Supplementary Materials for "Adaptive Change Point Monitoring for High-Dimensional Data

Teng Wu, Runmin Wang, Hao Yan, Xiaofeng Shao

Department of Statistics, University of Illinois at Urbana-Champaign

School of Computing Informatics & Decision Systems Engineering, Arizona State University

Supplementary Material

S1 Technical results

Proof of Theorem 1. We can directly apply the results shown in Wang et al. (2019) for the partial sum process

$$S_n(a,b) = \sum_{i=\lfloor na\rfloor+1}^{\lfloor nb\rfloor-1} \sum_{j=\lfloor na\rfloor+1}^{i} X_{i+1}^T X_j.$$

The partial sum process

$$\left\{\frac{\sqrt{2}}{n||\Sigma||_F} S_n(a,b)\right\}_{(a,b)\in[0,T]^2} \rightsquigarrow Q \quad in \quad l^\infty([0,T]^2)$$

where Q is a Gaussian process whose covariance structure is the following

$$Cov(Q(a_1,b_1),(a_2,b_2)) = \begin{cases} (\min(b_1,b_2)-\max(a_1,a_2))^2 & if \quad \max(a_1,a_2)\le \min(b_1,b_2) \\ 0 & otherwise \end{cases}$$

The test statistic is a continuous transformation of the Gaussian process and the results stated follows. □

Proof of Theorem 2. We now analyze the power of the first proposed test. Suppose the change point is at k^*, where $k^*/n \to r$ for some constant $r \in (1, T)$. This assures that the change point does not occur extremely early or late in the monitoring period. Under the alternative hypothesis, define a new sequence of random vectors Y_i,

$$Y_i = \begin{cases} X_i & i = 1, \dots, k^* \\ X_i - \Delta & i = k^* + 1, \dots, n \end{cases}.$$

This sequence does not have a change point. Without loss of generosity, assume Y_i's are centered.

Suppose that

$$\frac{n\Delta^T \Delta}{||\Sigma||_F} \to b \in [0 + \infty).$$

When $m < k < k^*$, $G_k(m)$ statistic will not be affected. It suffices to consider the case $m < k^* < k$ and $k^* < m < k$. Following the decomposition in Wang et al. (2019), under the fixed alternative when $k^* > m$,

$$\begin{aligned} G_k(m) = & G_k^Y(m) + (k - k^*)(k - k^* - 1)m(m - 1)||\Delta||_2^2 \\ & - 2(k - k^*)(k - m - 2)(m - 1) \sum_{j=1}^{m} Y_j^T \Delta \\ & - 4(m - 1)(m - 2)(k - k^*) \sum_{j=m+1}^{k^*} Y_j^T \Delta. \end{aligned}$$

$G_n^Y(m)$ is the statistic calculated for the Y_i sequence. Let $s_n(k) = \sum_{j=1}^{k} Y_j^T\Delta$. Then

$$\sup_{1\le l\le k\le nT} |\sum_{j=l}^{k} Y_j^T\Delta| \le 2 \sup_{1\le k\le nT} |s_n(k)| = O_p(n^{1/2}(\Delta^T\Sigma\Delta)^{1/2}).$$

The last part is obtained by Kolmogorov's inequality. This implies that when $k^* > m$,

$$\frac{1}{n^3\|\Sigma\|_F} G_k(m) = \frac{1}{n^3\|\Sigma\|_F} G_k^Y(m) + \frac{(k-k^*)(k-k^*-1)m(m-1)}{n^3}\frac{||\Delta||_2^2}{||\Sigma||_F} + O_p(\frac{n^{1/2}(\Delta^T\Sigma\Delta)^{1/2}}{||\Sigma||_F}).$$

Similarly, we can show when $k^* > m$

$$\frac{1}{n^3\|\Sigma\|_F} G_k(m) = \frac{1}{n\|\Sigma\|_F} G_k^Y(m) + \frac{k^*(k^*-1)(k-m)(k-m-1)}{n^3}\frac{||\Delta||_2^2}{||\Sigma||_F} + O_p(\frac{n^{1/2}(\Delta^T\Sigma\Delta)^{1/2}}{||\Sigma||_F}).$$

The last part is converging to 0 in probability. Therefore, the test statistic T_n can be viewed as an extension to the original process. The second terms are also a process depend on m and k^*. Under the fixed alternative, the $G_k(m)$ converge to the process

$$\frac{1}{n^3\|\Sigma\|_F}\{G_{\lfloor nt\rfloor}(\lfloor ns\rfloor)\}_{s\in[0,1]} \to G(s,t) + b\Lambda(s,t),$$

where

$$\Lambda(s,t) = \begin{cases} (t-r)^2 s^2 & s \le r \\ r^2(t-s)^2 & s > r \\ 0 & otherwise \end{cases}.$$

This implies that, when $b = 0$, the process is the same with the null process, and the proposed monitoring scheme will have trivial power. When the b is not zero, since the remainder term is positive, we will have non -trivial power.

