Testing the Linear Mean and Constant Variance Conditions in Sufficient Dimension Reduction

Zhejiang University of Finance and Economics[1],

Temple University[2] *and Renmin University of China*[3]

Supplementary Material

This supplement contains the proofs of Propositions 1 and 2 and Theorems 1 and 2.

S1 Proof of Proposition 1

For part 1, note that $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ because $E(\mathbf{x}) = \mathbf{0}$ and $\boldsymbol{\varepsilon} = \mathbf{x} - \mathbf{P_B x}$. All we need to show is that the LCM condition holds if and only if $E(\boldsymbol{\varepsilon} \mid \mathbf{B}^{\top}\mathbf{x}) = \mathbf{0}$. For the "only if" part, suppose the LCM condition holds. The LCM condition guarantees that $E(\mathbf{x} \mid \mathbf{B}^{\top}\mathbf{x}) = \mathbf{P_B x}$. Also note that $E(\mathbf{P_B x} \mid \mathbf{B}^{\top}\mathbf{x}) = \mathbf{P_B x}$ because $\mathbf{P_B x}$ is a function of $\mathbf{B}^{\top}\mathbf{x}$. Thus $E(\boldsymbol{\varepsilon} \mid \mathbf{B}^{\top}\mathbf{x}) = E(\mathbf{x} \mid \mathbf{B}^{\top}\mathbf{x}) - \mathbf{P_B x} = \mathbf{0}$. For the "if" part, suppose $E(\boldsymbol{\varepsilon} \mid \mathbf{B}^{\top}\mathbf{x}) = \mathbf{0}$. Then $\mathbf{0} = E\{(\mathbf{x} - \mathbf{P_B x}) \mid \mathbf{B}^{\top}\mathbf{x}\} = E(\mathbf{x} \mid \mathbf{B}^{\top}\mathbf{x}) - \mathbf{P_B x}$. It follows that $E(\mathbf{x} \mid \mathbf{B}^{\top}\mathbf{x}) = \mathbf{P_B x}$, which is a linear function of $\mathbf{B}^{\top}\mathbf{x}$.

From the definition of $\boldsymbol{\zeta} = \{\boldsymbol{\varepsilon}^{\top}, (\boldsymbol{\varepsilon} \otimes \boldsymbol{\varepsilon})^{\top}\}^{\top}$ and the result of part 1, the statement in part 2 is equivalent to the following: under the LCM condition, the CCV condition holds if and only if $E(\boldsymbol{\varepsilon} \otimes \boldsymbol{\varepsilon} \mid \mathbf{B}^{\mathrm{T}}\mathbf{x}) = E(\boldsymbol{\varepsilon} \otimes \boldsymbol{\varepsilon})$. By the property of the kronecker product, $E(\boldsymbol{\varepsilon} \otimes \boldsymbol{\varepsilon} \mid \mathbf{B}^{\mathrm{T}}\mathbf{x}) = E(\boldsymbol{\varepsilon} \otimes \boldsymbol{\varepsilon})$ is equivalent to $E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}} \mid \mathbf{B}^{\mathrm{T}}\mathbf{x}) = E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}})$. It remains to show that under the LCM condition, the CCV condition holds if and only if $E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}} \mid \mathbf{B}^{\mathrm{T}}\mathbf{x}) = E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}})$.

For the "only if" part, suppose $\mathrm{var}(\mathbf{x} \mid \mathbf{B}^{\mathrm{T}}\mathbf{x})$ is constant. Then

$$\begin{aligned}\mathrm{var}(\mathbf{x} \mid \mathbf{B}^{\mathrm{T}}\mathbf{x}) &= E\{\mathrm{var}(\mathbf{x} \mid \mathbf{B}^{\mathrm{T}}\mathbf{x})\} = \mathrm{var}(\mathbf{x}) - \mathrm{var}\{E(\mathbf{x} \mid \mathbf{B}^{\mathrm{T}}\mathbf{x})\} \\ &= \mathbf{I}_p - \mathbf{P}_{\mathbf{B}} = \mathbf{Q}_{\mathbf{B}}. \end{aligned} \tag{S1.1}$$

