Regime-Switching Factor Models for High-Dimensional Time Series

Rutgers University

Supplementary Material

The detailed proofs of Lemma 1-4, Theorem 1-5 and Corollary 1 are provided in the Supplementary Material.

Here we use Cs to denote the generic uniformly positive constants. Define

$$\boldsymbol{\Sigma}_{x,k,j}(l) = \frac{1}{n-l}\sum_{t=1}^{n-l}\mathrm{Cov}(\mathbf{x}_t,\, \mathbf{x}_{t+l}I(z_{t+l}=j) \mid z_t = k) = \pi_{k,j}^{(l)}\boldsymbol{\Sigma}_x(l),$$

$$\boldsymbol{\Sigma}_{f,k,j}(l) = \frac{1}{n-l}\sum_{t=1}^{n-l}\mathrm{Cov}\{\mathbf{f}_t,\, \mathbf{f}_{t+l}I(z_{t+l}=j) \mid z_t = k)\},$$

$$\hat{\boldsymbol{\Sigma}}_{x,k,j}(l) = \frac{\sum_{t=1}^{n-l}(\mathbf{x}_t-\bar{\mathbf{x}}_k)(\mathbf{x}_{t+l}-\bar{\mathbf{x}}_j)'I(z_t=k, z_{t+l}=j)}{\sum_{t=1}^{n-l}I(z_t=k)},$$

$$\hat{\boldsymbol{\Sigma}}_{\varepsilon,k,j} = \frac{\sum_{t=1}^{n-l}(\boldsymbol{\varepsilon}_t^{(k)}-\bar{\boldsymbol{\varepsilon}}_k)(\boldsymbol{\varepsilon}_{t+l}^{(j)}-\bar{\boldsymbol{\varepsilon}}_j)'I(z_t=k, z_{t+l}=j)}{\sum_{t=1}^{n-l}I(z_t=k)},$$

$$\hat{\boldsymbol{\Sigma}}_{f,k,j}(l) = \frac{\sum_{t=1}^{n-l}(\mathbf{f}_t-\bar{\mathbf{f}}_k)(\mathbf{f}_{t+l}-\bar{\mathbf{f}}_j)'I(z_t=k, z_{t+l}=j)}{\sum_{t=1}^{n-l}I(z_t=k)},$$

$$\hat{\boldsymbol{\Sigma}}_{f,\varepsilon,k,j}(l) = \frac{\sum_{t=1}^{n-l}(\mathbf{f}_t-\bar{\mathbf{f}}_k)(\boldsymbol{\varepsilon}_{t+l}^{(j)}-\bar{\boldsymbol{\varepsilon}}_j)'I(z_t=k, z_{t+l}=j)}{\sum_{t=1}^{n-l}I(z_t=k)},$$

$$\hat{\boldsymbol{\Sigma}}_{\varepsilon,f,k,j}(l) = \frac{\sum_{t=1}^{n-l}(\boldsymbol{\varepsilon}_t^{(k)}-\bar{\boldsymbol{\varepsilon}}_k)(\mathbf{f}_{t+l}-\bar{\mathbf{f}}_j)'I(z_t=k, z_{t+l}=j)}{\sum_{t=1}^{n-l}I(z_t=k)},$$

where $\bar{\mathbf{x}}_k = \sum_{t=1}^{n}\mathbf{x}_t I(z_t=k)/\sum_{t=1}^{n}I(z_t=k)$, $\bar{\mathbf{f}}_k = \sum_{t=1}^{n}\mathbf{f}_t I(z_t=k)/\sum_{t=1}^{n}I(z_t=k)$, and $\bar{\boldsymbol{\varepsilon}}_k = \sum_{t=1}^{n}\boldsymbol{\varepsilon}_t^{(k)}I(z_t=k)/\sum_{t=1}^{n}I(z_t=k)$, for $k=1,\ldots,m$.

We introduce some lemmas first.

Lemma 1. *Under Conditions 1-2 and Condition 5, if $\pi_{k,j}^{(l)} > 0$, we have*

$$\|\hat{\boldsymbol{\Sigma}}_{x,k,j}(l) - \boldsymbol{\Sigma}_{x,k,j}(l)\|_2 = O_p(n^{-1/2}), \quad \text{for} \quad k,j = 1,\ldots,m. \tag{S.2}$$

Proof. Since $\mathbf{z}$ is irreducible, positive current and aperiodic under Condition 5, by Theorem 3.5 in Bradley (2005) and Theorem 17.0.1 in Meyne and Tweedie (2009), it follows

$$\frac{n}{\sum_{t=1}^{n-l}I(z_t=k)} - \frac{1}{\pi_k} = O_p(n^{-1/2}), \tag{S.3}$$

$$\frac{\sum_{t=1}^{n-l} I(z_t = k, z_{t+l} = j)}{\sum_{t=1}^{n-l} I(z_t = k)} - \pi_{k,j}^{(l)} = O_p(n^{-1/2}). \tag{S.4}$$

$$\begin{aligned}
\hat{\mathbf{\Sigma}}_{x,k,j}(l) - \mathbf{\Sigma}_{x,k,j}(l) &= \frac{\sum_{t=1}^{n-l}(\mathbf{x}_t - \bar{\mathbf{x}}_k)(\mathbf{x}_{t+l} - \bar{\mathbf{x}}_j)' I(z_t = k, z_{t+l} = j)}{\sum_{t=1}^{n-l} I(z_t = k)} - \frac{\sum_{t=1}^{n-l} \pi_{k,j}^{(l)} \mathrm{E}(\mathbf{x}_t \mathbf{x}_{t+l}')}{n-l} \\
&= \frac{\sum_{t=1}^{n-l} [(\mathbf{x}_t - \bar{\mathbf{x}}_k)(\mathbf{x}_{t+l} - \bar{\mathbf{x}}_j)' - \mathrm{E}(\mathbf{x}_t \mathbf{x}_{t+l}') I(z_t = k, z_{t+l} = j)]}{\sum_{t=1}^{n-l} I(z_t = k)} \\
&\quad + \sum_{t=1}^{n-l} \left[\left(\frac{I(z_t = k, z_{t+l} = j)}{\sum_{t=1}^{n-l} I(z_t = k)} - \frac{\pi_{k,j}^{(l)}}{n-l} \right) \mathrm{E}(\mathbf{x}_t \mathbf{x}_{t+l}') \right] \\
&= I_1 + I_2,
\end{aligned}$$

