SURE INDEPENDENCE SCREENING ADJUSTED FOR CONFOUNDING COVARIATES WITH ULTRAHIGH DIMENSIONAL DATA

Canhong Wen[1], Wenliang Pan[1], Mian Huang[2], and Xueqin Wang[1,*]

[1]*Sun Yat-Sen University and* [2]*Shanghai University of Finance and Economics*

Supplementary Material

This supplementary file presents additional simulation results that are not included in the main article due to page limit. We also present details of the technique proof in the main paper.

S1 Additional simulation Results

Figures **??**-**??** summary the comparison results in Example 1 with compound symmetric (CS) covariance matrix of X when $p = 5000$. Figure **??** gives the boxplots of $\log(\mathcal{S})$ and Figures **??** presents the proportion of $\mathcal{P}_d$ with $d = \gamma\lceil n/\log(n)\rceil$ against γ.

*To whom correspondence should be addressed.

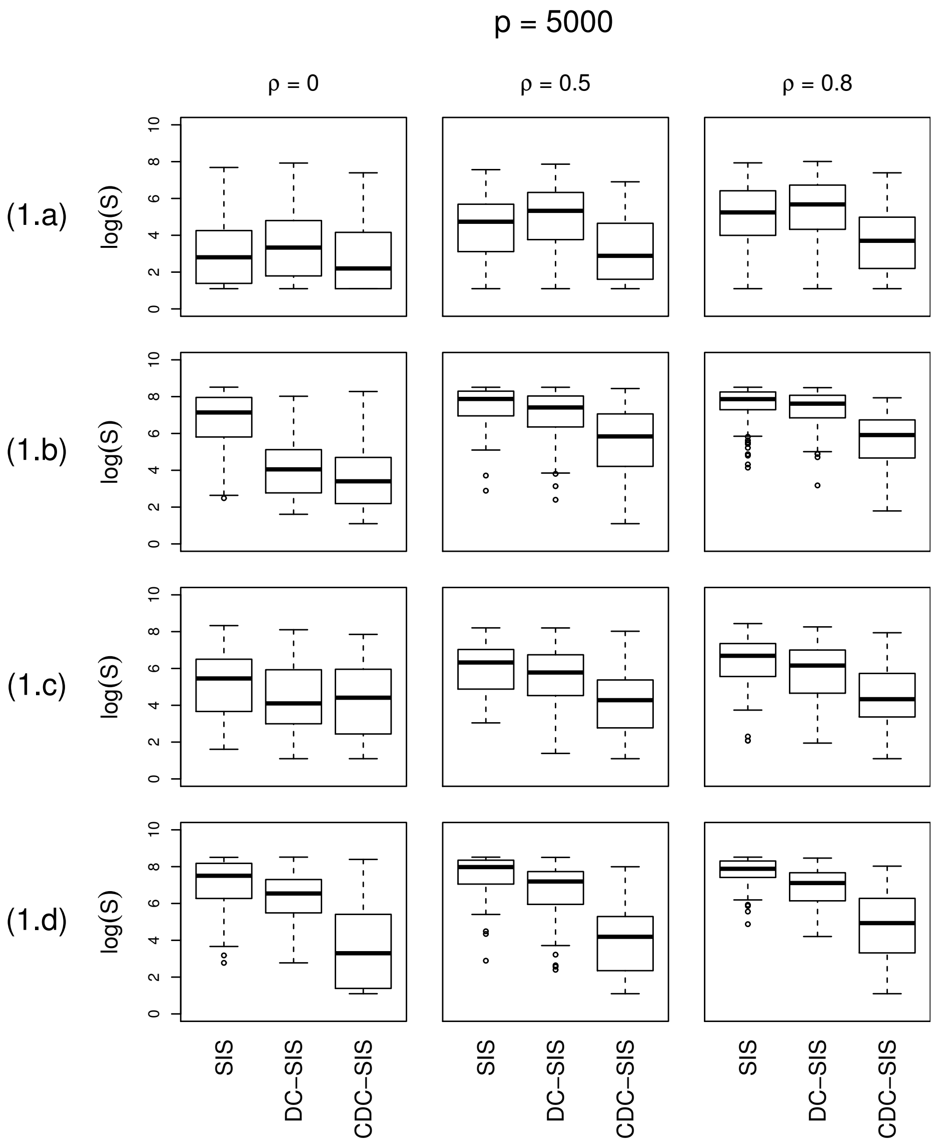

Figure S1: Example 1: Boxplots of $\log(\mathcal{S})$ for the SIS, DC-SIS and CDC-SIS methods for different values of ρ based on 100 replications under different models with $p = 5000$ and compound symmetric (CS) covariance matrix.

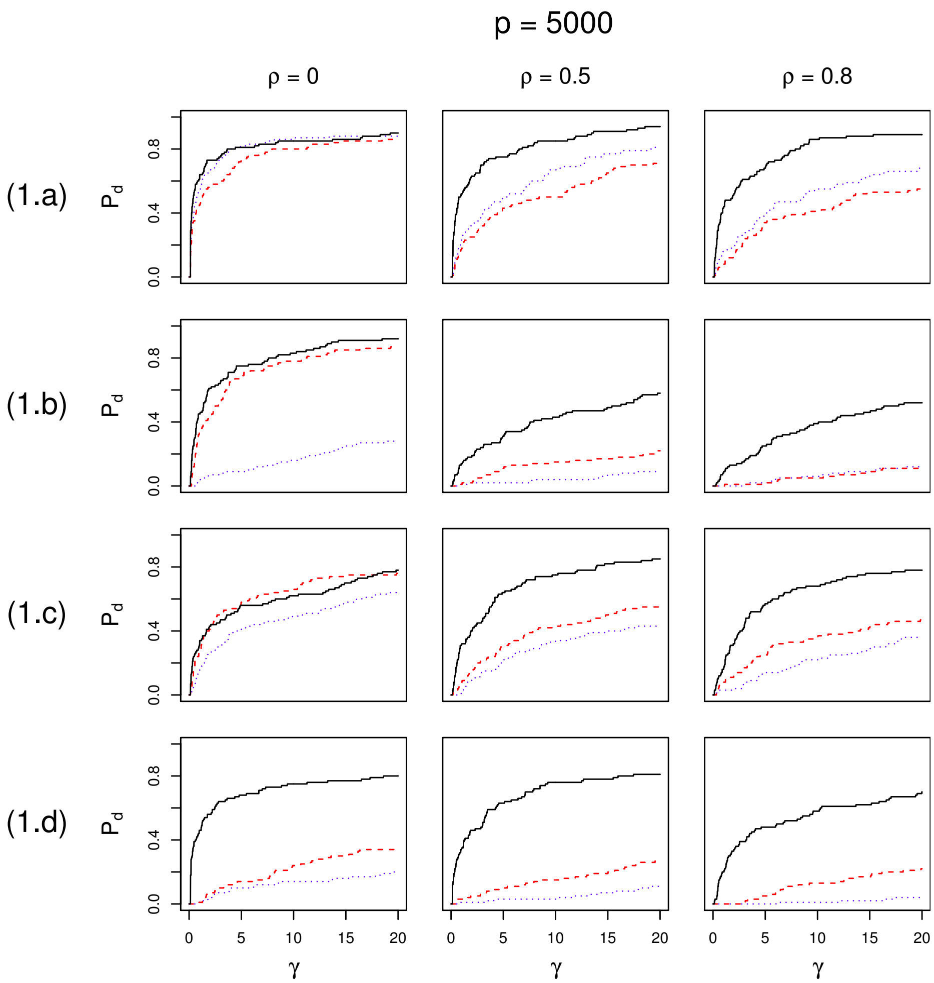

Figure S2: Examples 1: Summary results of the proportion of $\mathcal{P}_d^{\gamma}$ for the SIS (dotted line), DC-SIS (dashed line) and CDC-SIS (solid line) methods for different values of ρ based on 100 replications under different models with $p = 5000$ and compound symmetric (CS) covariance matrix.