When

$$\frac{n\Delta^T\Delta}{||\Sigma||_F} \to \infty.$$

Following above decomposition, we have

$$\max_k T_n(k) \geq T_n(k^*) = \frac{1}{n\|\Sigma\|_F} D^Y_{nT}(k^*) + O(\frac{n||\Delta||_2^2}{||\Sigma||_F}) \to \infty$$

Since the first term is pivotal and is bounded in probability, the test have power converging to 1. □

Proof of Theorem 3. We can directly apply the results in Theorem 2.1 and 2.2 in Zhang et al.(2020), which stated that for

$$S_{n,q,c}(r;[a,b]) = \sum_{l=1}^{p} \sum^{*}_{\lfloor na\rfloor+1\leq i_1,\dots,i_c\leq\lfloor nr\rfloor} \sum^{*}_{\lfloor nr\rfloor+1\leq j_1,\dots,j_{q-c}\leq\lfloor nb\rfloor} \prod_{t=1}^{c} X_{i_t,l} \prod_{g=1}^{q-c} X_{j_g,l},$$

we have

$$\frac{1}{\sqrt{n^q\|\Sigma\|_q^q}} S_{n,q,c}(r;[a,b]) \rightsquigarrow Q_{q,c}(r;[a,b]),$$

where $Q_{q,c}$ is the Gaussian process stated in Theorem 4. The monitoring statistic is a continuous transformation of process $S_{n,q,c}$'s and the asymptotic result follows. □

Proof of Theorem 4. We first discuss the case when $\frac{n^{q/2}\|\Delta\|_q^q}{\|\Sigma\|_q^{q/2}} \to \gamma \in [0,+\infty)$ and the true change point is at location $k^* = \lfloor nr\rfloor$. Here we adopt the process convergence results in Theorem 2.3 of Zhang et al.(2020), which stated that for $(k,m) =$

$(\lfloor ns \rfloor, \lfloor nt \rfloor)$,

$$\frac{1}{\sqrt{n^{3q}\|\Sigma\|_q^q}} D_{n,q}(s;[0,b]) = \frac{1}{\sqrt{n^{3q}\|\Sigma\|_q^q}} \sum_{l=1}^{p} \sum_{0 \le i_1,\ldots,i_q \le k}^{*} \sum_{k+1 \le j_1,\ldots,j_q \le m}^{*} (X_{i_1,l} - X_{j_1,l}) \cdots (X_{i_q,l} - X_{j_q,l}),$$

$$\rightsquigarrow G_q(s,t) + \gamma J_q(s;[0,t])$$

where

$$J_q(s;[0,t]) = \begin{cases} r^q(t-s)^q & r \le s < t \\ s^q(t-r)^q & s \le r < t \\ 0 & otherwise \end{cases}$$

Therefore, by continuous mapping theorem, when $\gamma \in [0, +\infty)$, the results in the theorem hold.

For the case $\frac{n^{q/2}\|\Delta\|_q^q}{\|\Sigma\|_q^{q/2}} \to +\infty$

$$\max_k T_{n,q}(k) \ge T_{n,q}(k^*) = \frac{1}{n\|\Sigma\|_F} D_{nT}^{Y}(k^*) + C\frac{n^{q/2}\|\Delta\|_q^q}{\|\Sigma\|_q^{q/2}} \to \infty$$