where

$$\begin{aligned}
I_1 &= \frac{\sum_{t=1}^{n-l}(\mathbf{x}_t \mathbf{x}_{t+l}' - \mathrm{E}\mathbf{x}_t \mathbf{x}_{t+l}') I(z_t = k, z_{t+l} = j)}{\sum_{t=1}^{n-l} I(z_t = k)} - \frac{\sum_{t=1}^{n-l} \mathbf{x}_t \bar{\mathbf{x}}_j' I(z_t = k, z_{t+l} = j)}{\sum_{t=1}^{n-l} I(z_t = k)} \\
&\quad - \frac{\sum_{t=1}^{n-l} \bar{\mathbf{x}}_k \mathbf{x}_{t+l}' I(z_t = k, z_{t+l} = j)}{\sum_{t=1}^{n-l} I(z_t = k)} + \frac{\sum_{t=1}^{n-l} \bar{\mathbf{x}}_k \bar{\mathbf{x}}_j' I(z_t = k, z_{t+l} = j)}{\sum_{t=1}^{n-l} I(z_t = k)} \\
&= L_1 + L_2 + L_3 + L_4.
\end{aligned}$$

For L_1, since $\mathbf{x}$ and $\mathbf{z}$ are independent, for $i, q = 1, \ldots, d$, by (S.3) and Davydov inequality, under Condition 1, it follows that,

$$\begin{aligned}
&\mathrm{E}\left\{ \left[\frac{\sum_{t=1}^{n-l} (x_{i,t} x_{q,t+l} - \mathrm{E}(x_{i,t} x_{q,t+l})) I(z_t = k, z_{t+l} = j)}{\sum_{t=1}^{n-l} I(z_t = k)} \right]^2 \right\} \\
&\leq \frac{C}{n^2} \mathrm{E}\left\{ \left[\sum_{t=1}^{n-l} (x_{i,t} x_{q,t+l} - \mathrm{E}(x_{i,t} x_{q,t+l})) I(z_t = k, z_{t+l} = j) \right]^2 \right\} \\
&\leq \frac{C}{n^2} \sum_{|t_1 - t_2| > l} \left| \mathrm{E}\left\{ [x_{i,t_1} x_{q,t_1+l} - \mathrm{E}(x_{i,t_1} x_{q,t_1+l})] \cdot [x_{i,t_2} x_{q,t_2+l} - \mathrm{E}(x_{i,t_2} x_{q,t_2+l})] \right\} \right| \\
&\quad + \frac{C}{n^2} \sum_{|t_1 - t_2| \leq l} \left| \mathrm{E}\left\{ [x_{i,t_1}, x_{q,t_1+l} - \mathrm{E}(x_{i,t_1} x_{q,t_1+l})] \cdot [x_{i,t_2} x_{q,t_2+l} - \mathrm{E}(x_{i,t_2} x_{q,t_2+l})] \right\} \right| \\
&\leq \frac{C}{n^2} \sum_{t_1 \neq t_2} \alpha(|t_1 - t_2|)^{1-2/\gamma} + \frac{C}{n} = O(1/n).
\end{aligned}$$

Then $\mathrm{E}\left(\|L_1\|_F^2\right) = O(1/n)$. Since $\|L_1\|_2 \leq \|L_1\|_F \leq \sqrt{d}\|L_1\|_2$, it follows that $\|L_1\|_2 = O_p(n^{-1/2})$.

For L_2, under Conditions 1 and 2, by (S.3) and Davydov inequality, we have

$$\mathrm{E}\|\bar{\mathbf{x}}_j\|_2^2 = \sum_{q=1}^{d} \mathrm{E}\left(\frac{\sum_{t=1}^{n} x_{q,t} I(z_t = j)}{\sum_{t=1}^{n} I(z_t = j)} \right)^2 \leq \frac{C}{n^2} \sum_{q=1}^{d} \left(\sum_{t=1}^{n} \mathrm{E}(x_{q,t}^2) + \sum_{t_1 \neq t_2}^{n} \left| \mathrm{Cov}(x_{q,t_1}, x_{q,t_2}) \right| \right) = O(1/n), \tag{S.5}$$

and

$$\mathrm{E}\left\|\frac{\sum_{t=1}^{n-l}\mathbf{x}_t I(z_t=k,z_{t+l}=j)}{\sum_{t=1}^{n-l}I(z_t=k)}\right\|_2^2 = \sum_{q=1}^{d}\mathrm{E}\left(\frac{\sum_{t=1}^{n-l}x_{q,t}I(z_t=k,z_{t+l}=j)}{\sum_{t=1}^{n-l}I(z_t=k)}\right)^2$$
$$\leq \frac{C}{n^2}\sum_{q=1}^{d}\left(\sum_{t=1}^{n-l}\mathrm{E}(x_{q,t}^2)+\sum_{t_1\neq t_2}^{n-l}\left|\mathrm{Cov}(x_{q,t_1},x_{q,t_2})\right|\right)=O(1/n). \quad \text{(S.6)}$$

Hence,

$$\|L_2\|_2 \leq \|\bar{\mathbf{x}}_j\|_2\cdot\left\|\frac{\sum_{t=1}^{n-l}\mathbf{x}_t I(z_t=k,z_{t+l}=j)}{\sum_{t=1}^{n-l}I(z_t=k)}\right\|_2=O_p(1/n).$$

Similarly $\|L_3\|_2=O_p(1/n)$ and $\|L_4\|_2=O_p(1/n^2)$.

For $i,q=1,\dots,d$, since the second moment of $\mathbf{x}_t$ is bounded under Condition 2, together with (S.4),

$$\mathrm{E}\left\{\sum_{t=1}^{n-l}\left[\left(\frac{I(z_t=k,z_{t+l}=j)}{\sum_{t=1}^{n-l}I(z_t=k)}-\frac{\pi_{k,j}^{(l)}}{n-l}\right)\mathrm{E}(x_{i,t}x_{q,t+l})\right]\right\}^2\leq\frac{C}{n}.$$

(S.2) follows by combining the above results. □

Lemma 2. *Under Conditions 1-5, if* $\pi_{k,j}^{(l)}>0$*, we have, for* $k,j=1,\dots,m$.

$$\|\hat{\mathbf{\Sigma}}_{f,k,j}(l)-\mathbf{\Sigma}_{f,k,j}(l)\|_2 = O_p(p^{1-\delta_k/2-\delta_j/2}n^{-1/2}), \quad \text{(S.7)}$$
$$\|\hat{\mathbf{\Sigma}}_{f,\varepsilon,k,j}(l)\|_2 = O_p(p^{1-\delta_k/2}n^{-1/2}), \quad \text{(S.8)}$$
$$\|\hat{\mathbf{\Sigma}}_{\varepsilon,f,k,j}(l)\|_2 = O_p(p^{1-\delta_j/2}n^{-1/2}). \quad \text{(S.9)}$$