Table **??** and Figures **??**-**??** summary the comparison results in Example 1 with first-order autoregressive (AR) covariance matrix of X. Table **??** shows the median value of the minimum model size S. Figures **??**-**??** give the boxplots of $\log(\mathcal{S})$ and Figures **??**-**??** present the proportion of $\mathcal{P}_d$ with $d = \gamma\lceil n/\log(n)\rceil$ against γ.

Figures **??**-**??** summary the comparison results in Example 2 when $p = 5000$. Figure **??** gives the boxplots of $\log(\mathcal{S})$ and Figures **??** presents the proportion of $\mathcal{P}_d$ with $d = \gamma\lceil n/\log(n)\rceil$ against γ.

Table S1: Example 1: Median of the minimum model size S for the SIS, DC-SIS and CDC-SIS methods for different values of p and ρ based on 100 replications under different models with first-order autoregressive (AR) covariance matrix.

Model	p	ρ	SIS	DC-SIS	CDC-SIS
(1.a)	1000	0	4	6	4
		0.5	3	3	3
		0.8	5	5	4
	5000	0	15	28	11
		0.5	4	5	3
		0.8	5	5	5
(1.b)	1000	0	338	14	8
		0.5	230	20	5
		0.8	5	5	4
	5000	0	1953	69	35
		0.5	1497	82	27
		0.8	5	5	5
(1.c)	1000	0	74	33	17
		0.5	30	13	5
		0.8	5	5	4
	5000	0	406	152	61
		0.5	114	52	12
		0.8	5	5	5
(1.d)	1000	0	396	147	6
		0.5	302	52	3
		0.8	6	5	3
	5000	0	1850	807	16
		0.5	1700	366	3
		0.8	12	5	4

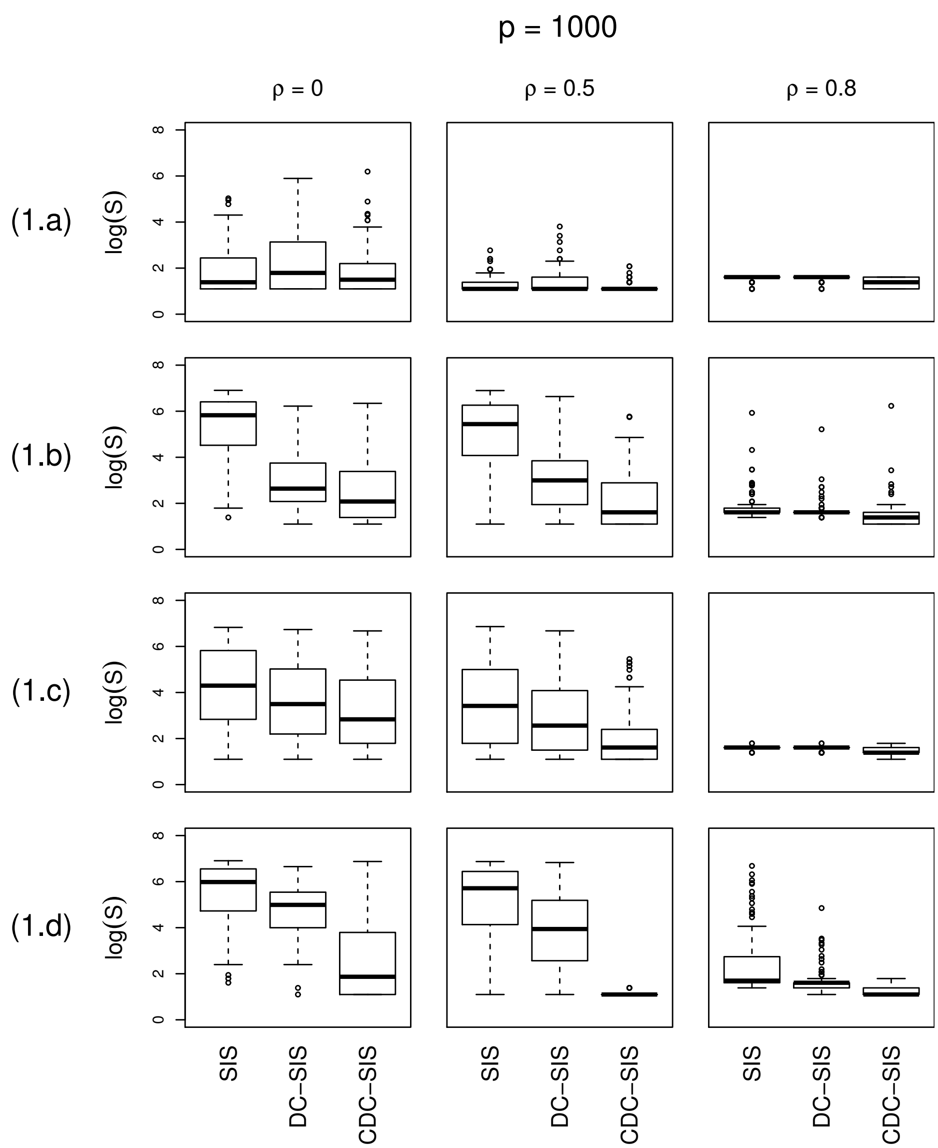

Figure S3: Example 1: Boxplots of $\log(\mathcal{S})$ for the SIS, DC-SIS and CDC-SIS methods for different values of ρ based on 100 replications under different models with $p = 1000$ and first-order autoregressive (AR) covariance matrix.