Here the first equality is true because $\mathrm{var}(\mathbf{x} \mid \mathbf{B}^{\mathrm{T}}\mathbf{x})$ is constant. The second equality follows from the EV-VE formula. The third equality is true because $\mathrm{var}(\mathbf{x}) = \mathbf{I}_p$, $\mathrm{var}\{E(\mathbf{x} \mid \mathbf{B}^{\mathrm{T}}\mathbf{x})\} = \mathrm{var}(\mathbf{P}_{\mathbf{B}}\mathbf{x})$, and $\mathbf{P}_{\mathbf{B}}$ is idempotent. The last equality is from the definition of $\mathbf{Q}_{\mathbf{B}}$. Under the LCM condition, we have $\boldsymbol{\varepsilon} = \mathbf{x} - \mathbf{P}_{\mathbf{B}}\mathbf{x} = \mathbf{x} - E(\mathbf{x} \mid \mathbf{B}^{\mathrm{T}}\mathbf{x})$. The definition of conditional variance leads to

$$\begin{aligned}E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}} \mid \mathbf{B}^{\mathrm{T}}\mathbf{x}) &= E[\{\mathbf{x} - E(\mathbf{x} \mid \mathbf{B}^{\mathrm{T}}\mathbf{x})\}\{\mathbf{x} - E(\mathbf{x} \mid \mathbf{B}^{\mathrm{T}}\mathbf{x})\}^{\mathrm{T}} \mid \mathbf{B}^{\mathrm{T}}\mathbf{x}] \\ &= \mathrm{var}(\mathbf{x} \mid \mathbf{B}^{\mathrm{T}}\mathbf{x}). \end{aligned} \tag{S1.2}$$

On the other hand, note that $\boldsymbol{\varepsilon} = \mathbf{x} - \mathbf{P}_{\mathbf{B}}\mathbf{x} = (\mathbf{I}_p - \mathbf{P}_{\mathbf{B}})\mathbf{x} = \mathbf{Q}_{\mathbf{B}}\mathbf{x}$. It follows that

$$E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}}) = \mathbf{Q}_{\mathbf{B}}\mathrm{var}(\mathbf{x})\mathbf{Q}_{\mathbf{B}} = \mathbf{Q}_{\mathbf{B}}\mathbf{Q}_{\mathbf{B}} = \mathbf{Q}_{\mathbf{B}}. \tag{S1.3}$$

(S1.1), (S1.2), and (S1.3) together imply that $E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\top} \mid \mathbf{B}^{\top}\mathbf{x}) = E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\top})$. For the "if" part, suppose $E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\top} \mid \mathbf{B}^{\top}\mathbf{x}) = E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\top})$. Under the LCM condition, both (S1.2) and (S1.3) are true. Together they imply $\mathrm{var}(\mathbf{x} \mid \mathbf{B}^{\top}\mathbf{x}) = \mathbf{Q}_{\mathbf{B}}$ is a constant matrix. □

S2 Proof of Proposition 2

The proof is similar to Theorem 1 of Shao and Zhang (2014), and is thus omitted. □

S3 Proof of Theorem 1

For part 1, define $\boldsymbol{\xi}_n(\mathbf{s}) = n^{-1}\sum_{j=1}^{n} \widehat{\boldsymbol{\varepsilon}}_j \exp(i\mathbf{s}^{\top}\widehat{\mathbf{B}}^{\top}\mathbf{x}_j)$ and $\boldsymbol{\phi}_n(\mathbf{s}) = n^{1/2}\boldsymbol{\xi}_n(\mathbf{s})$. From the proof of Theorem 4 in Shao and Zhang (2014), we have $n\widehat{\omega}_n = \|\boldsymbol{\phi}_n(\mathbf{s})\|^2$. It remains to show that $\|\boldsymbol{\phi}_n(\mathbf{s})\|^2 \xrightarrow{d} \|\boldsymbol{\phi}(\mathbf{s})\|^2$ as $n \to \infty$. First we have

$$\exp(i\mathbf{s}^{\top}\widehat{\mathbf{B}}^{\top}\mathbf{x}_j) = \cos(\mathbf{s}^{\top}\widehat{\mathbf{B}}^{\top}\mathbf{x}_j) + i\sin(\mathbf{s}^{\top}\widehat{\mathbf{B}}^{\top}\mathbf{x}_j). \tag{S3.4}$$

