily in $(0, 1/5)$. Therefore, if d_0 is fixed, for $p = o(\exp(n^\xi))$, $0 < \xi < 1/5$, (S.8) is simplified as $P\{\hat{\mathcal{A}}_n(\alpha_n) \neq \mathcal{A}\} \leq O\{\exp(-n^\nu / C_{12})\}$, provided $\xi < \nu < 1/5$. If $d_0 = O(n^b)$, $0 < b < 1/5$, for $p = o(\exp(n^\xi))$, $0 < \xi < 1/5 - b$, (S.8) becomes $P\{\hat{\mathcal{A}}_n(\alpha_n) \neq \mathcal{A}\} \leq O\{\exp(-n^\nu / C_{12})\}$, provided that $\xi + b < \nu < 1/5$. This completes the proof of Theorem 2 with $C = C_{12}$.

Proof of Theorem 3. We only need to consider the first step of the thresholded partial correlation approach, where we have

$$P\left(\left| \frac{\hat{Z}(y, x_j)}{\sqrt{1 + \hat{\kappa}}} - \frac{Z(y, x_j)}{\sqrt{1 + \kappa}} \right| > \epsilon\right) \leq C_{13} \exp(-n^\nu / C_{13})$$

for some $C_{13} > 0$. Define $E_j^{II} = \{\text{fail to include } x_j \text{ when } x_j \text{ is a true predictor}\}$, then using the same technique as the proof of Theorem 2,

$$\begin{aligned} P(E_j^{II}) &= \left\{ \left| \frac{\hat{Z}_n(y, x_j)}{\sqrt{1+\hat{\kappa}}} \right| \leq \frac{\Phi^{-1}(1-\alpha_n/2)}{(n-1)^{1/2}} \text{ when } \beta_j \neq 0 \right\} \\ &\leq P\left\{ \left| \frac{\hat{Z}_n(y, x_j)}{\sqrt{1+\hat{\kappa}}} - \frac{Z_n(y, x_j)}{\sqrt{1+\kappa}} \right| \geq \frac{c_n}{2\sqrt{1+\kappa}} \right\} \\ &\leq C_{13} \exp(-n^{\nu}/C_{13}). \end{aligned}$$

Then

$$P\{\hat{\mathcal{A}}_n^{[1]} \not\supseteq \mathcal{A}\} = P\left\{ \bigcup_{j=1}^{p} E_j^{II} \right\} \leq \sum_{j=1}^{p} P(E_j^{II}) \leq pC_{13} \exp(-n^{\nu}/C_{13}),$$

for any $0 < \nu < 1/5$. Therefore, for $p = o(\exp(n^{\xi}))$ and $0 < \xi < 1/5$, $P\{\hat{\mathcal{A}}_n^{[1]} \not\supseteq \mathcal{A}\} \leq O\{\exp(-n^{\nu}/C_{13})\}$, provided $\xi < \nu < 1/5$. This completes the proof of Theorem 3 with $C^* = C_{13}$.

S.4 Proof of Lemma 3

The first inequality is easy to obtain by

$$\begin{aligned} P\{|(\hat{\gamma}_1 - \hat{\gamma}_2) - (\gamma_1 - \gamma_2)| > \epsilon\} &\leq P\{|\hat{\gamma}_1 - \gamma_1| > \epsilon/2\} + P\{|\hat{\gamma}_2 - \gamma_2| > \epsilon/2\} \\ &\leq b_3 \exp(-n^{\nu}/b_3), \end{aligned}$$

where $b_3 = b_1 + b_2$. To study $\hat{\gamma}_1\hat{\gamma}_2$, we first show that $\hat{\gamma}_1$ and $\hat{\gamma}_2$ are bounded in probability. Denote $M_1 = \max\{|\gamma_1| + 1/2, |\gamma_2| + 1/2\}$, then

$$\begin{aligned} P\{|\hat{\gamma}_1| > M_1\} &\le P\{|\hat{\gamma}_1 - \gamma_1| + |\gamma_1| > M_1\} \\ &\le P\{|\hat{\gamma}_1 - \gamma_1| > 1/2\} \le b_1 \exp(-n^{\nu}/b_1). \end{aligned} \tag{S.9}$$