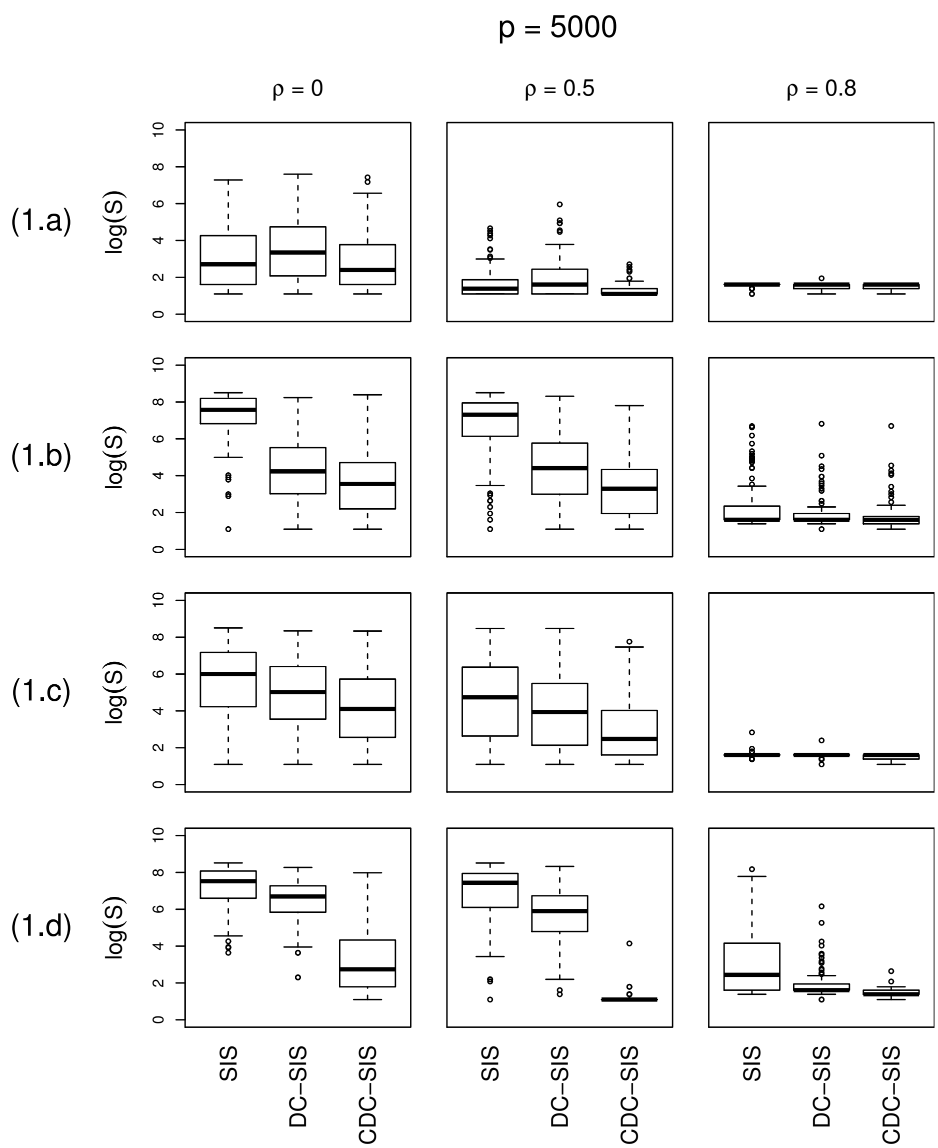

Figure S4: Example 1: Boxplots of $\log(\mathcal{S})$ for the SIS, DC-SIS and CDC-SIS methods for different values of ρ based on 100 replications under different models with $p = 5000$ and first-order autoregressive (AR) covariance matrix.

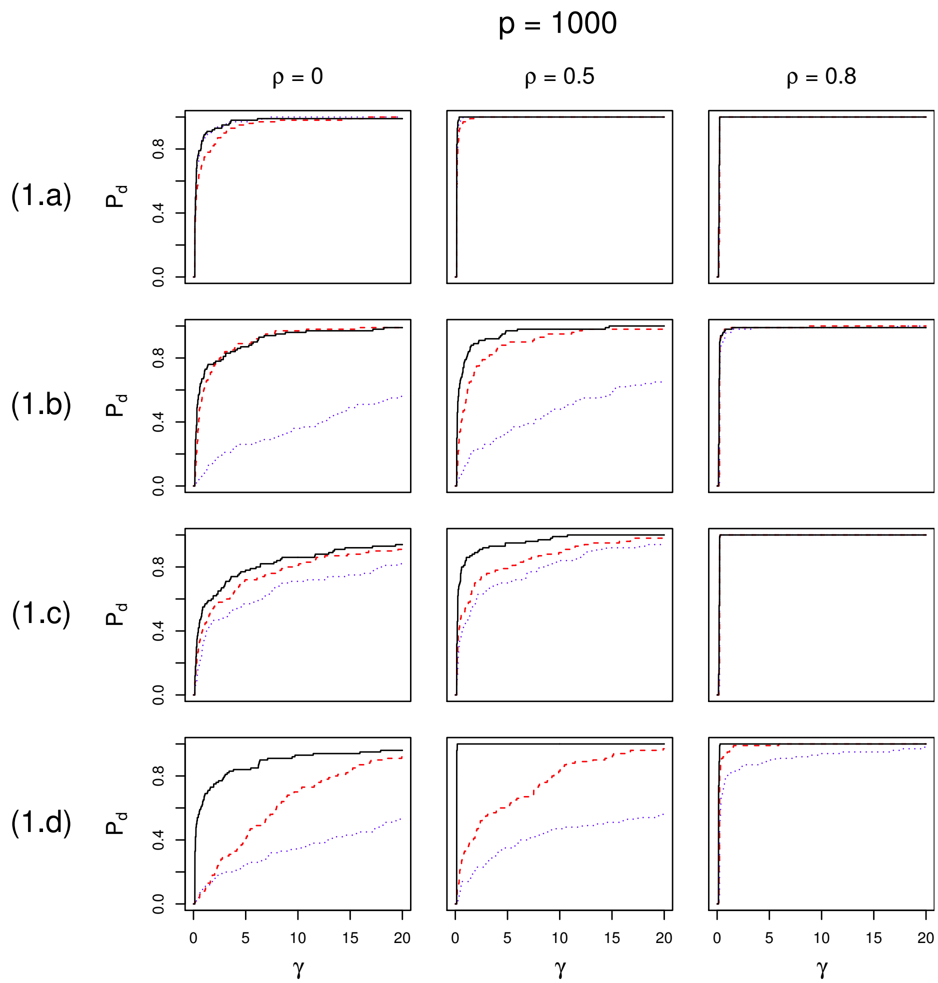

Figure S5: Examples 1: Summary results of the proportion of $\mathcal{P}_d$ for the SIS (dotted line), DC-SIS (dashed line) and CDC-SIS (solid line) methods for different values of ρ based on 100 replications under different models with $p = 1000$ and first-order autoregressive (AR) covariance matrix.

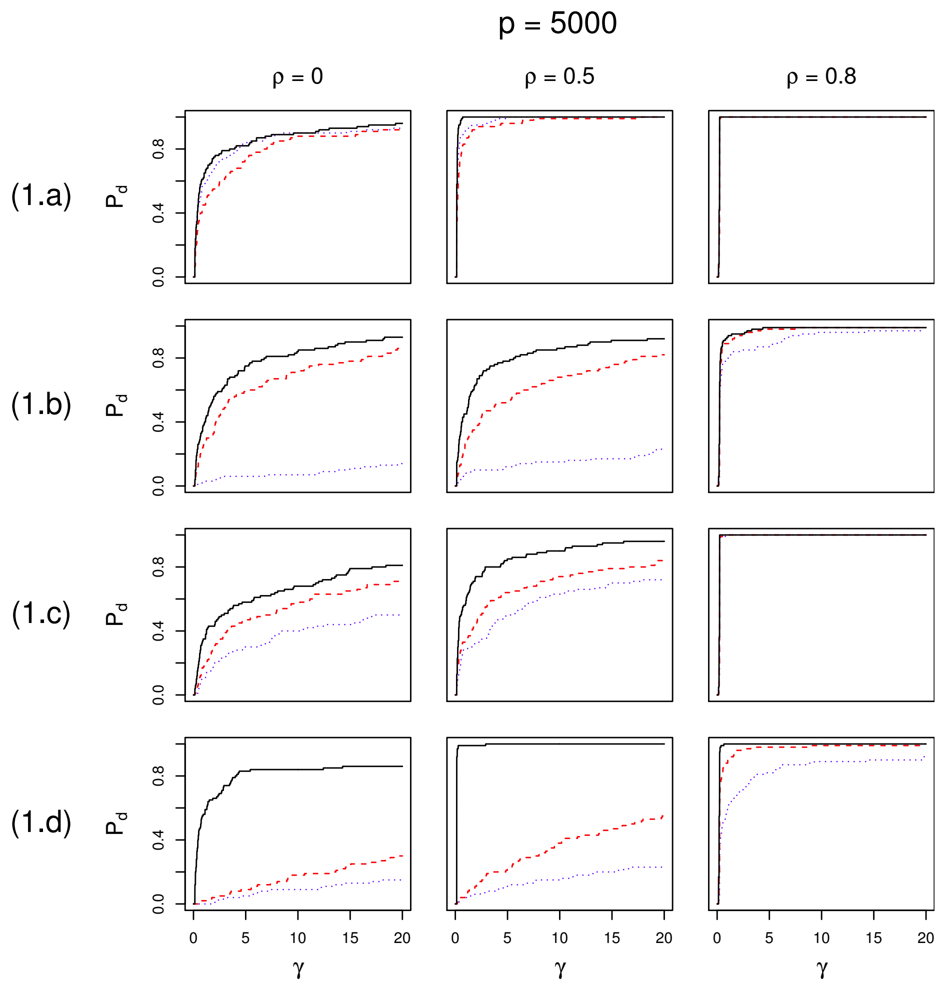

Figure S6: Examples 1: Summary results of the proportion of $\mathcal{P}_d$ for the SIS (dotted line), DC-SIS (dashed line) and CDC-SIS (solid line) methods for different values of ρ based on 100 replications under different models with $p = 1000$ and first-order autoregressive (AR) covariance matrix.