Let $\theta_1 = \mathbf{s}^{\top}\widehat{\mathbf{B}}^{\top}\mathbf{x}_j$, $\theta_2 = \mathbf{s}^{\top}\mathbf{B}^{\top}\mathbf{x}_j$, and $\theta_3 = \mathbf{s}^{\top}(\widehat{\mathbf{B}} - \mathbf{B})^{\top}\mathbf{x}_j$. Because $\widehat{\mathbf{B}} - \mathbf{B} = O_p(n^{-1/2})$, we have $\theta_3 = O_p(n^{-1/2})$. Note that $\cos\theta = \sum_{j=0}^{\infty}\{j(2j)!\}^{-1}\theta^{2j}$

and $\sin\theta = \sum_{j=0}^{\infty}\{j(2j+1)!\}^{-1}\theta^{2j+1}$ for any $\theta \in \mathbb{R}$. It follows that

$$\cos\theta_3 = 1 + o_p(n^{-1/2}) \text{ and } \sin\theta_3 = \theta_3 + o_p(n^{-1/2}). \tag{S3.5}$$

Note that $\theta_1 = \theta_2 + \theta_3$. By the angle sum identities, we have $\cos\theta_1 = \cos\theta_2\cos\theta_3 - \sin\theta_2\sin\theta_3$ and $\sin\theta_1 = \sin\theta_2\cos\theta_3 + \cos\theta_2\sin\theta_3$. Together with (S3.4) and (S3.5), we obtain

$$\begin{aligned}\exp(i\theta_1) &= \cos\theta_2 - \theta_3\sin\theta_2 + i(\sin\theta_2 + \theta_3\cos\theta_2) + o_p(n^{-1/2}) \\ &= \exp(i\theta_2) + \theta_3(-\sin\theta_2 + i\cos\theta_2) + o_p(n^{-1/2}).\end{aligned}$$

Plug in $\theta_1 = \mathbf{s}^{\mathrm{T}}\widehat{\mathbf{B}}^{\mathrm{T}}\mathbf{x}_j$, $\theta_2 = \mathbf{s}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{x}_j$, $\theta_3 = \mathbf{s}^{\mathrm{T}}(\widehat{\mathbf{B}} - \mathbf{B})^{\mathrm{T}}\mathbf{x}_j$, and we get

$$\begin{aligned}\exp(i\mathbf{s}^{\mathrm{T}}\widehat{\mathbf{B}}^{\mathrm{T}}\mathbf{x}_j) &= \exp(i\mathbf{s}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{x}_j) + \{i\cos(\mathbf{s}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{x}_j) - \sin(\mathbf{s}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{x}_j)\} \\ &\quad \mathbf{s}^{\mathrm{T}}(\widehat{\mathbf{B}} - \mathbf{B})^{\mathrm{T}}\mathbf{x}_j + o_p(n^{-1/2}),\end{aligned}$$

where the second term above is of order $O_p(n^{-1/2})$. On the other hand,

$$\widehat{\boldsymbol{\varepsilon}}_j = (\mathbf{I}_p - \mathbf{P}_{\widehat{\mathbf{B}}})\mathbf{x}_j = (\mathbf{I}_p - \mathbf{P}_{\mathbf{B}})\mathbf{x}_j + (\mathbf{P}_{\mathbf{B}} - \mathbf{P}_{\widehat{\mathbf{B}}})\mathbf{x}_j = \boldsymbol{\varepsilon}_j + (\mathbf{P}_{\mathbf{B}} - \mathbf{P}_{\widehat{\mathbf{B}}})\mathbf{x}_j,$$