$\square$

Proof of Theorem 5. By straightforward calculation, we have

$$\begin{aligned}
\widehat{\|\Sigma\|_F^2} &= \frac{1}{4\binom{n}{4}} \sum_{1\le j_1<j_2<j_3<j_4\le n} tr\left((X_{j_1}-X_{j_2})(X_{j_1}-X_{j_2})^T(X_{j_3}-X_{j_4})(X_{j_3}-X_{j_4})^T\right)\\
&= \frac{1}{4\binom{n}{4}} \sum_{1\le j_1<j_2<j_3<j_4\le n} [(X_{j_1}-X_{j_2})^T(X_{j_3}-X_{j_4})]^2\\
&= \frac{1}{4\binom{n}{4}} \sum_{1\le j_1<j_2<j_3<j_4\le n} [(X_{j_1}^TX_{j_3})^2+(X_{j_2}^TX_{j_3})^2+(X_{j_2}^TX_{j_4})^2+(X_{j_1}^TX_{j_4})^2]\\
&- \frac{2}{4\binom{n}{4}} \sum_{1\le j_1<j_2<j_3<j_4\le n} [X_{j_1}^TX_{j_3}X_{j_1}^TX_{j_4}+X_{j_2}^TX_{j_3}X_{j_2}^TX_{j_4}+X_{j_1}^TX_{j_3}X_{j_2}^TX_{j_3}+X_{j_1}^TX_{j_4}X_{j_2}^TX_{j_4}]\\
&+ \frac{2}{4\binom{n}{4}} \sum_{1\le j_1<j_2<j_3<j_4\le n} [X_{j_1}^TX_{j_3}X_{j_2}^TX_{j_4}+X_{j_2}^TX_{j_3}X_{j_1}^TX_{j_4}]\\
&= I_{n,1}+I_{n,2}+I_{n,3}+I_{n,4}-(I_{n,5}+I_{n,6}+I_{n,7}+I_{n,8})+(I_{n,9}+I_{n,10}).
\end{aligned}$$

For $I_{n,1}$,

$$\mathbb{E}[I_{n,1}] = \frac{1}{4\binom{n}{4}} \sum_{1\le j_1<j_2<j_3<j_4\le n} \mathbb{E}[(X_{j_1}^TX_{j_3})^2] = \frac{1}{4}tr(\mathbb{E}[X_{j_3}X_{j_3}^TX_{j_1}X_{j_1}^T]) = \|\Sigma\|_F^2/4.$$

Thus $\mathbb{E}[I_{n,1}/\|\Sigma\|_F^2] = 1/4$. By similar arguments, it is obvious to see that $\mathbb{E}[I_{n,i}/\|\Sigma\|_F^2] = 1/4$ for $i = 1,2,3,4$, and $\mathbb{E}[I_{n,i}/\|\Sigma\|_F^2] = 0$ for $i = 5, ..., 10$.

The outline of the proof is as following. We will show that $4I_{n,i}/\|\Sigma\|_F^2 \to_p 1$ for $i = 1,2,3,4$, and $I_{n,i}/\|\Sigma\|_F^2 \to_p 0$, for $i = 5, ..., 10$. Since some of the $I_{n,i}$ share very similar structures, we will only present the proof for (1) $4I_{n,1}/\|\Sigma\|_F^2 \to_p 1$ and (2) $I_{n,5}/\|\Sigma\|_F^2 \to_p 0$. Other terms can be proved by similar arguments.

To show (1), it suffices to show that $\mathbb{E}[16I_{n,1}^2/\|\Sigma\|_F^4] \to 1$. To see this,

$$\mathbb{E}[16I_{n,1}^2/\|\Sigma\|_F^4] = \frac{1}{\binom{n}{4}^2\|\Sigma\|_F^4} \sum_{1\leq j_1<j_2<j_3<j_4\leq n} \sum_{1\leq j_5<j_6<j_7<j_8\leq n} \mathbb{E}[(X_{j_1}^T X_{j_3})^2 (X_{j_5}^T X_{j_7})^2]$$
$$= \frac{1}{\binom{n}{4}^2\|\Sigma\|_F^4} \sum_{1\leq j_1<j_2<j_3<j_4\leq n} \sum_{1\leq j_5<j_6<j_7<j_8\leq n} \sum_{l_1,l_2,l_3,l_4=1}^{p} \mathbb{E}[X_{j_1,l_1}X_{j_3,l_1}X_{j_1,l_2}X_{j_3,l_2}X_{j_5,l_3}X_{j_7,l_3}X_{j_5,l_4}X_{j_7,l_4}].$$