Similarly, $P\{|\hat{\gamma}_2| > M_1\} \le b_2 \exp(-n^{\nu}/b_2)$. Then for any $0 < \epsilon < 1$,

$$\begin{aligned} & P\left\{|\hat{\gamma}_1\hat{\gamma}_2 - \gamma_1\gamma_2| > \epsilon\right\} \\ = \;& P\left\{|\hat{\gamma}_1\hat{\gamma}_2 - \hat{\gamma}_1\gamma_2 + \hat{\gamma}_1\gamma_2 - \gamma_1\gamma_2| > \epsilon\right\} \\ \le \;& P\left\{|\hat{\gamma}_1| \cdot |\hat{\gamma}_2 - \gamma_2| > \epsilon/2\right\} + P\left\{|\gamma_2| \cdot |\hat{\gamma}_1 - \gamma_1| > \epsilon/2\right\} \\ \le \;& P\left\{|\hat{\gamma}_1| \cdot |\hat{\gamma}_2 - \gamma_2| > \epsilon/2, |\hat{\gamma}_1| \le M_1\right\} + P\{|\hat{\gamma}_1| > M_1\} + P\left\{|\hat{\gamma}_1 - \gamma_1| > \epsilon/(2M_1)\right\} \\ \le \;& P\left\{|\hat{\gamma}_2 - \gamma_2| > \epsilon/(2M_1)\right\} + P\{|\hat{\gamma}_1| > M_1\} + P\left\{|\hat{\gamma}_1 - \gamma_1| > \epsilon/(2M_1)\right\} \end{aligned}$$

Thus by (S.4) and (S.9), $P\left\{|\hat{\gamma}_1\hat{\gamma}_2 - \gamma_1\gamma_2| > \epsilon\right\} \le b_4 \exp(-n^{\nu}/b_4)$, where $b_4 = 2b_1 + b_2$.

Now consider $\hat{\gamma}_1/\hat{\gamma}_2$ when $\gamma_2 \ne 0$. Note that $\hat{\gamma}_2$ is bounded away from 0 with probability tending to 1. This is because $P(|\hat{\gamma}_2| < |\gamma_2|/2) = P(|\gamma_2 - (\gamma_2 - \hat{\gamma}_2)| < |\gamma_2|/2) \le P(|\hat{\gamma}_2 - \gamma_2| > |\gamma_2|/2) \le b_2 \exp(-n^{\nu}/b_2)$, which tends to

0. Then

$$\begin{aligned}
& P\left\{\left|\frac{\hat\gamma_1}{\hat\gamma_2}-\frac{\gamma_1}{\gamma_2}\right|>\epsilon\right\} \\
\leq\; & P\left\{\left|\frac{\hat\gamma_1}{\hat\gamma_2}-\frac{\gamma_1}{\hat\gamma_2}\right|>\epsilon/2\right\}+P\left\{\left|\frac{\gamma_1}{\hat\gamma_2}-\frac{\gamma_1}{\gamma_2}\right|>\epsilon/2\right\} \\
\leq\; & P\left\{|\hat\gamma_1-\gamma_1|>\frac{\epsilon|\hat\gamma_2|}{2}\right\}+P\left\{\left|\frac{\gamma_1}{\gamma_2\hat\gamma_2}\right|\cdot|\hat\gamma_2-\gamma_2|>\epsilon/2\right\} \\
\leq\; & P\left\{|\hat\gamma_1-\gamma_1|>\frac{\epsilon|\hat\gamma_2|}{2},|\hat\gamma_2|\geq\frac{|\gamma_2|}{2}\right\}+P\left\{|\hat\gamma_2-\gamma_2|>\frac{\epsilon|\gamma_2\hat\gamma_2|}{2|\gamma_1|},|\hat\gamma_2|\geq\frac{|\gamma_2|}{2}\right\}+2P\left\{|\hat\gamma_2|<\frac{|\gamma_2|}{2}\right\} \\
\leq\; & P\left\{|\hat\gamma_1-\gamma_1|>\frac{\epsilon|\gamma_2|}{4}\right\}+P\left\{|\hat\gamma_2-\gamma_2|>\frac{\epsilon\gamma_2^2}{4|\gamma_1|}\right\}+2P\left\{|\hat\gamma_2|<\frac{|\gamma_2|}{2}\right\} \\
\leq\; & b_1\exp(-n^\nu/b_1)+b_2\exp(-n^\nu/b_2)+2b_2\exp(-n^\nu/b_2).
\end{aligned}$$

Therefore, $P\left\{|\hat\gamma_1/\hat\gamma_2-\gamma_1/\gamma_2|>\epsilon\right\}\leq b_5\exp(-n^\nu/b_5)$, where $b_5=b_1+3b_2$.