Table S2: Example 1: Accuracy of SIS, ISIS, DC-SIS, CDC-SIS in including the true model $\{X_1, X_2, X_5\}$ for different values of ρ and p with first–order autoregressive (AR) covariance matrix.

Model	p	ρ	SIS	ISIS	DC-SIS	CDC-SIS
(1.a)	1000	0	0.87	1	0.73	0.88
		0.5	1	1	0.97	1
		0.8	1	0.97	1	1
	5000	0	0.57	0.85	0.45	0.62
		0.5	0.92	1	0.84	1
		0.8	1	1	1	1
(1.b)	1000	0	0.08	0.08	0.6	0.68
		0.5	0.13	0.17	0.52	0.79
		0.8	0.96	0.45	0.99	0.98
	5000	0	0.03	0.02	0.26	0.38
		0.5	0.08	0.09	0.28	0.45
		0.8	0.79	0.4	0.89	0.93
(1.c)	1000	0	0.34	0.52	0.41	0.55
		0.5	0.45	0.59	0.55	0.84
		0.8	1	0.82	1	1
	5000	0	0.1	0.21	0.18	0.35
		0.5	0.28	0.39	0.33	0.57
		0.8	1	0.87	1	1
(1.d)	1000	0	0.10	0.12	0.08	0.65
		0.5	0.14	0.14	0.33	1
		0.8	0.8	0.41	0.95	1
	5000	0	0	0.01	0.02	0.56
		0.5	0.04	0.07	0.05	0.99
		0.8	0.59	0.19	0.88	1

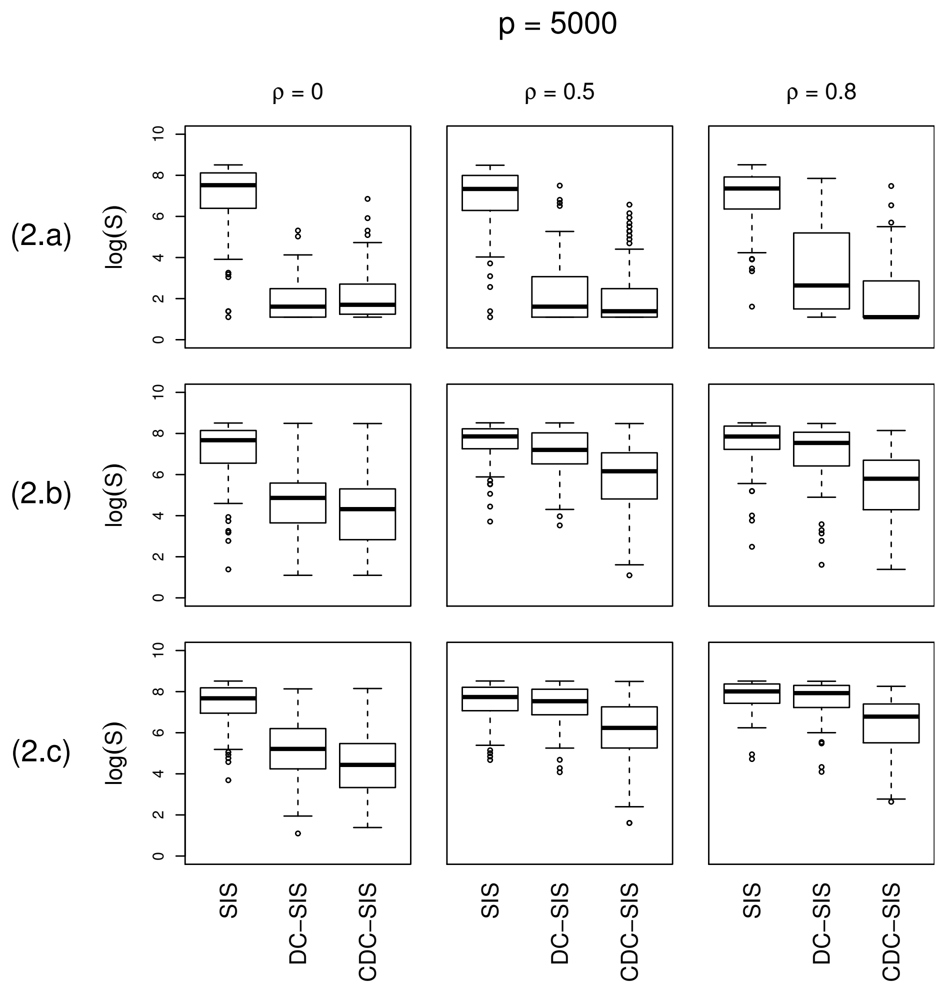

Figure S7: Example 2: Boxplots of $\log(\mathcal{S})$ for the SIS, DC-SIS and CDC-SIS methods for different values of ρ based on 100 replications under different models with $p = 5000$.

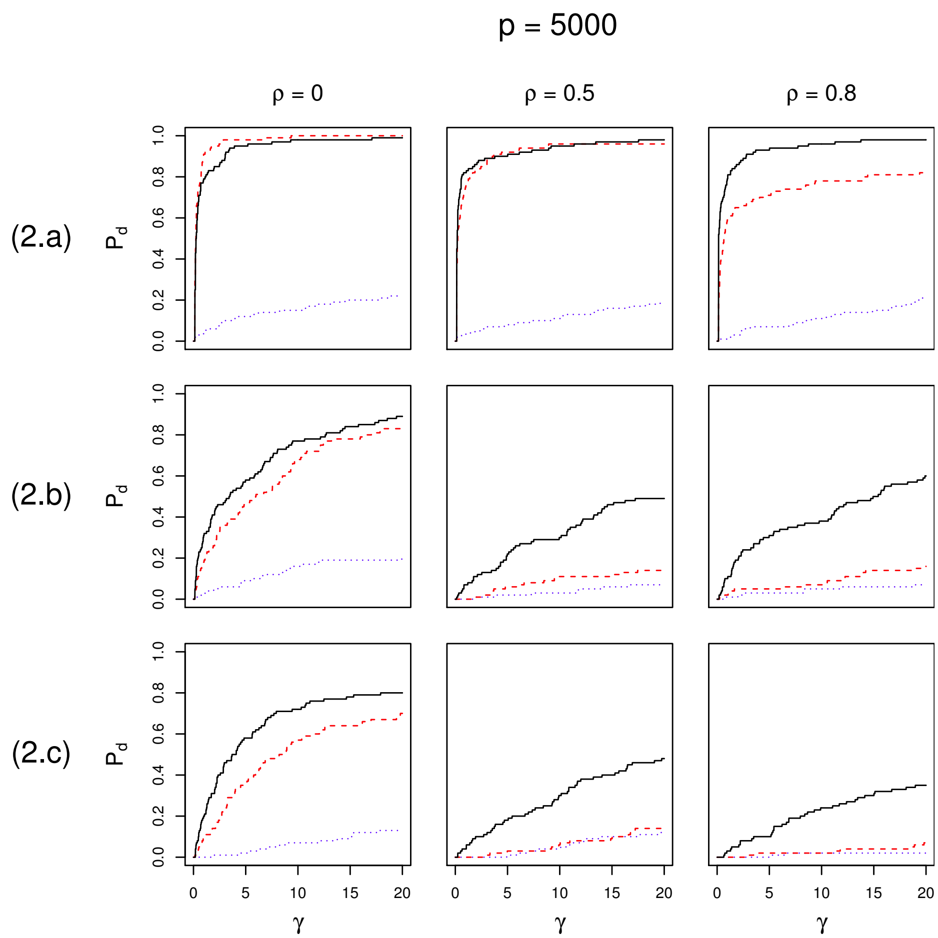

Figure S8: Examples 2: Summary results of the proportion of $\mathcal{P}_d^\gamma$ for the SIS (dotted line), DC-SIS (dashed line) and CDC-SIS (solid line) methods for different values of ρ based on 100 replications under different models with $p = 5000$.