As we know that the expectation of a product of random variables can be expressed in terms of joint cumulants, we have

$$\mathbb{E}[X_{j_1,l_1}X_{j_3,l_1}X_{j_1,l_2}X_{j_3,l_2}X_{j_5,l_3}X_{j_7,l_3}X_{j_5,l_4}X_{j_7,l_4}] = \sum_{\pi}\prod_{B\in\pi} cum(X_{j,l} : (j,l) \in B),$$

where π runs through the list of all partitions of $\{(j_1,l_1),(j_1,l_2),...,(j_7,l_3),(j_7,l_4)\}$ and B runs through the list of all blocks of the partition π. Since $j_1 < j_3$ and $j_5 < j_7$, it is impossible to have three or more indices in $\{j_1,j_3,j_5,j_7\}$ such that they are identical. Thus for the right hand side of the above expression, we only need to take the sum over all partitions with all block sizes smaller than 5, because for joint cumulants with order greater than 5, it must contain at least 3 indices from j_1,j_3,j_5,j_7 and at least one is not identical to the other two. And the joint cumulants will equal to zero since it involves two or more independent random variables.

Also since the mean of all random variables included in the left hand side of the above expression are all zero, we do not need to consider the partition with block size 1. Thus the expression can be simplified as

$$\begin{aligned}&\mathbb{E}[X_{j_1,l_1}X_{j_3,l_1}X_{j_1,l_2}X_{j_3,l_2}X_{j_5,l_3}X_{j_7,l_3}X_{j_5,l_4}X_{j_7,l_4}]\\=&C_1^{(j_1,j_3,j_5,j_7)}\mathbb{E}[X_{0,l_1}X_{0,l_2}X_{0,l_3}X_{0,l_4}]^2+C_2^{(j_1,j_3,j_5,j_7)}\mathbb{E}[X_{0,l_1}X_{0,l_2}X_{0,l_3}X_{0,l_4}]\Sigma_{l_1,l_2}\Sigma_{l_3,l_4}\\&+\Sigma_{l_1,l_2}^2\Sigma_{l_3,l_4}^2,\end{aligned}$$

where $C_1^{(j_1,j_3,j_5,j_7)}$, $C_2^{(j_1,j_3,j_5,j_7)}$ are finite positive constants purely based on the value of j_1, j_3, j_5, j_7. $C_1^{(j_1,j_3,j_5,j_7)}$ can only be nonzero if $j_1 = j_5$ and $j_3 = j_7$, and $C_2^{(j_1,j_3,j_5,j_7)}$ is nonzero if at least two of (j_1, j_3, j_5, j_7) are equal. This implies that

$$\sum_{1\le j_1<j_2<j_3<j_4\le n}\ \sum_{1\le j_5<j_6<j_7<j_8\le n} C_1^{(j_1,j_3,j_5,j_7)} = o(n^8),$$

and

$$\sum_{1\le j_1<j_2<j_3<j_4\le n}\ \sum_{1\le j_5<j_6<j_7<j_8\le n} C_2^{(j_1,j_3,j_5,j_7)} = o(n^8).$$

Furthermore, according to Assumption 2, $\sum_{l_1,l_2,l_3,l_4=1}^{p} cum(X_{0,l_1}, X_{0,l_2}, X_{0,l_3}, X_{0,l_4})^2 \le C\|\Sigma\|_F^4$. It can be verified that

$$\begin{aligned}\sum_{l_1,l_2,l_3,l_4=1}^{p}\mathbb{E}[X_{0,l_1}X_{0,l_2}X_{0,l_3}X_{0,l_4}]^2 \lesssim& \sum_{l_1,l_2,l_3,l_4=1}^{p} cum(X_{0,l_1}, X_{0,l_2}, X_{0,l_3}, X_{0,l_4})^2\\&+\sum_{l_1,l_2,l_3,l_4=1}^{p}\Sigma_{l_1,l_2}^2\Sigma_{l_3,l_4}^2\\ \lesssim& \|\Sigma\|_F^4,\end{aligned}$$

and by using the Cauchy-Schwartz inequaility,

$$\sum_{l_1,l_2,l_3,l_4=1}^{p} \mathbb{E}[X_{0,l_1}X_{0,l_2}X_{0,l_3}X_{0,l_4}]\Sigma_{l_1,l_2}\Sigma_{l_3,l_4}$$
$$\leq\sqrt{\sum_{l_1,l_2,l_3,l_4=1}^{p} \mathbb{E}[X_{0,l_1}X_{0,l_2}X_{0,l_3}X_{0,l_4}]^2}\sqrt{\sum_{l_1,l_2,l_3,l_4=1}^{p} \Sigma_{l_1,l_2}^2\Sigma_{l_3,l_4}^2} \leq \sqrt{C}\|\Sigma\|_F^4. \tag{S1.1}$$