If further assume $\gamma_2>0$, using the same technique as above,

$$\begin{aligned}
& P\left\{|\sqrt{\hat\gamma_2}-\sqrt{\gamma_2}|>\epsilon\right\} \\
\leq\; & P\left\{\frac{|\hat\gamma_2-\gamma_2|}{\sqrt{\hat\gamma_2}+\sqrt{\gamma_2}}>\epsilon,|\hat\gamma_2|\geq\frac{\gamma_2}{2}\right\}+P\left\{|\hat\gamma_2|<\frac{\gamma_2}{2}\right\} \\
\leq\; & P\left\{|\hat\gamma_2-\gamma_2|>\epsilon\sqrt{\gamma_2}(1+\frac{1}{\sqrt{2}})\right\}+P\left\{|\hat\gamma_2|<\frac{\gamma_2}{2}\right\}.
\end{aligned}$$

Thus $P\left\{|\sqrt{\hat\gamma_2}-\sqrt{\gamma_2}|>\epsilon\right\}\leq b_6\exp(-n^\nu/b_6)$, where $b_6=2b_2$.

At last, since $\hat\gamma_2$ is consistent with γ_2, we can apply Taylor's expansion to $\log\hat\gamma_2$, i.e. $\log\hat\gamma_2=\log\gamma_2+(\hat\gamma_2-\gamma_2)/\gamma_2+o_p(\hat\gamma_2-\gamma_2)$. Thus for large n,

$$P\left\{|\log\hat\gamma_2-\log\gamma_2|>\epsilon\right\}\leq P\left\{\frac{2}{\gamma_2}|\hat\gamma_2-\gamma_2|>\epsilon\right\}\leq P\left\{|\hat\gamma_2-\gamma_2|>\delta'''\right\},$$

where $\delta'''=\min\{\epsilon,\epsilon\gamma_2/2\}$. Therefore, $P\left\{|\log\hat\gamma_2-\log\gamma_2|>\epsilon\right\}\leq b_2\exp(-n^\nu/b_2)$.

Table S1: Computational time (in minutes) of 1000 simulations when $p = 2000$

ρ	SCAD	LASSO	PC-simple	TPC
0	344.31	31.37	349.84	27.90
0.3	341.52	14.14	245.88	28.01
0.8	288.43	214.18	218.22	29.42

S.5 Additional Simulation Results

This section provides additional simulation results. Table S1 depicts the computing time of 1000 simulation with $p = 2000$ when data were generated from an elliptical distribution. Table S2 depicts the results from the normal linear model in the simulation examples. Table S3 reports the simulation results when data were generated a population in which x's with even subscripts were generated in the same fashion as that for elliptical distribution, and x's with odd subscripts take discrete values 0, 1 and 2 with probabilities 0.25, 0.5 and 0.25, respectively. In this simulation study, we take $\rho = 0.3$ low correlation and $\rho = 0.8$ for high correlation.