where $(\mathbf{P}_{\mathbf{B}} - \mathbf{P}_{\widehat{\mathbf{B}}})\mathbf{x}_j = O_p(n^{-1/2})$. Together with the definition of $\boldsymbol{\phi}_n(\mathbf{s})$, we have

$$\begin{aligned}\boldsymbol{\phi}_n(\mathbf{s}) &= n^{-1/2}\sum_{j=1}^{n}\widehat{\boldsymbol{\varepsilon}}_j\exp(i\mathbf{s}^{\mathrm{T}}\widehat{\mathbf{B}}^{\mathrm{T}}\mathbf{x}_j) \\ &= \boldsymbol{\phi}_n^{(1)}(\mathbf{s}) + \boldsymbol{\phi}_n^{(2)}(\mathbf{s}) + \boldsymbol{\phi}_n^{(3)}(\mathbf{s}) + o_p(1),\end{aligned} \tag{S3.6}$$

where $\boldsymbol{\phi}_n^{(1)}(\mathbf{s}) = n^{-1/2}\sum_{j=1}^{n}\boldsymbol{\varepsilon}_j\exp(i\mathbf{s}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{x}_j)$, $\boldsymbol{\phi}_n^{(2)}(\mathbf{s}) = n^{-1/2}(\mathbf{P}_{\mathbf{B}} - \mathbf{P}_{\widehat{\mathbf{B}}})\sum_{j=1}^{n}\mathbf{x}_j\exp(i\mathbf{s}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{x}_j)$, and $\boldsymbol{\phi}_n^{(3)}(\mathbf{s}) = n^{-1/2}\sum_{j=1}^{n}\boldsymbol{\varepsilon}_j\{i\cos(\mathbf{s}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{x}_j) - \sin(\mathbf{s}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{x}_j)\}$

$\mathbf{s}^{\mathrm{T}}(\widehat{\mathbf{B}}-\mathbf{B})^{\mathrm{T}}\mathbf{x}_j$. Because $\mathbf{P}_{\widehat{\mathbf{B}}}-\mathbf{P}_{\mathbf{B}} = n^{-1}\sum_{j=1}^{n}\boldsymbol{\ell}_2(\mathbf{x}_j,Y_j)+o_p(n^{-1/2})$, $\boldsymbol{\phi}_n^{(2)}(\mathbf{s})$ becomes

$$\boldsymbol{\phi}_n^{(2)}(\mathbf{s}) = n^{-1/2}E\left\{\mathbf{x}\exp(i\mathbf{s}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{x})\right\}\sum_{j=1}^{n}\boldsymbol{\ell}_2(\mathbf{x}_j,Y_j)+o_p(1). \tag{S3.7}$$

Because $\widehat{\mathbf{B}}-\mathbf{B} = n^{-1}\sum_{i=1}^{n}\boldsymbol{\ell}_1(\mathbf{x}_i,Y_i)+o_p(n^{-1/2})$, $\boldsymbol{\phi}_n^{(3)}(\mathbf{s})$ becomes

$$\begin{aligned}\boldsymbol{\phi}_n^{(3)}(\mathbf{s}) &= n^{-1/2}E\left[\boldsymbol{\varepsilon}\{i\cos(\mathbf{s}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{x})-\sin(\mathbf{s}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{x})\}\mathbf{x}^{\mathrm{T}}\right]\left\{\sum_{j=1}^{n}\boldsymbol{\ell}_1(\mathbf{x}_j,Y_j)\right\}\mathbf{s}\\ &+ o_p(1).\end{aligned} \tag{S3.8}$$