This indicates that

$$\begin{aligned}
&\mathbb{E}[16I_{n,1}^2/\|\Sigma\|_F^4] \\
=&\frac{1}{\binom{n}{4}^2\|\Sigma\|_F^4}\sum_{1\leq j_1<j_2<j_3<j_4\leq n}\ \sum_{1\leq j_5<j_6<j_7<j_8\leq n} C_1^{(j_1,j_3,j_5,j_7)} \sum_{l_1,l_2,l_3,l_4=1}^{p} \mathbb{E}[X_{0,l_1}X_{0,l_2}X_{0,l_3}X_{0,l_4}]^2 \\
+&\frac{1}{\binom{n}{4}^2\|\Sigma\|_F^4}\sum_{1\leq j_1<j_2<j_3<j_4\leq n}\ \sum_{1\leq j_5<j_6<j_7<j_8\leq n} C_2^{(j_1,j_3,j_5,j_7)} \sum_{l_1,l_2,l_3,l_4=1}^{p} \mathbb{E}[X_{0,l_1}X_{0,l_2}X_{0,l_3}X_{0,l_4}]\Sigma_{l_1,l_2}\Sigma_{l_3,l_4} \\
+&\frac{1}{\binom{n}{4}^2\|\Sigma\|_F^4}\sum_{1\leq j_1<j_2<j_3<j_4\leq n}\ \sum_{1\leq j_5<j_6<j_7<j_8\leq n} \sum_{l_1,l_2,l_3,l_4=1}^{p} \Sigma_{l_1,l_2}^2\Sigma_{l_3,l_4}^2 = o(1)+o(1)+1 \to 1.
\end{aligned}$$

Thus, $4I_{n,1}/\|\Sigma\|_F^2 \to_p 1$, and (1) is proved. By similar arguments, $4I_{n,i}/\|\Sigma\|_F^2 \to_p 1$ holds for $i = 2, 3, 4$.

To show (2), we need to prove $\mathbb{E}[I_{n,5}^2/\|\Sigma\|_F^4] \to 0$. To see this,

$$\begin{aligned}
\mathbb{E}[I_{n,5}^2/\|\Sigma\|_F^4] =& \frac{1}{4\binom{n}{4}^2\|\Sigma\|_F^4}\sum_{1\leq j_1<j_2<j_3<j_4\leq n}\ \sum_{1\leq j_5<j_6<j_7<j_8\leq n} \mathbb{E}[(X_{j_1}^TX_{j_3}X_{j_1}^TX_{j_4})(X_{j_5}^TX_{j_7}X_{j_5}^TX_{j_8})] \\
=&\frac{1}{\binom{n}{4}^2\|\Sigma\|_F^4}\sum_{1\leq j_1<j_2<j_3<j_4\leq n}\ \sum_{1\leq j_5<j_6<j_7<j_8\leq n} \sum_{l_1,l_2,l_3,l_4=1}^{p} \mathbb{E}[X_{j_1,l_1}X_{j_3,l_1}X_{j_1,l_2}X_{j_4,l_2}X_{j_5,l_3}X_{j_7,l_3}X_{j_5,l_4}X_{j_8,l_4}].
\end{aligned}$$

By similar arguments for the joint cumulants we provided in the the proof for

(1), it can be proved that

$$
\begin{aligned}
&\mathbb{E}[X_{j_1,l_1}X_{j_3,l_1}X_{j_1,l_2}X_{j_4,l_2}X_{j_5,l_3}X_{j_7,l_3}X_{j_5,l_4}X_{j_8,l_4}]\\
=&C_1^{(j_1,j_3,j_4,j_5,j_7,j_8)}\mathbb{E}[X_{0,l_1}X_{0,l_2}X_{0,l_3}X_{0,l_4}]\Sigma_{l_1,l_3}\Sigma_{l_2,l_4}+C_2^{(j_1,j_3,j_4,j_5,j_7,j_8)}\Sigma_{l_1,l_2}\Sigma_{l_3,l_4}\Sigma_{l_1,l_3}\Sigma_{l_2,l_4}.
\end{aligned}
$$