S2 Technique proof

Before proving Theorem 3.1 in the main paper, let us first revisit the conditions needed.

(C1): The kernel function $K(\cdot)$ is bounded uniformly such that $K(u) \geq 0$, $\int K(u)du = 1$, $\int uK(u)du = 0$, and $\int \|u\|^2 K(u)du < \infty$.

(C2): There exists a positive constant s_0 such that for all $0 < s \leq s_0$,

$$\sup_p \max_{1\leq r\leq p} E(\exp(s\|X_r\|_p^2)) < \infty, \quad E(\exp(s\|Y\|_{q_y}^2)) < \infty,$$

where p and q_y are the dimensions of the predictor X_r and the response variable Y, respectively.

(C3): If Z_1, Z_2, Z_3, Z_4 are independent copies of Z, then for $1 \leq r \leq p$, there exists a positive constant L, such that

$$\sup_r |E(d_{1234,r}|Z_1, Z_2, Z_3, Z_4) - E(d_{1234,r}|Z_1', Z_2, Z_3, Z_4)| \leq L|Z_1 - Z_1'|.$$

(C4): There exist some constants $c > 0$ and $0 \leq \kappa < 1/2$ such that

$$\min_{r\in\mathcal{A}} \rho_r^* \geq 2cn^{-\kappa}.$$

S3 Some lemmas

This section introduces three useful lemmas which are needed in the proof of the sure screening property. Lemma **??** is extracted from Lemma 5.6.1.A of **?** and we omit its proof here.

Lemma 1 *Let $\mu = E(Y)$. If $pr(a \leq Y \leq b) = 1$, then*

$$E[\exp\{s(Y-\mu)\}] \leq \exp\{s^2(b-a)^2/8\}, \text{ for any } s > 0.$$

Lemma 2 *Under Condition (C**??**), for any s satisfies $0 < s \leq s_0/32$, we have*

$$\sup_p \max_{1\leq r\leq p} (\exp(s|d_{1234,r}|)) < \infty.$$

Proof With the Holder's inequality, we can show that

$$E(\exp(s|d_{1234,r}|)) \leq [E\{\exp(16s|X_{r1}-X_{r2}||Y_1-Y_2|)\}E\{\exp(16s|X_{r1}-X_{r2}||Y_3-Y_4|)\}]^{1/4}$$
$$[E\{\exp(16s|X_{r1}-X_{r2}||Y_1-Y_3|)\}]^{1/2}.$$

For the last term in the right hand size, we can bound it by

$$
\begin{aligned}
&E(\exp(16s|X_{r1}-X_{r2}||Y_1-Y_2|))\\
\le& E[\exp\{16s(|X_{r1}-X_{r2}|^2+|Y_1-Y_2|^2)/2\}]\\
\le& [E\{\exp(16s|X_{r1}-X_{r2}|^2)\}]^{1/2}[E\{\exp(16s|Y_1-Y_2|^2)\}]^{1/2}\\
\le& E[\exp(32sX_{r1}^2)]E[\exp(32sY_1^2)].
\end{aligned}
$$

Following similar arguments, we can obtain

$$
\begin{aligned}
E(\exp(16s|X_{r1}-X_{r2}||Y_3-Y_4|)) &\le E(\exp(32sX_{r1}^2))E(\exp(32sY_1^2))\\
E(\exp(16s|X_{r1}-X_{r2}||Y_1-Y_3|)) &\le E(\exp(32sX_{r1}^2))E(\exp(32sY_1^2)).
\end{aligned}
$$

Therefore,

$$E(\exp(s|d_{1234,r}|)) \le E(\exp(32sX_{r1}^2))E(\exp(32sY_1^2)).$$

Then, under Condition (C**??**),

$$\sup_p \max_{1\le r\le p} E(\exp(s|d_{1234,r}|)) < \infty.$$

Lemma 3 *Under Condition (C**??**), for any s satisfies $0 < s \le s_0/32$, we have*

$$\sup_p \max_{1\le r\le p} E[\exp(s|E(d_{1234,r} \mid Z_1, Z_2, Z_3, Z_4)|)] < \infty.$$

Proof According to the Jensen's inequality, we have

$$\exp(s|E(d_{1234,r} \mid Z_1, Z_2, Z_3, Z_4)|) \le E(\exp(s|d_{1234,r}| \mid Z_1, Z_2, Z_3, Z_4)).$$

This together with Lemma **??** suggests that

$$E[\exp(s|E(d_{1234,r} \mid Z_1, Z_2, Z_3, Z_4)|)] \leq E(\exp(s|d_{1234,r}|)) < \infty.$$

S4 Proof of Theorem 3.1

For notation simplicity, we define $h_r(z) = \mathrm{CDCov}^2(X_r, Y \mid Z = z)$, $\hat{h}_r(z) = \widehat{\mathrm{CDCov}}^2(\mathbf{X}_r, \mathbf{Y}, \mathbf{Z} \mid Z = z)$, $K_i = K_H(z - Z_i)(i = 1, \ldots, n)$, $K = \sum_{i=1}^n K_H(z - Z_i)$ and $g_r(Z_1, Z_2, Z_3, Z_4) = E(d_{1234,r} \mid Z_1, Z_2, Z_3, Z_4)$, where Z_1, Z_2, Z_3, Z_4 are independent copies of Z. Furthermore, we suppose that $K_i \leq C_0$ and $f(z) \leq C_0'$, where C_0 and C_0' are positive constants. Throughout the proof, the notation C is a generic constant which may take different values at each appearance. The s denotes an arbitrary positive number in the interval $(0, s_0/64]$. We consider the decomposition

$$d_{1234,r} = E(d_{1234,r} \mid Z_1, Z_2, Z_3, Z_4) + e_{1234,r} = g_r(Z_1, Z_2, Z_3, Z_4) + e_{1234,r},$$

where $e_{1234,r}$ satisfies $E(e_{1234,r} \mid Z_1, Z_2, Z_3, Z_4) = 0$.