Table S2: Simulation Results for Example 1: Normal Distribution

p	ρ	Method	MedME(Devi)	TPN	TFN	UF	CF	OF
		SCAD	0.013 (0.006)	3.00	1.04	0.00	0.66	0.34
		LASSO	8.936 (0.148)	3.00	16.91	0.00	0.01	0.99
200	0	PC-simple	0.012 (0.006)	3.00	0.03	0.00	0.97	0.03
		TPC	0.012 (0.006)	3.00	0.03	0.00	0.97	0.03
		SCAD	0.014 (0.006)	3.00	0.86	0.00	0.73	0.27
		LASSO	11.105 (0.151)	3.00	15.60	0.00	0.02	0.98
200	0.3	PC-simple	0.011 (0.006)	3.00	0.00	0.00	1.00	0.00
		TPC	0.011 (0.006)	3.00	0.01	0.00	0.99	0.01
		SCAD	0.010 (0.006)	3.00	0.67	0.00	0.72	0.28
		LASSO	20.731 (0.069)	3.00	9.52	0.00	0.03	0.97
200	0.8	PC-simple	0.009 (0.006)	2.92	0.10	0.08	0.92	0.00
		TPC	0.009 (0.006)	2.92	0.10	0.08	0.92	0.00
		SCAD	0.013 (0.008)	3.00	1.26	0.00	0.77	0.23
		LASSO	9.046 (0.121)	3.00	20.75	0.00	0.02	0.98
500	0	PC-simple	0.014 (0.008)	3.00	0.14	0.00	0.87	0.13
		TPC	0.014 (0.008)	3.00	0.15	0.00	0.86	0.14
		SCAD	0.014 (0.007)	3.00	1.33	0.00	0.72	0.28
		LASSO	11.231 (0.101)	3.00	19.07	0.00	0.00	1.00
500	0.3	PC-simple	0.013 (0.007)	3.00	0.07	0.00	0.93	0.07
		TPC	0.013 (0.008)	3.00	0.10	0.00	0.90	0.10
		SCAD	0.011 (0.007)	3.00	0.92	0.00	0.71	0.29
		LASSO	20.777 (0.085)	3.00	11.74	0.00	0.02	0.98
500	0.8	PC-simple	0.012 (0.008)	2.86	0.18	0.14	0.86	0.00
		TPC	0.012 (0.008)	2.87	0.16	0.13	0.87	0.00
		SCAD	0.013 (0.007)	3.00	2.25	0.00	0.66	0.34
		LASSO	9.080 (0.120)	3.00	31.21	0.00	0.00	1.00
2000	0	PC-simple	0.023 (0.017)	3.00	0.41	0.00	0.62	0.38
		TPC	0.022 (0.016)	3.00	0.41	0.00	0.62	0.38
		SCAD	0.010 (0.006)	3.00	1.65	0.00	0.69	0.31
		LASSO	11.277 (0.129)	3.00	26.97	0.00	0.00	1.00
2000	0.3	PC-simple	0.010 (0.006)	3.00	0.12	0.00	0.89	0.11
		TPC	0.010 (0.006)	3.00	0.13	0.00	0.88	0.12
		SCAD	0.011 (0.006)	3.00	1.32	0.00	0.66	0.34
		LASSO	20.828 (0.098)	3.00	15.50	0.00	0.03	0.97
2000	0.8	PC-simple	0.011 (0.007)	2.90	0.11	0.10	0.90	0.00
		TPC	0.011 (0.007)	2.90	0.11	0.10	0.90	0.00

$*$ The numbers in the parentheses are median absolute deviations over 1000 simulations.

Table S3: Simulation Results for Elliptical Distribution with half x's being discrete

p	ρ	Method	MedME(Devi)	TPN	FPN	UF	CF	OF
		SCAD	1.040 (0.892)	3.00	8.52	0.00	0.13	0.87
		LASSO	11.209 (0.220)	3.00	20.40	0.00	0.00	1.00
200	0.3	PC-simple	0.218 (0.054)	3.00	0.47	0.00	0.59	0.41
		TPC	0.187 (0.053)	3.00	0.14	0.00	0.87	0.13
		SCAD	0.114 (0.058)	3.00	4.71	0.00	0.22	0.78
		LASSO	20.619 (0.172)	3.00	16.53	0.00	0.00	1.00
200	0.8	PC-simple	0.090 (0.042)	2.97	0.33	0.03	0.72	0.25
		TPC	0.091 (0.038)	2.96	0.17	0.04	0.84	0.12
		SCAD	1.425 (1.160)	3.00	14.11	0.00	0.08	0.92
		LASSO	11.252 (0.221)	3.00	33.78	0.00	0.00	1.00
500	0.3	PC-simple	0.222 (0.063)	3.00	0.67	0.00	0.45	0.55
		TPC	0.198 (0.054)	2.99	0.32	0.01	0.69	0.30
		SCAD	0.119 (0.049)	3.00	8.48	0.00	0.13	0.87
		LASSO	20.659 (0.199)	3.00	24.43	0.00	0.00	1.00
500	0.8	PC-simple	0.099 (0.037)	2.99	0.37	0.01	0.69	0.30
		TPC	0.096 (0.033)	3.00	0.21	0.00	0.80	0.20
		SCAD	1.584 (0.893)	3.00	25.60	0.00	0.02	0.98
		LASSO	11.383 (0.219)	3.00	55.84	0.00	0.00	1.00
2000	0.3	PC-simple	0.295 (0.067)	3.00	1.51	0.00	0.11	0.89
		TPC	0.234 (0.066)	3.00	0.67	0.00	0.44	0.56
		SCAD	0.183 (0.079)	3.00	18.16	0.00	0.02	0.98
		LASSO	20.720 (0.168)	3.00	38.64	0.00	0.00	1.00
2000	0.8	PC-simple	0.127 (0.048)	2.98	0.74	0.02	0.41	0.57
		TPC	0.116 (0.037)	2.99	0.36	0.01	0.69	0.30