If $C_1^{(j_1,j_3,j_4,j_5,j_7,j_8)} \neq 0$, then $j_1 = j_5$. And if $C_2^{(j_1,j_3,j_4,j_5,j_7,j_8)} \neq 0$, $j_3 = j_5$ and $j_4 = j_8$. These two properties guarantee that

$$
\sum_{1\le j_1<j_2<j_3<j_4\le n}\ \sum_{1\le j_5<j_6<j_7<j_8\le n} C_1^{(j_1,j_3,j_4,j_5,j_7,j_8)} = o(n^8),
$$

and

$$
\sum_{1\le j_1<j_2<j_3<j_4\le n}\ \sum_{1\le j_5<j_6<j_7<j_8\le n} C_2^{(j_1,j_3,j_4,j_5,j_7,j_8)} = o(n^8).
$$

Furthermore we have shown the bound for $\sum_{l_1,l_2,l_3,l_4=1}^{p}\mathbb{E}[X_{0,l_1}X_{0,l_2}X_{0,l_3}X_{0,l_4}]\Sigma_{l_1,l_2}\Sigma_{l_3,l_4}$ in (S1.2). And

$$
\begin{aligned}
&\sum_{l_1,l_2,l_3,l_4=1}^{p}\Sigma_{l_1,l_2}\Sigma_{l_3,l_4}\Sigma_{l_1,l_3}\Sigma_{l_2,l_4} = \sum_{l_1,l_4=1}^{p}\left(\sum_{l_2=1}^{p}\Sigma_{l_1,l_2}\Sigma_{l_2,l_4}\right)\left(\sum_{l_3=1}^{p}\Sigma_{l_1,l_3}\Sigma_{l_3,l_4}\right)\\
&=\sum_{l_1,l_4=1}^{p}[(\Sigma^2)_{l_1,l_4}]^2 = tr(\Sigma^4) = o(\|\Sigma\|_F^4),
\end{aligned}
$$

by Assumption 1. Thus,

$$
\begin{aligned}
&\mathbb{E}[I_{n,5}^2/\|\Sigma\|_F^4]\\
=&\frac{1}{4\binom{n}{4}^2\|\Sigma\|_F^4}\sum_{1\le j_1<j_2<j_3<j_4\le n}\ \sum_{1\le j_5<j_6<j_7<j_8\le n} C_1^{(j_1,j_3,j_4,j_5,j_7,j_8)}\sum_{l_1,l_2,l_3,l_4=1}^{p}\mathbb{E}[X_{0,l_1}X_{0,l_2}X_{0,l_3}X_{0,l_4}]\Sigma_{l_1,l_3}\Sigma_{l_2,l_4}\\
+&\frac{1}{4\binom{n}{4}^2\|\Sigma\|_F^4}\sum_{1\le j_1<j_2<j_3<j_4\le n}\ \sum_{1\le j_5<j_6<j_7<j_8\le n} C_2^{(j_1,j_3,j_4,j_5,j_7,j_8)}\sum_{l_1,l_2,l_3,l_4=1}^{p}\Sigma_{l_1,l_2}\Sigma_{l_3,l_4}\Sigma_{l_1,l_3}\Sigma_{l_2,l_4}\\
=&o(1)+o(1)\to 1.
\end{aligned}
$$

This indicates $I_{n,5}/\|\Sigma\|_F^2 \to_p 0$. And by similar arguments we can prove that $I_{n,i}/\|\Sigma\|_F^2 \to_p 0$, for all $i = 6, ..., 10$. Combine the above results, we have $\widehat{\|\Sigma\|_F^2}/\|\Sigma\|_F^2 \to_p 1$. This completes the proof.