Recall that $\mathbf{g}(\mathbf{s}) = E\left\{\mathbf{x}\exp(i\mathbf{s}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{x})\right\}$ and $\mathbf{h}(\mathbf{s}) = E[\boldsymbol{\varepsilon}\{i\cos(\mathbf{s}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{x}) - \sin(\mathbf{s}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{x})\}\mathbf{x}^{\mathrm{T}}]$. (S3.6), (S3.7), and (S3.8) together lead to

$$\boldsymbol{\phi}_n(\mathbf{s}) = n^{-1/2}\sum_{j=1}^{n}\boldsymbol{\ell}_3(\mathbf{x}_j,Y_j,\mathbf{s})+o_p(1), \tag{S3.9}$$

where $\boldsymbol{\ell}_3(\mathbf{x}_j,Y_j,\mathbf{s}) = \boldsymbol{\varepsilon}_j\exp(i\mathbf{s}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{x}_j)-\boldsymbol{\ell}_2(\mathbf{x}_j,Y_j)\mathbf{g}(\mathbf{s})+\mathbf{h}(\mathbf{s})\boldsymbol{\ell}_1(\mathbf{x}_j,Y_j)\mathbf{s}$. Under H_0, we have $E(\boldsymbol{\varepsilon}\mid\mathbf{B}^{\mathrm{T}}\mathbf{x}) = \mathbf{0}$. Thus $E\{\boldsymbol{\varepsilon}\exp(i\mathbf{s}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{x})\} = E\{E(\boldsymbol{\varepsilon}\mid\mathbf{B}^{\mathrm{T}}\mathbf{x})\exp(i\mathbf{s}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{x})\} = \mathbf{0}$. Also $E\{\boldsymbol{\ell}_k(\mathbf{x},Y)\} = \mathbf{0}$ for $k=1,2$. Take expectation on both sides of (S3.9),

$$E\{\boldsymbol{\phi}_n(\mathbf{s})\} = \mathbf{0} \text{ as } n\to\infty. \tag{S3.10}$$

For $\mathrm{cov}_{\boldsymbol{\phi}_n}(\mathbf{s},\mathbf{s}_0) = \mathrm{cov}\left\{\boldsymbol{\phi}_n(\mathbf{s}),\overline{\boldsymbol{\phi}_n(\mathbf{s}_0)}\right\}$, $\mathrm{cov}_{\boldsymbol{\phi}_n}(\mathbf{s},\mathbf{s}_0) = E\left\{\boldsymbol{\phi}_n(\mathbf{s})\overline{\boldsymbol{\phi}_n(\mathbf{s}_0)}^{\mathrm{T}}\right\}$ as $n\to\infty$. Because $(\mathbf{x}_j,Y_j)\perp\!\!\!\perp(\mathbf{x}_k,Y_k)$ for $j\neq k$ and $E\{\boldsymbol{\ell}_3(\mathbf{x},Y,\mathbf{s})\} = \mathbf{0}$,

$$E\left\{\sum_{j=1}^{n}\sum_{k=1}^{n}\boldsymbol{\ell}_3(\mathbf{x}_j,Y_j,\mathbf{s})\overline{\boldsymbol{\ell}_3(\mathbf{x}_k,Y_k,\mathbf{s}_0)}\right\} = E\left\{\sum_{j=1}^{n}\boldsymbol{\ell}_3(\mathbf{x}_j,Y_j,\mathbf{s})\overline{\boldsymbol{\ell}_3(\mathbf{x}_j,Y_j,\mathbf{s}_0)}\right\}.$$

Thus as $n \to \infty$, we have

$$\begin{aligned} \mathrm{cov}_{\boldsymbol{\phi}_n}(\mathbf{s}, \mathbf{s}_0) &= n^{-1} E\left\{\sum_{j=1}^{n} \boldsymbol{\ell}_3(\mathbf{x}_j, Y_j, \mathbf{s})\overline{\boldsymbol{\ell}_3(\mathbf{x}_j, Y_j, \mathbf{s}_0)}\right\} \\ &= E\left\{\boldsymbol{\ell}_3(\mathbf{x}, Y, \mathbf{s})\overline{\boldsymbol{\ell}_3(\mathbf{x}, Y, \mathbf{s}_0)}\right\} \qquad \text{(S3.11)} \end{aligned}$$