Proof.

$$\begin{aligned}
&\hat{\mathbf{\Sigma}}_{f,k,j}(l)-\mathbf{\Sigma}_{f,k,j}(l)\\
&= \frac{\sum_{t=1}^{n-l}\mathbf{A}_k(\mathbf{x}_t-\bar{\mathbf{x}}_k)(\mathbf{x}_{t+l}-\bar{\mathbf{x}}_j)'\mathbf{A}_j'I(z_t=k,z_{t+l}=j)}{\sum_{t=1}^{n-l}I(z_t=k)}-\pi_{k,j}^{(l)}\mathbf{A}_k\mathbf{\Sigma}_{x,k,j}(l)\mathbf{A}_j'\\
&= \frac{\mathbf{A}_k\sum_{t=1}^{n-l}\left(\hat{\mathbf{\Sigma}}_{x,k,j}(l)-\mathbf{\Sigma}_{x,k,j}(l)\right)\mathbf{A}_j'}{\sum_{t=1}^{n-l}I(z_t=k)}\\
&\quad +\mathbf{A}_k\mathbf{\Sigma}_{x,k,j}(l)\mathbf{A}_j'\left(\frac{\sum_{t=1}^{n}I(z_t=k,z_{t+l}=j)}{\sum_{t=1}^{n-l}I(z_t=k)}-\pi_{k,j}^{(l)}\right).
\end{aligned}$$

Hence, under Conditions 2 and 4, by Lemma 1 and (S.4), for $k=1,\dots,m$,

$$\begin{aligned}
\|\hat{\mathbf{\Sigma}}_{f,k,j}(l)-\mathbf{\Sigma}_{f,k,j}(l)\|_2 &\leq \|\mathbf{A}_k\|_2\cdot\|\hat{\mathbf{\Sigma}}_{x,k,j}(l)-\mathbf{\Sigma}_{x,j,k}(l)\|_2\cdot\|\mathbf{A}_j\|_2\\
&\quad +\|\mathbf{A}_k\|_2\cdot\|\mathbf{\Sigma}_{x,k,j}(l)\|_2\cdot\|\mathbf{A}_j\|_2\cdot O(n^{-1/2})\\
&= O_p(p^{1-\delta_k/2-\delta_j/2}n^{-1/2}).
\end{aligned}$$

For (S.8), we expand $\hat{\mathbf{\Sigma}}_{f,\varepsilon,k,j}(l)$,

$$\begin{aligned}\hat{\mathbf{\Sigma}}_{f,\varepsilon,k,j}(l) &= \frac{\mathbf{A}_k\sum_{j=1}^m\sum_{t=1}^{n-l}(\mathbf{x}_t-\bar{\mathbf{x}}_k)(\boldsymbol{\varepsilon}_{t+l}^{(j)}-\bar{\boldsymbol{\varepsilon}}_j)'I(z_t=k,z_{t+l}=j)}{\sum_{t=1}^{n-l}I(z_t=k)}\\ &= \frac{\mathbf{A}_k\sum_{j=1}^m\sum_{t=1}^{n-l}\mathbf{x}_t{\boldsymbol{\varepsilon}_{t+l}^{(j)}}'I(z_t=k,z_{t+l}=j)}{\sum_{t=1}^{n-l}I(z_t=k)} - \frac{\mathbf{A}_k\sum_{j=1}^m\sum_{t=1}^{n-l}\mathbf{x}_t\bar{\boldsymbol{\varepsilon}}_j'I(z_t=k,z_{t+l}=j)}{\sum_{t=1}^{n-l}I(z_t=k)}\\ &\quad -\frac{\mathbf{A}_k\sum_{j=1}^m\sum_{t=1}^{n-l}\bar{\mathbf{x}}_k{\boldsymbol{\varepsilon}_{t+l}^{(j)}}'I(z_t=k,z_{t+l}=j)}{\sum_{t=1}^{n-l}I(z_t=k)} + \frac{\mathbf{A}_k\sum_{j=1}^m\sum_{t=1}^{n-l}\bar{\mathbf{x}}_k\bar{\boldsymbol{\varepsilon}}_j'I(z_t=k,z_{t+l}=j)}{\sum_{t=1}^{n-l}I(z_t=k)}\\ &= L_1+L_2+L_3+L_4.\end{aligned}$$

For L_1, since $\mathbf{x}$, $\boldsymbol{\varepsilon}^{(j)}$ and $\mathbf{z}$ are independent, $j=1,\ldots,m$, under Conditions 1-3, for $i=1,\ldots,d$, $q=1,\ldots,p$,

$$\begin{aligned}&\mathrm{E}\left[\left(\frac{\sum_{t=1}^{n-l}x_{i,t}\varepsilon_{q,t+l}^{(j)}I(z_t=k,z_{t+l}=j)}{\sum_{t=1}^{n-l}I(z_t=k)}\right)^2\right] \le \frac{C}{n^2}\mathrm{E}\left[\left(\sum_{t=1}^{n-l}x_{i,t}\varepsilon_{q,t+l}^{(j)}I(z_t=k,z_{t+l}=j)\right)^2\right]\\ &\le \frac{C}{n^2}\sum_{t=1}^{n-l}\mathrm{E}\left[x_{i,t}^2\left(\varepsilon_{q,t+l}^{(j)}\right)^2\right]+\sum_{t_1\ne t_2}^{n-l}\left|\mathrm{Cov}\left(x_{i,t_1}\varepsilon_{q,t_1+l}^{(j)},x_{i,t_2}\varepsilon_{q,t_2+l}^{(j)}\right)\right| = O(1/n).\end{aligned}$$

So $\mathrm{E}\left\|\sum_{j=1}^m\sum_{t=1}^{n-l}\mathbf{x}_t{\boldsymbol{\varepsilon}_{t+l}^{(j)}}'I(z_t=k,z_{t+l}=j)/\sum_{t=1}^{n-l}I(z_t=k)\right\|_F^2=O(pn^{-1})$. Under Condition 4, we have

$$\begin{aligned}\|L_1\|_2 &\le \|\mathbf{A}_k\|_2\cdot\left\|\frac{\sum_{j=1}^m\sum_{t=1}^{n}\mathbf{x}_t{\boldsymbol{\varepsilon}_{t+l}^{(j)}}'I(z_t=k,z_{t+l}=j)}{\sum_{t=1}^nI(z_t=k)}\right\|_F\\ &= O(p^{1/2-\delta_k/2})O_p(p^{1/2}n^{-1/2})=O_p(p^{1-\delta_k/2}n^{-1/2}).\end{aligned}$$