$\square$

Proof of Theorem 6. We can rewrite $\widehat{\|\Sigma\|_q^q}$ as

$$
\begin{aligned}
\widehat{\|\Sigma\|_q^q} &= \frac{1}{2^q\binom{n}{2q}} \sum_{l_1,l_2=1}^{p} \sum_{1\le i_1<\cdots<i_q<j_1<\cdots<j_q\le n} \prod_{k=1}^{q} (X_{i_k,l_1}X_{i_k,l_2} + X_{j_k,l_1}X_{j_k,l_2} - X_{i_k,l_1}X_{j_k,l_2} - X_{j_k,l_1}X_{i_k,l_2}) \\
&= \frac{1}{2^q\binom{n}{2q}} \sum_{1\le i_1<\cdots<i_q<j_1<\cdots<j_q\le n} \sum_{t_1,s_1\in\{i_1,j_1\}} \cdots \sum_{t_q,s_q\in\{i_q,j_q\}} \sum_{l_1,l_2=1}^{p} \prod_{k=1}^{q} (-1)^{\mathbf{1}\{t_k\neq s_k\}} X_{t_k,l_1}X_{s_k,l_2} \\
&= \frac{1}{2^q\binom{n}{2q}} \sum_{1\le i_1<\cdots<i_q<j_1<\cdots<j_q\le n} \sum_{t_1\in\{i_1,j_1\}} \cdots \sum_{t_q\in\{i_q,j_q\}} \sum_{l_1,l_2=1}^{p} \prod_{k=1}^{q} X_{t_k,l_1}X_{t_k,l_2} \\
&\quad + \frac{1}{2^q\binom{n}{2q}} \sum_{1\le i_1<\cdots<i_q<j_1<\cdots<j_q\le n} \sum_{t_1,s_1\in\{i_1,j_1\}} \cdots \sum_{t_q,s_q\in\{i_q,j_q\}} \\
&\qquad \sum_{l_1,l_2=1}^{p} \mathbf{1}\{\cup_{k=1}^{q}\{t_k\neq s_k\}\} \prod_{k=1}^{q} (-1)^{\mathbf{1}\{t_k\neq s_k\}} X_{t_k,l_1}X_{s_k,l_2}.
\end{aligned}
$$

The second equality in the above expression is by calculating the cross products among q brackets, and the third equality is splitting the terms based on different values of t_k, s_k for $k = 1, ..., q$. The first term in the third equality contains all products with $t_k = s_k$ for all $k = 1, ..., q$, and the second term contains products with at least one $k = 1, ..., q$ such that $t_k \neq s_k$.

The outline of the proof is as follows. We want to show:

1. for every $t_1 \in \{i_1, j_1\}, ..., t_q \in \{i_q, j_q\}$,

$$I(t_1, ..., t_q) = \frac{1}{\binom{n}{2q}\|\Sigma\|_q^q} \sum_{1 \le i_1 < \cdots < i_q < j_1 < \cdots < j_q \le n} \sum_{l_1, l_2 = 1}^{p} \prod_{k=1}^{q} X_{t_k, l_1} X_{t_k, l_2} \to_p 1;$$

2. for every $t_1, s_1 \in \{i_1, j_1\}, ..., t_q, s_q \in \{i_q, j_q\}$ and there exists at least one $k = 1, ..., q$ such that $t_k \neq s_k$,

$$J(t_1, s_1, ..., t_q, s_q) = \frac{1}{\binom{n}{2q}\|\Sigma\|_q^q} \sum_{1 \le i_1 < \cdots < i_q < j_1 < \cdots < j_q \le n} \sum_{l_1, l_2 = 1}^{p} \prod_{k=1}^{q} X_{t_k, l_1} X_{s_k, l_2} \to_p 0.$$

And it is easy to see that if these two results hold, then $\widehat{\|\Sigma\|_q^q}/\|\Sigma\|_q^q \to_p 1$. As we observe that most of terms are structurally very similar, we shall only present the proof for $I(i_1, ..., i_q) \to_p 1$ and a general proof for (2).

It is trivial to see that $\mathbb{E}[\sum_{l_1, l_2 = 1}^{p} \prod_{k=1}^{q} X_{t_k, l_1} X_{t_k, l_2} / \|\Sigma\|_q^q] = 1$. This indicates that to show (1), it suffices to show that $\mathbb{E}[I(i_1, ..., i_q)^2] \to 1$. To show this,

$$\begin{aligned}\mathbb{E}[I(i_1, ..., i_q)^2] =& \frac{1}{\binom{n}{2q}^2\|\Sigma\|_q^{2q}} \sum_{1 \le i_1 < \cdots < i_q < j_1 < \cdots < j_q \le n} \sum_{1 \le i_1' < \cdots < i_q' < j_1' < \cdots < j_q' \le n} \\ & \sum_{l_1, l_2, l_3, l_4 = 1}^{p} \mathbb{E}\left[\prod_{k=1}^{q} X_{i_k, l_1} X_{i_k, l_2} X_{i_k', l_3} X_{i_k', l_4}\right].\end{aligned}$$