Note that $\overline{\exp(i\mathbf{s}_0^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{x})} = \exp(-i\mathbf{s}_0^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{x})$, $\overline{\mathbf{g}(\mathbf{s}_0)} = \mathbf{g}(-\mathbf{s}_0)$, and $\overline{\mathbf{h}(\mathbf{s}_0)} = \mathbf{h}(-\mathbf{s}_0)$. We have $\overline{\boldsymbol{\ell}_3(\mathbf{x}, Y, \mathbf{s}_0)} = \boldsymbol{\varepsilon}\exp(-i\mathbf{s}_0^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{x}) - \boldsymbol{\ell}_2(\mathbf{x}, Y)\mathbf{g}(-\mathbf{s}_0) + \mathbf{h}(-\mathbf{s}_0)$ $\boldsymbol{\ell}_1(\mathbf{x}, Y)\mathbf{s}_0$. Plug them into (S3.11), together with the definition of $\mathrm{cov}_{\boldsymbol{\phi}}(\mathbf{s}, \mathbf{s}_0)$, we have

$$\mathrm{cov}_{\boldsymbol{\phi}_n}(\mathbf{s}, \mathbf{s}_0) = \mathrm{cov}_{\boldsymbol{\phi}}(\mathbf{s}, \mathbf{s}_0) \text{ as } n \to \infty. \qquad \text{(S3.12)}$$

From (S3.10) and (S3.12), we know the two complex-valued Gaussian processes $\boldsymbol{\phi}_n(\mathbf{s})$ and $\boldsymbol{\phi}(\mathbf{s})$ have the same mean function and the same covariance function as $n \to \infty$. From the proof of Theorem 5 and Corollary 2 in Székely et al. (2007), we know $\|\boldsymbol{\phi}_n(\mathbf{s})\|^2 \xrightarrow{d} \|\boldsymbol{\phi}(\mathbf{s})\|^2$ as n goes to infinity.

Now we turn to part 2. First note that $\widehat{\omega}_n \xrightarrow{p} m(\boldsymbol{\varepsilon} \mid \mathbf{B}^{\mathrm{T}}\mathbf{x})$ as n goes to infinity. Under $H_1 : E(\boldsymbol{\varepsilon} \mid \mathbf{B}^{\mathrm{T}}\mathbf{x}) \neq E(\boldsymbol{\varepsilon})$ almost surely, $m(\boldsymbol{\varepsilon} \mid \mathbf{B}^{\mathrm{T}}\mathbf{x}) > 0$ according to Proposition 1. Thus $n\widehat{\omega}_n \xrightarrow{p} \infty$ under H_1. $\square$

S4 Proof of Theorem 2

From Theorem 1, we have $n\widehat{\omega}_n \xrightarrow{d} \|\boldsymbol{\phi}(\mathbf{s})\|^2$. Recall that $\widehat{\mathbf{B}}^{(t)}$ is an estimator of $\mathbf{B}$ based on $\{(\mathbf{x}_j^{(t)}, Y_j) : j = 1, \ldots, n\}$. Let $\boldsymbol{\phi}_n^{(t)}(\mathbf{s}) = n^{1/2}\boldsymbol{\xi}_n^{(t)}(\mathbf{s})$, where $\boldsymbol{\xi}_n^{(t)}(\mathbf{s}) = n^{-1}\sum_{j=1}^n \widehat{\boldsymbol{\varepsilon}}_j^{(t)} \exp\left\{i\mathbf{s}^{\mathrm{T}}(\widehat{\mathbf{B}}^{(t)})^{\mathrm{T}}\mathbf{x}_j^{(t)}\right\}$. Then $n\widehat{\omega}_n^{(t)} = \|\boldsymbol{\phi}_n^{(t)}(\mathbf{s})\|^2$. Following the proof of Theorem 1, where $\mathbf{B}$ and $\widehat{\mathbf{B}}$ are replaced by $\widehat{\mathbf{B}}$ and $\widehat{\mathbf{B}}^{(t)}$ respectively, we have $\|\boldsymbol{\phi}_n^{(t)}(\mathbf{s})\|^2 \xrightarrow{d} \|\boldsymbol{\phi}^*(\mathbf{s})\|^2$ as long as $E(\mathbf{x}^* \mid \widehat{\mathbf{B}}^{\mathrm{T}}\mathbf{x}^*) = \mathbf{0}$. If we have the additional condition that $\mathrm{cov}_{\boldsymbol{\phi}^*}(\mathbf{s}, \mathbf{s}_0) = \mathrm{cov}_{\boldsymbol{\phi}}(\mathbf{s}, \mathbf{s}_0)$, then $\|\boldsymbol{\phi}^*(\mathbf{s})\|^2 = \|\boldsymbol{\phi}(\mathbf{s})\|^2$ and we get the desired result. It remains to show that (i) $E(\mathbf{x}^* \mid \widehat{\mathbf{B}}^{\mathrm{T}}\mathbf{x}^*) = \mathbf{0}$ as $n \to \infty$, and (ii) $\mathrm{cov}_{\boldsymbol{\phi}^*}(\mathbf{s}, \mathbf{s}_0) = \mathrm{cov}_{\boldsymbol{\phi}}(\mathbf{s}, \mathbf{s}_0)$.