Before proving Theorem 3.1, we first show that $\sup_p \max_{1 \leq r \leq p} E(\exp(s|e_{1234,r}|)) <$

∞. Using the results from Lemma **??** and Lemma **??**, we have

$$
\begin{aligned}
&\sup_p \max_{1\le r\le p} E(s|e_{1234,r}|)\\
=&\sup_p \max_{1\le r\le p} E[\exp\{s(|d_{1234,r} - g_r(Z_1, Z_2, Z_3, Z_4)|)\}]\\
\le&\sup_p \max_{1\le r\le p} [E\{\exp(2s|d_{1234,r}|)\}]^{1/2}[E\{\exp(2s|g_r(Z_1, Z_2, Z_3, Z_4)|)\}]^{1/2}\\
<&\infty.
\end{aligned}
\tag{S4.1}
$$

Our aim is to show the uniform consistency of the denominator and the numerator of $\hat{\rho}_r$ under regularity conditions. Because the denominator of $\hat{\rho}_r$ has a similar form as the numerator, we only consider its numerator below. We decomposed the $\hat{h}_r(z)$ into three parts as follows:

$$
\begin{aligned}
\hat{h}_r(z) &= \sum_{i,j,k,l} d_{ijkl,r} \frac{K_i K_j K_k K_l}{K^4}\\
&= \sum_{i,j,k,l} \frac{K_i K_j K_k K_l}{K^4} \left(g_r(Z_i, Z_j, Z_k, Z_l) + e_{ijkl,r}\right)\\
&= \sum_{i,j,k,l} \frac{K_i K_j K_k K_l}{K^4} h_r(z) + \sum_{i,j,k,l} \frac{K_i K_j K_k K_l}{K^4} \left(g_r(Z_i, Z_j, Z_k, Z_l) - h_r(z)\right)\\
&\quad + \sum_{i,j,k,l} \frac{K_i K_j K_k K_l}{K^4} e_{ijkl,r},
\end{aligned}
$$

Subtracting $h_r(z)$ from both sides, we have

$$
\hat{h}_r(z) - h_r(z) = \sum_{i,j,k,l} \frac{K_i K_j K_k K_l}{K^4}(g_r(Z_i, Z_j, Z_k, Z_l) - h_r(z)) + \sum_{i,j,k,l} \frac{K_i K_j K_k K_l}{K^4} e_{ijkl,r}.
$$

Thus

$$\begin{aligned}
&\mathrm{pr}\big(|\frac{K^4}{n^4}\sum_{i,j,k,l} d_{ijkl,r}\frac{K_iK_jK_kK_l}{K^4} - \frac{K^4}{n^4}E(d_{ijkl,r} \mid z)| \geq \varepsilon\big) \\
=&\mathrm{pr}(|\frac{K^4}{n^4}(\hat{h}_r(z) - h_r(z))| \geq \varepsilon) \\
\leq&\mathrm{pr}(|\frac{1}{n^4}\sum_{i,j,k,l} K_iK_jK_kK_l(g_r(Z_i, Z_j, Z_k, Z_l) - h_r(z))| \geq \varepsilon/2) \\
&+ \mathrm{pr}(|\frac{1}{n^4}\sum_{i,j,k,l} K_iK_jK_kK_l e_{ijkl,r}| \geq \varepsilon/2).
\end{aligned} \tag{S4.2}$$

Next we will obtain the convergence rate of the two terms separably in (**??**). We first deal with the second term in (**??**), which can be decomposed as

$$\begin{aligned}
\frac{1}{n^4}\sum_{i,j,k,l} K_iK_jK_kK_l e_{ijkl,r} =& \frac{1}{n^4}\sum_{i\neq j\neq k\neq l} K_iK_jK_kK_l e_{ijkl,r} + \frac{1}{n^4}\sum_{i\neq j} 3K_i^2K_jK_k e_{ijkl,r} \\
&+ \frac{1}{n^4}\sum_{i\neq j}(2K_i^2K_j^2 e_{ijkl,r} + 2K_i^3K_j e_{ijkl,r}) + \frac{1}{n^4}\sum_{i} K_i^4 e_{ijkl,r}.
\end{aligned} \tag{S4.3}$$

For the first term in (**??**), it is easy to show that

$$\mathrm{pr}(|\frac{1}{n^4}\sum_{i\neq j\neq k\neq l} K_iK_jK_kK_l e_{ijkl,r}| \geq \varepsilon) \leq \mathrm{pr}(|\frac{1}{A_n^4}\sum_{i\neq j\neq k\neq l} K_iK_jK_kK_l e_{ijkl,r}| \geq \varepsilon).$$

Denote

$$\begin{aligned}
A_1 &= \{\forall i, j, k, l, 1 \leq i, j, k, l \leq n, |K_iK_jK_kK_l e_{ijkl,r}| \leq M\}, \\
B_1 &= \{\exists i, j, k, l, 1 \leq i, j, k, l \leq n, |K_iK_jK_kK_l e_{ijkl,r}| > M\}, \\
S_1 &= \frac{1}{A_n^4}\sum_{i\neq j\neq k\neq l} K_iK_jK_kK_l e_{ijkl,r}.
\end{aligned}$$

Note that the intersection of A_1 and B_1 is empty set. Then

$$\begin{aligned}\mathrm{pr}(|S_1| \geq \varepsilon) =& \mathrm{pr}(|S_1| \geq \varepsilon \mid A_1)\mathrm{pr}(A_1) + \mathrm{pr}(|S_1| \geq \varepsilon \mid B_1)\mathrm{pr}(B_1), \\ \leq& \mathrm{pr}(|S_1| \geq \varepsilon | A_1) + \mathrm{pr}(B_1).\end{aligned}$$

With the Markov's inequality, we can obtain that

$$\mathrm{pr}(|S_1| \geq \varepsilon \mid A_1) = 2\mathrm{pr}(S_1 \geq \varepsilon \mid A_1) \leq 2\exp(-t\varepsilon)E(\exp(tS_1) \mid A_1).$$

According to U-statistic representation theorem in **?**, S_1 can be represented as $(n!)^{-1}\sum_{n!}\Omega(W_1,\ldots,W_n)$, where $\Omega(W_1,\ldots,W_n) = m^{-1}\sum_{i=1}^m \varphi_i$, $m = [n/4]$. Since the exponential function is convex and $E(K_iK_jK_kK_le_{ijkl,r}) = 0$, we have

$$\begin{aligned}E(\exp(tS_1) \mid A_1) =& E[\exp(\frac{t}{n!}\sum_{n!}\Omega(W_1,\ldots,W_n)) \mid A_1] \\ \leq& \frac{1}{n!}\sum_{n!} E[\exp(t\Omega(W_1,\ldots,W_n)) \mid A_1] \\ =& E(\exp(\frac{t}{m}\sum_{i=1}^m \varphi_i) \mid A_1) \\ =& \{E[\exp(\frac{t}{m}\varphi_1) \mid A_1]\}^m,\end{aligned}$$

which together with Lemma **??** entails that

$$\begin{aligned}E(\exp(tS_1) \mid A_1) \leq& \exp\left(\frac{(2M)^2t^2}{8m}\right) \\ =& \exp\left(\frac{t^2M^2}{2m}\right),\end{aligned}$$

where M will be specified later. Then $\mathrm{pr}(|S_1| \geq \varepsilon \mid A_1) \leq 2\exp(-t\varepsilon + t^2M^2/2m)$. By choosing $t = \varepsilon m/M^2$, we can obtain that

$$\mathrm{pr}(|S_1| \geq \varepsilon \mid A_1) \geq 2\exp\left(-\frac{m\varepsilon^2}{2M^2}\right). \tag{S4.4}$$

For the term $\mathrm{pr}(B_1)$, using the result in (**??**),

$$\begin{aligned}\mathrm{pr}(B_1) \leq & A_n^4 \max_{1\leq i,j,k,l\leq n} \mathrm{pr}\left(|e_{ijkl,r}| > M/C_0^4\right) \\ \leq & 2A_n^4 \exp(-sM/C_0^4)E[\exp(s|e_{ijkl,r}|)] \\ \leq & 2Cn^4 \exp(-sM/C_0^4).\end{aligned} \tag{S4.5}$$

Combining the results from (**??**) and (**??**), we can obtain that

$$\mathrm{pr}(|S_1| \geq \varepsilon) \leq 2Cn^4 \exp(-sM/C_0^4) + 2\exp(-m\varepsilon^2/2M^2).$$

Let $M = n^\gamma$ for some $0 < \gamma < 1/2$. Together with $m = [n/4] > n/5$, we can show that

$$\mathrm{pr}(|S_1| \geq \varepsilon) \leq 2Cn^4 \exp(-sn^\gamma/C_0^4) + 2\exp(-n^{1-2\gamma}\varepsilon^2/10). \tag{S4.6}$$