For L_2, with (S.3) and independence of $\mathbf{z}$ and $\boldsymbol{\varepsilon}_t^{(j)}$, under Condition 3,

$$\begin{aligned}\mathrm{E}\|\bar{\boldsymbol{\varepsilon}}_j\|_2^2 &= \sum_{q=1}^p\mathrm{E}\left(\frac{\sum_{t=1}^n\boldsymbol{\varepsilon}_{q,t}^{(j)}I(z_t=j)}{\sum_{t=1}^nI(z_t=j)}\right)^2 \le \sum_{q=1}^p\frac{C}{n^2}\mathrm{E}\left[\left(\sum_{t=1}^n\varepsilon_{q,t}^{(j)}I(z_t=j)\right)^2\right]\\ &\le \sum_{q=1}^p\frac{C}{n^2}\left[\sum_{t=1}^n\mathrm{E}\left(\varepsilon_{q,t}^{(j)}\right)^2+\sum_{t_1\ne t_2}\left|\mathrm{E}(\varepsilon_{q,t_1}\varepsilon_{q,t_2})\right|\right]\le\frac{Cp}{n}=O(pn^{-1}),\end{aligned}$$

and with (S.5) and (S.6), under Condition 4, we have $\|L_2\|_2=O_p(p^{1-\delta_k/2}n^{-1})$, $\|L_3\|_2=O_p(p^{1-\delta_k/2}n^{-1})$, and $\|L_4\|_2=O_p(p^{1-\delta_k/2}n^{-2})$. Hence (S.8) follows. Similar to the proof of (S.8), we can prove (S.9). □

Lemma 3. *Under Conditions 4-7,*

$$\lambda_{\min}(\mathbf{M}_k)=O(p^{2-\delta_k-\delta_{\min}}),\ \textit{for } k=1,\ldots,m. \tag{S.10}$$

where $\mathbf{M}_k$ *is defined in (9) and* $\lambda_{\min}(\mathbf{M})_k$ *is the minimum eigenvalue of* $\mathbf{M}_k$*.*

Proof. Let $\sigma_{\max}(\mathbf{H})$ and $\sigma_{\min}(\mathbf{H})$ denote the maximum and minimum singular value of $\mathbf{H}$. Under Condition 6, using the inequality about the singular values in Merikoski and Kumar (2004) we can prove that $\sigma_{\min}(\mathbf{A}_k\boldsymbol{\Sigma}_x(l_k)\sum_{j\in\mathcal{C}}\pi_{k,j}^{(l_k)}\mathbf{A}_j') = O(p^{1-\delta_k/2-\delta_{\min}/2})$. Using the fact that $\sigma_{\max}(\mathbf{A}_k\boldsymbol{\Sigma}_x(l_k)\sum_{j\notin\mathcal{C}}\pi_{k,j}^{(l_k)}\mathbf{A}_j') = o(p^{1-\delta_k/2-\delta_{\min}/2})$ under Conditions 4 and 6, we have

$$\begin{aligned}\sigma_{\min}\left(\mathbf{A}_k\boldsymbol{\Sigma}_x(l_k)\sum_{j=1}^m\pi_{k,j}^{(l_k)}\mathbf{A}_j'\right) &\geq \sigma_{\min}\left(\mathbf{A}_k\boldsymbol{\Sigma}_x(l_k)\sum_{j\in\mathcal{C}}\pi_{k,j}^{(l_k)}\mathbf{A}_j'\right) - \sigma_{\max}\left(\mathbf{A}_k\boldsymbol{\Sigma}_x(l_k)\sum_{j\notin\mathcal{C}}\pi_{k,j}^{(l_k)}\mathbf{A}_j'\right)\\ &= O(p^{1-\delta_k/2-\delta_{\min}/2}). \qquad \text{(S.11)}\end{aligned}$$

It follows that

$$\lambda_{\min}(\mathbf{M}_k) \geq \max_{1\leq l\leq l_0}\sigma_{\min}^2\left(\mathbf{A}_k\boldsymbol{\Sigma}_x(l)\sum_{j=1}^m\pi_{k,j}^{(l)}\mathbf{A}_j'\right) = O(p^{2-\delta_k-\delta_{\min}}).$$

□

Lemma 4. *Under Conditions 1-7,*

$$\|\hat{\mathbf{M}}_k - \mathbf{M}_k\|_2 = O_p(p^{2-\delta_k/2-\delta_{\min}/2}n^{-1/2}),\ \textit{for } k=1,\ldots,m.$$

Proof.

$$\|\hat{\mathbf{M}}_k - \mathbf{M}_k\|_2 \leq \sum_{l=1}^{l_0}\left(\|\hat{\boldsymbol{\Sigma}}_{y,k}(l) - \boldsymbol{\Sigma}_{y,k}(l)\|_2^2 + 2\|\boldsymbol{\Sigma}_{y,k}(l)\|_2\cdot\|\hat{\boldsymbol{\Sigma}}_{y,k}(l) - \boldsymbol{\Sigma}_{y,k}(l)\|_2\right). \qquad \text{(S.12)}$$

Conditions 5, 6 and 7 indicate that

$$\|\boldsymbol{\Sigma}_{y,k}(l_k)\|_2 = O_p(p^{1-\delta_k/2-\delta_{\min}/2}). \qquad \text{(S.13)}$$