Because $\mathbf{x}^* = \mathbf{P}_{\widehat{\mathbf{B}}}\mathbf{x} + W^*\mathbf{Q}_{\widehat{\mathbf{B}}}\mathbf{x}$, we have $E(\mathbf{x}^*) = \mathbf{0} = E(\mathbf{x})$. Note that $\mathbf{x}^*(\mathbf{x}^*)^{\mathrm{T}} = \mathbf{P}_{\widehat{\mathbf{B}}}\mathbf{x}\mathbf{x}^{\mathrm{T}}\mathbf{P}_{\widehat{\mathbf{B}}} + (W^*)^2\mathbf{Q}_{\widehat{\mathbf{B}}}\mathbf{x}\mathbf{x}^{\mathrm{T}}\mathbf{Q}_{\widehat{\mathbf{B}}} + W^*\mathbf{P}_{\widehat{\mathbf{B}}}\mathbf{x}\mathbf{x}^{\mathrm{T}}\mathbf{Q}_{\widehat{\mathbf{B}}} + W^*\mathbf{Q}_{\widehat{\mathbf{B}}}\mathbf{x}\mathbf{x}^{\mathrm{T}}\mathbf{P}_{\widehat{\mathbf{B}}}$. Because $\mathrm{var}(\mathbf{x}) = \mathbf{I}_p$, $\mathbf{Q}_{\widehat{\mathbf{B}}}\mathbf{P}_{\widehat{\mathbf{B}}} = \mathbf{0}$, $E\{(W^*)^2\} = 1$ and $W^* \perp\!\!\!\perp \mathbf{x}$, we have $\mathrm{var}(\mathbf{x}^*) = E\{\mathbf{x}^*(\mathbf{x}^*)^{\mathrm{T}}\} = \mathbf{P}_{\widehat{\mathbf{B}}} + \mathbf{Q}_{\widehat{\mathbf{B}}} = \mathbf{I}_p = \mathrm{var}(\mathbf{x})$. Thus (ii) is true from condition (C3).

Define $\psi(\mathbf{B}) \overset{\text{def}}{=} E(\mathbf{Q}_{\mathbf{B}}\boldsymbol{\varepsilon}^* \mid \mathbf{P}_{\mathbf{B}}\mathbf{x}^*)$ and $\psi(\widehat{\mathbf{B}}) \overset{\text{def}}{=} E(\mathbf{Q}_{\widehat{\mathbf{B}}}\boldsymbol{\varepsilon}^* \mid \mathbf{P}_{\widehat{\mathbf{B}}}\mathbf{x}^*)$, where $\widehat{\mathbf{B}}$ can be any consistent estimator of $\mathbf{B}$. We thus have $\psi(\mathbf{B}) - \psi(\widehat{\mathbf{B}}) = \psi'(\boldsymbol{\kappa})(\mathbf{B} - \widehat{\mathbf{B}})$ where $\boldsymbol{\kappa}$ is between $\mathbf{B}$ and $\widehat{\mathbf{B}}$. According to condition (C4), $\psi'(\boldsymbol{\kappa})$ is bounded and for any $C > 0$ we have $\Pr(\|\widehat{\mathbf{B}} - \mathbf{B}\|_{\max} \le Cn^{-1/2}) \to 1$, where $\|\mathbf{A}\|_{\max} \overset{\text{def}}{=} \max\{| a_{ij} |\}$ for any matrix $\mathbf{A}$. Besides, we write $E(\mathbf{Q}_{\widehat{\mathbf{B}}}\boldsymbol{\varepsilon}^* \mid$