For the last two terms in (**??**), we have similar results:

$$\mathrm{pr}(|n^{-4}\sum_{i\neq j} K_i^2K_jK_k e_{ijkl,r}| \geq \varepsilon) \leq 2Cn^4\exp(-sn^\gamma/C_0^4) + 2\exp(-n^{3-2\gamma}\varepsilon^2/10),$$

$$\mathrm{pr}(|n^{-4}\sum_{i\neq j} K_i^2K_j^2 e_{ijkl,r}| \geq \varepsilon) \leq 2Cn^4\exp(-sn^\gamma/C_0^4) + 2\exp(-n^{5-2\gamma}\varepsilon^2/10),$$

$$\mathrm{pr}(|n^{-4}\sum_{i\neq j} K_i^3K_j e_{ijkl,r}| \geq \varepsilon) \leq 2Cn^4\exp(-sn^\gamma/C_0^4) + 2\exp(-n^{5-2\gamma}\varepsilon^2/10),$$

$$\mathrm{pr}(|n^{-4}\sum_{i} K_i^4 e_{ijkl,r}| \geq \varepsilon) \leq 2Cn^4\exp(-sn^\gamma/C_0^4) + 2\exp(-n^{7-2\gamma}\varepsilon^2/10).$$

Therefore, we can conclude that

$$\mathrm{pr}(|\frac{1}{n^4}\sum_{i,j,k,l} K_iK_jK_kK_l e_{ijkl,r}| \geq \varepsilon) \leq 2n^4 C\exp(-sn^{\gamma}/C_0^4) + 2\exp(-n^{1-2\gamma}\varepsilon^2/10) + o(1). \tag{S4.7}$$

In the sequel, we turn to $\mathrm{pr}(|n^{-4}\sum_{i,j,k,l} K_iK_jK_kK_l(g_r(Z_i,Z_j,Z_k,Z_l) - h_r(z))| \geq \varepsilon)$. We can decompose it into

$$\begin{aligned} n^{-4}&\sum_{i,j,k,l} K_iK_jK_kK_l(g_r(Z_i,Z_j,Z_k,Z_l) - h_r(z)) \\ &= n^{-4}\sum_{i\neq j\neq k\neq l} K_iK_jK_kK_l(g_r(Z_i,Z_j,Z_k,Z_l) - h_r(z)) \\ &\quad + n^{-4}\sum_{i\neq j\neq k} 3K_i^2K_jK_k(g_r(Z_i,Z_j,Z_k,Z_i) - h_r(z)) \\ &\quad + n^{-4}\sum_{i\neq j} 2K_i^2K_j^2(g_r(Z_i,Z_j,Z_i,Z_j) - h_r(z)) \\ &\quad + n^{-4}\sum_{i\neq j} 2K_i^3K_j(g_r(Z_i,Z_j,Z_i,Z_i) - h_r(z)) \\ &\quad + n^{-4}\sum_{i} K_i^4(g_r(Z_i,Z_i,Z_i,Z_i) - h_r(z)). \end{aligned}$$

We first obtain the upper bound of $\mathrm{pr}(|n^{-4}\sum_{i\neq j\neq k\neq l} K_iK_jK_kK_l(g_r(Z_i,Z_j,Z_k,Z_l) - h_r(z))| \geq \varepsilon)$. Under Condition (C**??**), we can prove that

$$|g_r(Z_1,Z_2,Z_3,Z_4) - h_r(z)| = |g_r(Z_1,Z_2,Z_3,Z_4) - g_r(z,z,z,z)| \leq L\sum_{i=1}^{4}|Z_i - z|.$$

Thus

$$|E(K_1K_2K_3K_4g_r(Z_1,Z_2,Z_3,Z_4)-h_r(z))| \leq L\sum_{i=1}^{4} E(K_1K_2K_3K_4|Z_i-z|).$$

With Condition (C??), we have

$$\begin{aligned}
E(K_1K_2K_3K_4|Z_1-z|) &= \int_{-\infty}^{\infty} K_1|z_1-z|f(z_1)dz_1 \int_{-\infty}^{\infty} K_2K_3K_4 f(z_2)f(z_3)f(z_4)dz_2dz_3dz_4 \\
&\leq C_0' \int_{-\infty}^{\infty} K_1|z_1-z|dz_1 \int_{-\infty}^{\infty} C_0^3 f(z_2)f(z_3)f(z_4)dz_2dz_3dz_3 \\
&= C_0'C_0^3 \int_{-\infty}^{\infty} K(z_{(1)})h^2|z_{(1)}|dz_{(1)} \\
&= 2C_0'C_0^3h^{2q_z} \int_{0}^{\infty} K(z_{(1)})z_{(1)}dz_{(1)} \\
&= Ch^{2q_z}.
\end{aligned}$$

Let $\phi(z) = E(K_1K_2K_3K_4(g_r(Z_1,Z_2,Z_3,Z_4)-h_r(z)))$, we have $|\phi(z)| \leq Ch^{2q_z}$. Note that when n is sufficiently large, $Ch^{2q_z} < \varepsilon/2$. Therefore,

$$\begin{aligned}
&\mathrm{pr}(|n^{-4}\sum_{i\neq j\neq k\neq l} K_iK_jK_kK_l(g_r(Z_i,Z_j,Z_k,Z_l)-h_r(z))| \geq \varepsilon) \\
\leq &\mathrm{pr}(|1/A_n^4\sum_{i\neq j\neq k\neq l} K_iK_jK_kK_l(g_r(Z_i,Z_j,Z_k,Z_l)-h_r(z))| \geq \varepsilon) \\
\leq &\mathrm{pr}(|1/A_n^4\sum_{i\neq j\neq k\neq l} K_iK_jK_kK_l(g_r(Z_i,Z_j,Z_k,Z_l)-h_r(z))-\phi(z)+\phi(z))| \geq \varepsilon) \\
\leq &\mathrm{pr}(|1/A_n^4\sum_{i\neq j\neq k\neq l} K_iK_jK_kK_l(g_r(Z_i,Z_j,Z_k,Z_l)-h_r(z))-\phi(z))| \geq \varepsilon/2).
\end{aligned}$$

Denote

$$A_2 = \{\forall i, j, k, l, 1 \le i, j, k, l \le n, |(g_r(Z_i, Z_j, Z_k, Z_l) - h_r(z))K_iK_jK_kK_l| \le M\},$$

$$B_2 = \{\exists i, j, k, l, 1 \le i, j, k, l \le n, |(g_r(Z_i, Z_j, Z_k, Z_l) - h_r(z))K_iK_jK_kK_l| > M\},$$

$$S_2 = \frac{1}{A_n^4} \sum_{i \ne j \ne k \ne l} K_iK_jK_kK_l(g_r(Z_i, Z_j, Z_k, Z_l) - h_r(z)) - \phi(z).$$

Following similar arguments to S_1, we have

$$\mathrm{pr}(|S_2| \ge \varepsilon) \ge \mathrm{pr}(|S_2| \ge \varepsilon | A_2) + \mathrm{pr}(B_2).$$

With the Markov's inequality, we can obtain that

$$\mathrm{pr}(|S_2| \ge \varepsilon \mid A_2) \le 2 \exp\left(-n^{1-2\gamma}\varepsilon^2/40\right).$$

For $\mathrm{pr}(B_2)$, we have

$$\begin{aligned}
\mathrm{pr}(B_2) &\le A_n^4 \max_{i,j,k,l} \mathrm{pr}(|K_iK_jK_kK_l(g_r(Z_i, Z_j, Z_k, Z_l) - h_r(z))| \ge M) \\
&\le n^4 \exp(-sM/4) E(\exp(s|K_iK_jK_kK_l(g_r(Z_i, Z_j, Z_k, Z_l) - h_r(z))|/4)).
\end{aligned}$$