When $\pi_{k,j}^{(l)} > 0$, for any l,

$$\begin{aligned}\hat{\boldsymbol{\Sigma}}_{y,k}(l) &= \frac{\sum_{j=1}^m\sum_{t=1}^{n-l}(\mathbf{y}_t - \hat{\boldsymbol{\mu}}_k)(\mathbf{y}_{t+l} - \hat{\boldsymbol{\mu}}_j)'I(z_t=k, z_{t+l}=j)}{\sum_{t=1}^{n-l}I(z_t=k)}\\ &= \frac{\sum_{j=1}^m\sum_{t=1}^{n-l}(\mathbf{y}_t - \boldsymbol{\mu}_k - \mathbf{A}_k\bar{\mathbf{x}}_k - \bar{\boldsymbol{\varepsilon}}_k)(\mathbf{y}_{t+l} - \boldsymbol{\mu}_j - \mathbf{A}_j\bar{\mathbf{x}}_j - \bar{\boldsymbol{\varepsilon}}_j)'I(z_t=k, z_{t+l}=j)}{\sum_{t=1}^{n-l}I(z_t=k)}\\ &= \frac{\sum_{j=1}^m\sum_{t=1}^{n-l}(\mathbf{A}_k(\mathbf{x}_t - \bar{\mathbf{x}}_k) + \boldsymbol{\varepsilon}_t^{(k)} - \bar{\boldsymbol{\varepsilon}}_k)(\mathbf{A}_j(\mathbf{x}_{t+l} - \bar{\mathbf{x}}_j) + \boldsymbol{\varepsilon}_{t+l}^{(j)} - \bar{\boldsymbol{\varepsilon}}_j)'I(z_t=k, z_{t+l}=j)}{\sum_{t=1}^{n-l}I(z_t=k)}\\ &= \sum_{j=1}^m\left(\hat{\boldsymbol{\Sigma}}_{f,k,j}(l) + \hat{\boldsymbol{\Sigma}}_{\varepsilon,k,j}(l) + \hat{\boldsymbol{\Sigma}}_{f,\varepsilon,k,j}(l) + \hat{\boldsymbol{\Sigma}}_{\varepsilon,f,k,j}(l)\right).\end{aligned}$$

By Lemma 2 we have

$$\begin{aligned}&\|\hat{\boldsymbol{\Sigma}}_{y,k}(l_k) - \boldsymbol{\Sigma}_{y,k}(l_k)\|_2\\ &= \sum_{j=1}^m\left(\|\hat{\boldsymbol{\Sigma}}_{f,k,j}(l_k) - \boldsymbol{\Sigma}_{f,k,j}(l_k)\|_2 + \|\hat{\boldsymbol{\Sigma}}_{f,\varepsilon,k,j}(l_k)\|_2 + \|\hat{\boldsymbol{\Sigma}}_{\varepsilon,f,k,j}(l_k)\|_2 + \|\hat{\boldsymbol{\Sigma}}_{\varepsilon,k,j}(l_k)\|_2\right)\\ &= O_p(p^{1-\delta_k/2-\delta_{\min}/2}n^{-1/2} + p^{1-\delta_k/2}n^{-1/2} + p^{1-\delta_{\min}/2}n^{-1/2} + \sum_{k=1}^m\|\boldsymbol{\Sigma}_{\varepsilon,k,j}(l_k)\|_2). \qquad \text{(S.14)}\end{aligned}$$

Since $\boldsymbol{\varepsilon}_t^{(k)}$ are independent noises, we have $\|\hat{\boldsymbol{\Sigma}}_{\varepsilon,k,j}(l_k)\|_2 \le \|\hat{\boldsymbol{\Sigma}}_{\varepsilon,k,j}(l_k)\|_F = O_p(pn^{-1/2})$, which implies from (S.14) that

$$\|\hat{\boldsymbol{\Sigma}}_{y,k}(l_k) - \boldsymbol{\Sigma}_{y,k}(l_k)\|_2 = O_p(pn^{-1/2}). \tag{S.15}$$

Together with (S.12), (S.13) and (S.15), the lemma follows. □

Proof of Theorem 1:

Proof. By Lemmas 1-4, and Lemma 3 in Lam et al. (2011), we can easily reach the conclusion of Theorem 1. □

Proof of Theorem 2:

Proof. From (13), when $z_t = k$,

$$\begin{aligned}\hat{\mathbf{f}}_t - \mathbf{f}_t &= \hat{\mathbf{Q}}_k\hat{\mathbf{R}}_t - \mathbf{Q}_k\mathbf{R}_t = \hat{\mathbf{Q}}_k\hat{\mathbf{Q}}_k'(\mathbf{y}_t - \hat{\boldsymbol{\mu}}_k) - \mathbf{Q}_k\mathbf{R}_t = \hat{\mathbf{Q}}_k\hat{\mathbf{Q}}_k'(\mathbf{Q}_k\mathbf{R}_t + \boldsymbol{\varepsilon}_t^{(k)} + \boldsymbol{\mu}_k - \hat{\boldsymbol{\mu}}_k) - \mathbf{Q}_k\mathbf{R}_t \\ &= (\hat{\mathbf{Q}}_k\hat{\mathbf{Q}}_k' - \mathbf{Q}_k\mathbf{Q}_k')\mathbf{Q}_k\mathbf{R}_t + \hat{\mathbf{Q}}_k(\hat{\mathbf{Q}}_k - \mathbf{Q}_k)'(\boldsymbol{\varepsilon}_t^{(k)} + \boldsymbol{\mu}_k - \hat{\boldsymbol{\mu}}_k) + \hat{\mathbf{Q}}_k\mathbf{Q}_k'(\boldsymbol{\varepsilon}_t^{(k)} + \boldsymbol{\mu}_k - \hat{\boldsymbol{\mu}}_k) \\ &= I_1 + I_2 + I_3.\end{aligned}$$

Note that when $z_t = k$, $\|\mathbf{R}_t\|_2 = \|\mathbf{A}_k\|_2 = O(p^{1/2-\delta_k/2})$ defined in (3), so $\|I_1\|_2 \le 2\|\hat{\mathbf{Q}}_k - \mathbf{Q}_k\|_2\|\mathbf{R}_t\|_2 = O_p(p^{1/2-\delta_k/2}\|\hat{\mathbf{Q}}_k - \mathbf{Q}_k\|_2) = O_p(p^{1/2+\delta_{\min}/2}n^{-1/2})$. I_2 is dominated by I_3 in probability.

$$\mathrm{E}(\|\hat{\mathbf{Q}}_k\mathbf{Q}_k'\boldsymbol{\varepsilon}_t^{(k)}\|_2^2) = \sum_{i=1}^{d}\mathrm{E}[(\mathbf{q}_i'\boldsymbol{\varepsilon}_t^{(k)})^2] \le d\lambda_{max}(\boldsymbol{\Sigma}_k) < \infty. \tag{S.16}$$

$$\hat{\boldsymbol{\mu}}_k - \boldsymbol{\mu}_k = \frac{\sum_{t=1}^n \mathbf{y}_t I(z_t = k)}{\sum_{t=1}^n I(z_t = k)} - \boldsymbol{\mu}_k = \frac{\sum_{t=1}^n (\mathbf{A}_k\mathbf{x}_t + \boldsymbol{\varepsilon}_t^{(k)}) I(z_t = k)}{\sum_{t=1}^n I(z_t = k)}.$$

By (S.16), we can easily have

$$\|\frac{\sum_{t=1}^n (\hat{\mathbf{Q}}_k\mathbf{Q}_k'\boldsymbol{\varepsilon}_t^{(k)} I(z_t = k)}{\sum_{t=1}^n I(z_t = k)}\|_2 = O_p(n^{-1/2}).$$

Under Condition 4, $\|\hat{\mathbf{Q}}_k\mathbf{Q}_k'(\hat{\boldsymbol{\mu}}_k - \boldsymbol{\mu}_k)\|_2 = O_p(p^{1/2-\delta_k/2}n^{-1/2}) + O_p(n^{-1/2})$. Hence, with (S.16), $\|I_3\|_2 = O_p(p^{1/2-\delta_k/2}n^{-1/2}) + O_p(1)$.