$\mathbf{P}_{\widehat{\mathbf{B}}}\mathbf{x}^*)$

$$\begin{aligned} &= E(\mathbf{Q}_{\widehat{\mathbf{B}}}\boldsymbol{\varepsilon}^* \mid \mathbf{P}_{\widehat{\mathbf{B}}}\mathbf{x}^*, \|\widehat{\mathbf{B}}-\mathbf{B}\|_{\max} \leq Cn^{-1/2})\Pr(\|\widehat{\mathbf{B}}-\mathbf{B}\|_{\max} \leq Cn^{-1/2}) \\ &+ E(\mathbf{Q}_{\widehat{\mathbf{B}}}\boldsymbol{\varepsilon}^* \mid \mathbf{P}_{\widehat{\mathbf{B}}}\mathbf{x}^*, \|\widehat{\mathbf{B}}-\mathbf{B}\|_{\max} > Cn^{-1/2})\Pr(\|\widehat{\mathbf{B}}-\mathbf{B}\|_{\max} > Cn^{-1/2}), \end{aligned}$$

together with the fact that

$$\sup_{\|\widehat{\mathbf{B}}-\mathbf{B}\|_{\max} \leq Cn^{-1/2}} E(\mathbf{Q}_{\widehat{\mathbf{B}}}\boldsymbol{\varepsilon}^* \mid \mathbf{P}_{\widehat{\mathbf{B}}}\mathbf{x}^*) \to E(\mathbf{Q}_{\mathbf{B}}\boldsymbol{\varepsilon}^* \mid \mathbf{P}_{\mathbf{B}}\mathbf{x}^*) \to \mathbf{0},$$

and

$$\Pr(\|\widehat{\mathbf{B}}-\mathbf{B}\|_{\max} > Cn^{-1/2}) \to 0,$$

thus we have $E(\mathbf{Q}_{\widehat{\mathbf{B}}}\boldsymbol{\varepsilon}^* \mid \mathbf{P}_{\widehat{\mathbf{B}}}\mathbf{x}^*, \|\widehat{\mathbf{B}}-\mathbf{B}\|_{\max} \leq Cn^{-1/2}) \to \mathbf{0}$ and $\Pr(\|\widehat{\mathbf{B}}-\mathbf{B}\|_{\max} > Cn^{-1/2}) \to 0$. Combing the above results, we have $E(\mathbf{Q}_{\widehat{\mathbf{B}}}\boldsymbol{\varepsilon}^* \mid \mathbf{P}_{\widehat{\mathbf{B}}}\mathbf{x}^*) \to \mathbf{0}$ and (i) is true. This completes the proof of Theorem 2. □

References

Shao, X. and Zhang, J. (2014). Martingale difference correlation and its use in high-dimensional variable screening. *Journal of the American Statistical Association*, **109**:1302–1318.

Székely, G. J., Rizzo, M. L., and Bakirov, N. K. (2007). Measuring and testing dependence by correlation of distances. *The Annals of Statistics*, **35**:2769–2794.

Tingyou Zhou, Assistant Professor, School of Data Sciences, Zhejiang University of Finance and Economics, 18 Xueyuan Road, Xiasha High Education District, Hangzhou 310018, China.

E-mail: zhoutingyou@zufe.edu.cn

Yuexiao Dong, Associate Professor, Department of Statistical Science, Fox School of Business, Temple University, 1801 Liacouras Walk, Philadelphia, PA 19122, USA.

E-mail: ydong@temple.edu

Liping Zhu, Center for Applied Statistics, Institute of Statistics and Big Data, Renmin University of China, Beijing 100872, China.

E-mail: zhu.liping@ruc.edu.cn