Together with Condition (C**??**) that and Lemma **??**, we can see that

$$
\begin{aligned}
& E(\exp(s|K_iK_jK_kK_l(g_r(Z_i,Z_j,Z_k,Z_l)-h_r(z))|/4)) \\
\leq\ & E[\exp(sK_iK_jK_kK_lL\sum_{i=1}^{4}|Z_i-z|)/4] \\
=\ & E[\prod_{i=1}^{4}\exp(sK_iK_jK_kK_lL|Z_i-z|)/4] \\
\leq\ & \prod_{i=1}^{4}[E(\exp(sK_iK_jK_kK_lL|Z_i-z|))]^{1/4} \\
\leq\ & \prod_{i=1}^{4}[E(\exp(sLC^3K_1|Z_1-z|))]^{1/4} \\
\leq\ & \prod_{i=1}^{4}[2h\int_0^{\infty}\exp(sLC^3hz_{(1)}K(z_{(1)}))dz_{(1)}]^{1/4}.
\end{aligned}
$$

According to (C**??**) that, $\int\|u\|^2K(u)du<\infty$, thus for large enough positive number A when $u>A, K(u)=o(u^{-2})$. So $\int_0^{\infty}\exp(sLC^3hz_{(1)}K(z_{(1)}))dz_{(1)}<\infty$. This implies

$$E(\exp(s|K_iK_jK_kK_l(g_r(Z_i,Z_j,Z_k,Z_l)-h_r(z))|/4))<\infty.$$

Let $M=n^{\gamma}$, we can obtain $\mathrm{pr}(B_2)\leq n^4C\exp(-sn^{\gamma}/4)$. Therefore,

$$
\begin{aligned}
&\mathrm{pr}(|n^{-4}\sum_{i\neq j\neq k\neq l}K_iK_jK_kK_l(g_r(Z_i,Z_j,Z_k,Z_l)-h_r(z))|\geq\varepsilon) \\
&\qquad\leq 2\exp(-n^{1-2\gamma}\varepsilon^2/40)+n^4C\exp(-sn^{\gamma}/4)+o(1).
\end{aligned}
$$

Following similar arguments as S_2, we can show that

$$\operatorname{pr}(|n^{-4}\sum_{i\neq j\neq k}3K_i^2K_jK_k(g(Z_i,Z_j,Z_k,Z_i)-h_r(z))|\geq\varepsilon)$$
$$\leq 2\exp(-n^{3-2\gamma}\varepsilon^2/40)+n^4C\exp(-sn^\gamma/4)+o(1),$$
$$\operatorname{pr}(|n^{-4}\sum_{i\neq j}2K_i^2K_j^2(g_r(Z_i,Z_j,Z_i,Z_j)-h_r(z))|\geq\varepsilon)$$
$$\leq 2\exp(-n^{5-2\gamma}\varepsilon^2/40)+n^4C\exp(-sn^\gamma/4)+o(1),$$

$$\operatorname{pr}(|n^{-4}\sum_{i\neq j}2K_i^3K_j(g_r(Z_i,Z_j,Z_i,Z_i)-h_r(z))|\geq\varepsilon)$$
$$\leq 2\exp(-n^{5-2\gamma}\varepsilon^2/40)+n^4C\exp(-sn^\gamma/4)+o(1),$$
$$\operatorname{pr}(|n^{-4}\sum_{i}K_i^4(g_r(Z_i,Z_i,Z_i,Z_i)-h_r(z))|\geq\varepsilon)$$
$$\leq 2\exp(-n^{7-2\gamma}\varepsilon^2/40)+n^4C\exp(-sn^\gamma/4)+o(1).$$

Therefore, we have

$$\operatorname{pr}(|n^{-4}\sum_{i,j,k,l}K_iK_jK_kK_l(g_r(Z_i,Z_j,Z_k,Z_l)-h_r(z))|\geq\varepsilon)$$
$$\leq 2\exp(-n^{1-2\gamma}\varepsilon^2/40)+n^4C\exp(-sn^\gamma/4)+o(1).$$

This inequality together with (**??**) and (**??**) implies that

$$\begin{aligned}&\operatorname{pr}(|n^{-4}\sum_{i,j,k,l}d_{ijkl,r}K_iK_jK_kK_l-\frac{K^4}{n^4}E(d_{ijkl,r}|z)|\geq\varepsilon)\\&\leq n^4\exp(-c_2n^\gamma)+\exp(-c_1n^{1-2\gamma}\varepsilon^2)+o(1),\end{aligned}\tag{S4.8}$$

for some positive constants c_1 and c_2.

Thus the convergence rate of the numerator of $\hat{\rho}_r$ is achieved. Following similar arguments, we can obtain the convergence rate of the denominator of $\hat{\rho}_r$. We can verify that the convergence rate of $\hat{\rho}_r$ has the same form of (**??**). We omit the details here.

Let $\varepsilon = cn^{-\kappa}$, where κ satisfies $0 < \kappa + \gamma < 1/2$. Then we have

$$\begin{aligned}
\mathrm{pr}(\max_{1\le r\le p} |\hat{\rho}_r(z) - \rho_r(z)| > cn^{-\kappa}) &\le p \max_{1\le r\le p} \mathrm{pr}(|\hat{\rho}_r(z) - \rho_r(z)| > cn^{-\kappa}) \\
&\le n^4 \exp(-c_2 n^{\gamma}) + \exp(-c_1 n^{1-2\gamma}\varepsilon^2) + o(1).
\end{aligned}$$

According to Hoeffdings inequality and $\mathrm{pr}(\mathcal{A} \subseteq \hat{\mathcal{M}}_{d_n}) \subseteq \{|\hat{\rho}_r - \rho_r| > cn^{-\kappa},\ \text{for some } r \in \mathcal{A}\}$, we can show that

$$\begin{aligned}
\mathrm{pr}(\mathcal{A} \subseteq \hat{\mathcal{M}}_{d_n}) \ge& \mathrm{pr}(\max_{r\in\mathcal{A}} |n^{-1} \sum_{i=1}^{n} \hat{\rho}_r(Z_i) - E(\rho_r(Z))| \le cn^{-\kappa}) \\
\ge& 1 - |\mathcal{A}| \cdot \max_{r\in\mathcal{A}} \mathrm{pr}(|n^{-1} \sum_{i=1}^{n} \hat{\rho}_r(Z_i) - E(\rho_r(Z))| \ge cn^{-\kappa}) \\
\ge& 1 - |\mathcal{A}| \cdot \max_{r\in\mathcal{A}} [\mathrm{pr}(|n^{-1} \sum_{i=1}^{n} \hat{\rho}_r(Z_i) - \frac{1}{n} \sum_{i=1}^{n} \rho_r(Z_i)| \ge cn^{-\kappa}/2) \\
&+ \mathrm{pr}(|n^{-1} \sum_{i=1}^{n} \rho_r(Z_i) - E(\rho_r(Z))| \ge cn^{-\kappa}/2)] \\
=& 1 - |\mathcal{A}| \max_{r\in\mathcal{A}} \mathrm{pr}(|\sum_{i=1}^{n} \hat{\rho}_r(Z_i) - \sum_{i=1}^{n} \rho_r(Z_i)| \ge cn^{1-\kappa}/2) - |\mathcal{A}| \exp(-c_3 n^{1-2\kappa}) + o(1) \\
\ge& 1 - n|\mathcal{A}| \cdot (\exp(-c_1 n^{1-2(\gamma+\kappa)}) + n^4 \exp(-c_2 n^{\gamma})) - |\mathcal{A}| \exp(-c_3 n^{1-2\kappa}) + o(1),
\end{aligned}$$

where $|\mathcal{A}|$ is the cardinality of $\mathcal{A}$, and c_1, c_2, c_3 are some positive constants.