We have $p^{-1/2}\|\hat{\mathbf{f}}_t - \mathbf{f}_t\|_2 = O_p(p^{\delta_{\min}/2}n^{-1/2} + p^{-1/2})$. □

Proof of Theorem 3:

Proof. We assume that $\mathbf{Q}_k$ is uniquely defined as in Remark 10 under Condition 7. Then

$$\mathrm{tr}\left[\mathbf{Q}_k'(\mathbf{I}_p - \hat{\mathbf{Q}}_k\hat{\mathbf{Q}}_k')\mathbf{Q}_k\right] = \mathrm{tr}(\mathbf{I}_d - \mathbf{Q}_k'\hat{\mathbf{Q}}_k\hat{\mathbf{Q}}_k'\mathbf{Q}_k) = d\left[\mathcal{D}(\mathcal{M}(\hat{\mathbf{Q}}_k), \mathcal{M}(\mathbf{Q}_k))\right]^2. \tag{S.17}$$

On the other hand,

$$\begin{aligned}\operatorname{tr}\left[\mathbf{Q}_k'(\mathbf{I}_p-\hat{\mathbf{Q}}_k\hat{\mathbf{Q}}_k')\mathbf{Q}_k\right]-\operatorname{tr}\left[\mathbf{Q}_k'(\mathbf{I}_p-\mathbf{Q}_k\mathbf{Q}_k')\mathbf{Q}_k\right] &= \operatorname{tr}\left[\mathbf{Q}_k'(\mathbf{Q}_k\mathbf{Q}_k'-\hat{\mathbf{Q}}_k\hat{\mathbf{Q}}_k')\mathbf{Q}_k\right]\\ &\leq d\|\mathbf{Q}_k'(\mathbf{Q}_k\mathbf{Q}_k'-\hat{\mathbf{Q}}_k\hat{\mathbf{Q}}_k')\mathbf{Q}_k\|_2.\end{aligned}$$

And since the diagonal entries in $\mathbf{Q}_k'\hat{\mathbf{Q}}_k\hat{\mathbf{Q}}_k'\mathbf{Q}_k$ are between 0 and 1,

$$\begin{aligned}\operatorname{tr}\left[\mathbf{Q}_k'(\mathbf{I}_p-\hat{\mathbf{Q}}_k\hat{\mathbf{Q}}_k')\mathbf{Q}_k\right]-\operatorname{tr}\left[\mathbf{Q}_k'(\mathbf{I}_p-\mathbf{Q}_k\mathbf{Q}_k')\mathbf{Q}_k\right] &= \operatorname{tr}\left[\mathbf{Q}_k'(\mathbf{Q}_k\mathbf{Q}_k'-\hat{\mathbf{Q}}_k\hat{\mathbf{Q}}_k')\mathbf{Q}_k\right]\\ &\geq \|\mathbf{Q}_k'(\mathbf{Q}_k\mathbf{Q}_k'-\hat{\mathbf{Q}}_k\hat{\mathbf{Q}}_k')\mathbf{Q}_k\|_2.\end{aligned}$$

Note that $\operatorname{tr}\left[\mathbf{Q}_k'(\mathbf{I}_p-\mathbf{Q}_k\mathbf{Q}_k')\mathbf{Q}_k\right]=0$. Hence,

$$\left[\mathcal{D}(\mathcal{M}(\hat{\mathbf{Q}}_k),\mathcal{M}(\mathbf{Q}_k))\right]^2 \asymp \|\mathbf{Q}_k'(\mathbf{Q}_k\mathbf{Q}_k'-\hat{\mathbf{Q}}_k\hat{\mathbf{Q}}_k')\mathbf{Q}_k\|_2.$$

Since

$$\mathbf{Q}_k'(\mathbf{Q}_k\mathbf{Q}_k'-\hat{\mathbf{Q}}_k\hat{\mathbf{Q}}_k')\mathbf{Q}_k=-\mathbf{Q}_k'(\mathbf{Q}_k-\hat{\mathbf{Q}}_k)(\mathbf{Q}_k-\hat{\mathbf{Q}}_k)'\mathbf{Q}_k+(\mathbf{Q}_k-\hat{\mathbf{Q}}_k)'(\mathbf{Q}_k-\hat{\mathbf{Q}}_k),$$

which is bounded by $2\|\hat{\mathbf{Q}}_k-\mathbf{Q}\|_2^2$, with (S.17) we have

$$\mathcal{D}(\hat{\mathbf{Q}}_k,\mathbf{Q}_k)=O_p(\|\hat{\mathbf{Q}}_k-\mathbf{Q}_k\|_2).$$

By Theorem 1, we have proved Theorem 3. □

Proof of Theorem 4:

Proof. The proof is quite similar to that of Theorem 1 of Lam and Yao (2012). We denote $\lambda_{k,j}$ and $\hat{\mathbf{q}}_{k,j}$ for the j-th largest eigenvalues of $\hat{\mathbf{M}}_k$ and its corresponding orthonormal eigenvectors, respectively, for $k=1,\dots,m$. The corresponding population values are denoted by $\lambda_{k,j}$ and $\mathbf{q}_{k,j}$ for the matrix $\mathbf{M}_k$. Let $\hat{\mathbf{Q}}_k=(\hat{\mathbf{q}}_{k,1},\dots,\hat{\mathbf{q}}_{k,d})$ and $\mathbf{Q}_k=(\mathbf{q}_{k,1},\dots,\mathbf{q}_{k,d})$. We have

$$\lambda_{k,j}=\mathbf{q}_{k,j}'\mathbf{M}_k\mathbf{q}_{k,j}, \quad \text{and} \qquad \hat{\lambda}_{k,j}=\hat{\mathbf{q}}_{k,j}'\hat{\mathbf{M}}_k\hat{\mathbf{q}}_{k,j}, \qquad j=1,\dots,p.$$

We can decompose $\hat{\lambda}_{k,j}-\lambda_{k,j}$ by

$$\hat{\lambda}_{k,j}-\lambda_{k,j}=\hat{\mathbf{q}}_{k,j}'\hat{\mathbf{M}}_k\hat{\mathbf{q}}_{k,j}-\mathbf{q}_{k,j}'\mathbf{M}_k\mathbf{q}_{k,j}=I_1+I_2+I_3+I_4+I_5,$$

where

$$I_1=(\hat{\mathbf{q}}_{k,j}-\mathbf{q}_{k,j})'(\hat{\mathbf{M}}_k-\mathbf{M}_k)\hat{\mathbf{q}}_{k,j}, \quad I_2=(\hat{\mathbf{q}}_{k,j}-\mathbf{q}_{k,j})'\mathbf{M}(\hat{\mathbf{q}}_{k,j}-\mathbf{q}_{k,j}),$$

$$I_3=(\hat{\mathbf{q}}_{k,j}-\mathbf{q}_{k,j})'\mathbf{M}_k\mathbf{q}_{k,j}, \quad I_4=\mathbf{q}_{k,j}'(\hat{\mathbf{M}}_k-\mathbf{M}_k)\hat{\mathbf{q}}_{k,j}, \quad I_5=\mathbf{q}_{k,j}'\mathbf{M}_k(\hat{\mathbf{q}}_{k,j}-\mathbf{q}_{k,j}).$$

For $j=1,\dots,d$, $\|\hat{\mathbf{q}}_{k,j}-\mathbf{q}_{k,j}\|_2\leq\|\hat{\mathbf{Q}}_k-\mathbf{Q}_k\|_2=O_p(h_{n,k})$, where $h_{n,k}=p^{\delta_k/2+\delta_{\min}/2}n^{-1/2}$ by Theorem 1, and $\|\mathbf{M}_k\|_2\leq\sum_{l=1}^{l_0}\|\mathbf{\Sigma}_y(l)\|^2=O_p(p^{2-\delta_k-\delta_{\min}})$. By Lemma 2 and Lemma 4, we have $\|I_1\|_2$ and $\|I_2\|_2$ are of order $O_p(p^{2-\delta_k-\delta_{\min}}h_{n,k}^2)$ and $\|I_3\|_2$, $\|I_4\|_2$ and $\|I_5\|_2$ are of order $O_p(p^{2-\delta_k-\delta_{\min}}h_{n,k})$. So $|\hat{\lambda}_{k,j}-\lambda_{k,j}|=O_p(p^{2-\delta_k-\delta\min}h_{n,k})=O_p(p^{2-\delta_k/2-\delta_{\min}/2}n^{-1/2})$.

For $j = d+1, \dots, p$, define,

$$\widetilde{\mathbf{M}}_k = \sum_{l=1}^{l_0} \hat{\mathbf{\Sigma}}_{y,k}(l)\mathbf{\Sigma}_{y,k}(l)', \quad \hat{\mathbf{B}}_k = (\hat{\mathbf{q}}_{k,d+1}, \dots, \hat{\mathbf{q}}_{k,p}), \quad \text{and} \quad \mathbf{B}_k = (\mathbf{q}_{k,d+1}, \dots, \mathbf{q}_{k,p}).$$

It can be shown that $\|\hat{\mathbf{B}}_k - \mathbf{B}_k\|_2 = O_p(h_{n,k})$, similar to proof of Theorem 1 with Lemma 3 in Lam et al. (2011). Hence, $\|\hat{\mathbf{q}}_{k,j} - \mathbf{q}_{k,j}\|_2 \le \|\hat{\mathbf{B}}_k - \mathbf{B}_k\|_2 = O_p(h_{n,k})$.
Since $\lambda_j = 0$, for $j = d+1, \dots, p$, consider the decomposition

$$\hat{\lambda}_j = \hat{\mathbf{q}}_{k,j}'\hat{\mathbf{M}}_k\hat{\mathbf{q}}_{k,j} = K_1 + K_2 + K_3,$$

where

$$K_1 = \hat{\mathbf{q}}_{k,j}'(\hat{\mathbf{M}}_k - \widetilde{\mathbf{M}}_k - \widetilde{\mathbf{M}}_k' + \mathbf{M}_k)'\hat{\mathbf{q}}_{k,j}, \quad K_2 = 2\hat{\mathbf{q}}_{k,j}'(\widetilde{\mathbf{M}}_k - \mathbf{M}_k)(\hat{\mathbf{q}}_{k,j} - \mathbf{q}_{k,j}),$$

$$K_3 = (\hat{\mathbf{q}}_{k,j} - \mathbf{q}_{k,j})'\mathbf{M}_k(\hat{\mathbf{q}}_{kj}' - \mathbf{q}_{k,j}).$$

By Lemma 2 and Lemma 4,

$$\begin{aligned}
K_1 &= \sum_{l=1}^{l_0} \|(\hat{\mathbf{\Sigma}}_{y,k}(l) - \mathbf{\Sigma}_{y,k}(l))\hat{\mathbf{q}}_{k,j}\|_2^2 \le \sum_{l=1}^{l_0} \|\hat{\mathbf{\Sigma}}_{y,k}(l) - \mathbf{\Sigma}_{y,k}(l)\|_2^2 = O_p(p^2 n^{-1}), \\
|K_2| &= O_p(\|\widetilde{\mathbf{M}}_k - \mathbf{M}_k\|_2 \cdot \|\hat{\mathbf{q}}_{k,j} - \mathbf{q}_{k,j}\|_2) = O_p(\|\widetilde{\mathbf{M}}_k - \mathbf{M}_k\|_2 \cdot \|\hat{\mathbf{B}}_k - \mathbf{B}_k\|_2) = O_p(p^2 n^{-1}), \\
|K_3| &= O_p(\|\hat{\mathbf{B}}_k - \mathbf{B}_k\|_2^2 \cdot \|\mathbf{M}_k\|_2) = O_p(p^{2-\delta_k-\delta_{\min}} h_n^2) = O_p(p^2 n^{-1}).
\end{aligned}$$

Hence $\lambda_{k,j} = O_p(p^2 n^{-1})$. □

Proof of Corollary 1:

Proof. The proof is similar to the proof of Corollary 1 of Lam and Yao (2012).
By Lemma 3 and Lemma 4, we have

$$\lambda_{k,1} = \|\mathbf{M}_k\|_2 = O(p^{2-\delta_k-\delta_{\min}}) \quad \text{and} \quad \lambda_{k,d} = O(p^{2-\delta_k-\delta_{\min}}).$$

So we have $\lambda_{k,i} \asymp p^{2-\delta_k-\delta_{\min}}$, for $i = 1, \dots, d$. From Theorem 4(i), we have $|\hat{\lambda}_{k,i} - \lambda_{k,i}| = O_p(p^{2-\delta_k-\delta_{\min}} n^{-1/2})$, then $\hat{\lambda}_{k,i} = O_p(p^{2-\delta_k-\delta_{\min}})$ for $i = 1, \dots, d$. It implies that $\hat{\lambda}_{k,i+1}/\hat{\lambda}_{k,i} \asymp 1$ for $i = 1, \dots, d-1$. By Theorem 4(ii),

$$\hat{\lambda}_{k,d+1}/\hat{\lambda}_{k,d} = O_p(p^2 n^{-1}/p^{2-\delta_k-\delta_{\min}}) = O_p(p^{\delta_k+\delta_{\min}} n^{-1}).$$

□

Proof of Theorem 5:

Proof.

$$
\begin{aligned}
w_{t,k,j} &= \frac{1}{2}(\log|\mathbf{\Sigma}_{B,j}| - \log|\mathbf{\Sigma}_{B,k}|) + \frac{1}{2}\boldsymbol{\varepsilon}_t^{(k)\prime}(\mathbf{B}_j\mathbf{\Sigma}_{B,j}{}^{-1}\mathbf{B}_j' - \mathbf{B}_k\mathbf{\Sigma}_{B,k}{}^{-1}\mathbf{B}_k')\boldsymbol{\varepsilon}_t^{(k)} \\
&\quad + \frac{1}{2}\left(\mathbf{B}_j'\mathbf{A}_k\mathbf{x}_t + \mathbf{B}_j'(\boldsymbol{\mu}_k - \boldsymbol{\mu}_j)\right)' \mathbf{\Sigma}_{B,j}{}^{-1}\left(\mathbf{B}_j'\mathbf{A}_k\mathbf{x}_t + \mathbf{B}_j'(\boldsymbol{\mu}_k - \boldsymbol{\mu}_j)\right) \\
&\quad + (\mathbf{B}_j'(\mathbf{A}_k\mathbf{x}_t + \boldsymbol{\mu}_k - \boldsymbol{\mu}_j))'\mathbf{\Sigma}_{B,j}^{-1}\mathbf{B}_j'\boldsymbol{\varepsilon}_t^{(k)} \\
&= L_1 + L_2 + L_3 + L_4. \qquad \text{(S.18)}
\end{aligned}
$$

We have

$$
\begin{aligned}
\mathrm{E}(L_2) &= \frac{1}{2}\mathrm{tr}\left[\mathbf{\Sigma}_k^{1/2}\left(\mathbf{B}_j\mathbf{\Sigma}_{B,j}^{-1}\mathbf{B}_j' - \mathbf{B}_k\mathbf{\Sigma}_{B,k}^{-1}\mathbf{B}_k'\right)\mathbf{\Sigma}_k^{1/2}\right] \\
&= \frac{1}{2}\mathrm{tr}\left(\mathbf{B}_j'\mathbf{\Sigma}_k\mathbf{B}_j\mathbf{\Sigma}_{B,j}^{-1} - \mathbf{I}_{p-d}\right) = \frac{1}{2}\mathrm{tr}(\mathbf{B}_j'\mathbf{\Sigma}_k\mathbf{B}_j\mathbf{\Sigma}_{B,j}^{-1}) - \frac{(p-d)}{2},
\end{aligned}
$$

$\mathrm{E}\left(L_3\right) = \frac{1}{2}\mathrm{tr}\left(\mathbf{B}_j'(\mathbf{\Sigma}_{f,k,t} + \mathbf{U}_{k,j})\mathbf{B}_j\mathbf{\Sigma}_{B,j}^{-1}\right)$, and $\mathrm{E}(L_4) = \mathbf{0}$, so we obtain (25).

To prove (26) , we refer to a fact about multivariate normal random vector. Let $\mathbf{v} \sim N(0, \mathbf{I}_p)$, then for a symmetric matrix $\mathbf{\Sigma}$, $\mathrm{Var}(\mathbf{v}'\mathbf{\Sigma}\mathbf{v}) = 2\|\mathbf{\Sigma}\|_F^2$. Note that $\mathrm{Cov}(L_2, L_3) = \mathrm{Cov}(L_2, L_4) = \mathrm{Cov}(L_3, L_4) = 0$ in (24), define $\mathbf{W}_j = \mathbf{B}_j\mathbf{\Sigma}_{B,j}^{-1}\mathbf{B}_j$, and we have

$$
\begin{aligned}
\mathrm{Var}(w_{t,k,j}) &= \mathrm{Var}(L_2) + \mathrm{Var}(L_3) + \mathrm{Var}(L_4) \\
&= \frac{1}{2}\|\mathbf{\Sigma}_k^{1/2}(\mathbf{W}_j - \mathbf{W}_k)\mathbf{\Sigma}_k^{1/2}\|_F^2 + \frac{1}{4}\mathrm{Var}\left(\mathbf{x}_t'\mathbf{A}_k'\mathbf{W}_j\mathbf{A}_k\mathbf{x}_t\right) + \frac{1}{2}\mathrm{Var}\left(\mathbf{x}_t'\mathbf{A}_k'\mathbf{W}_j(\boldsymbol{\mu}_k - \boldsymbol{\mu}_j)\right) \\
&\quad + \mathrm{Var}\left(\mathbf{x}_t'\mathbf{A}_k'\mathbf{W}_j\boldsymbol{\varepsilon}_t^{(k)}\right) + \mathrm{Var}\left((\boldsymbol{\mu}_k - \boldsymbol{\mu}_j)'\mathbf{W}_j\boldsymbol{\varepsilon}_t^{(k)}\right) \\
&= \frac{1}{2}\|\mathbf{\Sigma}_k^{1/2}(\mathbf{W}_j - \mathbf{W}_k)\mathbf{\Sigma}_k^{1/2}\|_F^2 + \frac{1}{2}\|\mathbf{\Sigma}_k^{1/2}\mathbf{A}_k'\mathbf{W}_j\mathbf{A}_k\mathbf{\Sigma}_k^{1/2}\|_F^2 + \frac{1}{2}\mathrm{tr}\left(\mathbf{\Sigma}_{f,k,t}\mathbf{W}_j\mathbf{U}_{k,j}\mathbf{W}_j\right) \\
&\quad + \mathrm{tr}\left((\mathbf{\Sigma}_{f,k,t} + \mathbf{U}_{k,j})\mathbf{W}_j\mathbf{\Sigma}_k\mathbf{W}_j\right).
\end{aligned}
$$

□