Due to the special structure of our statistic,

$$\mathbb{E}\left[\prod_{k=1}^{q} X_{i_k, l_1} X_{i_k, l_2} X_{i_k', l_3} X_{i_k', l_4}\right] = \sum_{m=0}^{q} C_m E(X_{0, l_1} X_{0, l_2} X_{0, l_3} X_{0, l_4})^m (\Sigma_{l_1, l_2} \Sigma_{l_3, l_4})^{q-m},$$

where $C_m = C_m(i_1, ..., i_q, i_1', ..., i_q') \ge 0$ is a function of all indices for all $m = 1, 2, ..., q$. $C_m = 1$ if there are exact m indices in $\{i_1, ..., i_q\}$ which equal to m indices in

$\{i'_1, ..., i'_q\}$, and $C_m = 0$ otherwise. These events are mutually exclusive which indicates that $\sum_{m=0}^{q} C_m = 1$. This indicates that for all $m = 1, ..., q$,

$$\sum_{1 \le i_1 < \cdots < i_q < j_1 < \cdots < j_q \le n} \sum_{1 \le i'_1 < \cdots < i'_q < j'_1 < \cdots < j'_q \le n} C_m(i_1, ..., i_q, i'_1, ..., i'_q) = o(n^{4q}),$$

and

$$\frac{1}{\binom{n}{2q}^2} \sum_{1 \le i_1 < \cdots < i_q < j_1 < \cdots < j_q \le n} \sum_{1 \le i'_1 < \cdots < i'_q < j'_1 < \cdots < j'_q \le n} C_0(i_1, ..., i_q, i'_1, ..., i'_q) \to 1.$$

Furthermore, for any $m = 1, ..., q$, by Hölder's inequaility for vector spaces, we have

$$\begin{aligned}
&\sum_{l_1,l_2,l_3,l_4=1}^{p} |E(X_{0,l_1}X_{0,l_2}X_{0,l_3}X_{0,l_4})|^m |\Sigma_{l_1,l_2}\Sigma_{l_3,l_4}|^{q-m} \\
\le&\left(\sum_{l_1,l_2,l_3,l_4=1}^{p} |E(X_{0,l_1}X_{0,l_2}X_{0,l_3}X_{0,l_4})^m|^{q/m}\right)^{m/q} \left(\sum_{l_1,l_2,l_3,l_4=1}^{p} (|\Sigma_{l_1,l_2}\Sigma_{l_3,l_4}|^{q-m})^{q/(q-m)}\right)^{(q-m)/q} \\
=&\left(\sum_{l_1,l_2,l_3,l_4=1}^{p} |E(X_{0,l_1}X_{0,l_2}X_{0,l_3}X_{0,l_4})|^q\right)^{m/q} \left(\sum_{l_1,l_2,l_3,l_4=1}^{p} |\Sigma_{l_1,l_2}|^q|\Sigma_{l_3,l_4}|^q\right)^{(q-m)/q} \\
\le& C\|\Sigma\|_q^{2m}\|\Sigma\|_q^{2(q-m)} = C\|\Sigma\|_q^{2q},
\end{aligned}$$

where the last inequality is due to Assumption 5, and to see this,

$$\begin{aligned}
&\sum_{l_1,l_2,l_3,l_4=1}^{p} |E(X_{0,l_1}X_{0,l_2}X_{0,l_3}X_{0,l_4})|^q \\
\le& C\sum_{l_1,l_2,l_3,l_4=1}^{p} |cum(X_{0,l_1}, X_{0,l_2}, X_{0,l_3}, X_{0,l_4})|^q + |\Sigma_{l_1,l_2}\Sigma_{l_3,l_4}|^q + |\Sigma_{l_1,l_3}\Sigma_{l_2,l_4}|^q + |\Sigma_{l_1,l_4}\Sigma_{l_2,l_3}|^q \\
\le& C\sum_{1 \le l_1 \le l_2 \le l_3 \le l_4 \le p} (1 \vee (l_4 - l_1))^{-2rq} + 3C\|\Sigma\|_q^{2q} \le Cp^2 + 3C\|\Sigma\|_q^{2q} \le C\|\Sigma\|_q^{2q},
\end{aligned}$$

for some generic positive constant C, since $\|\Sigma\|_q^{2q} = (\sum_{i,j=1}^{p} \Sigma_{i,j}^q)^2 \ge Cp^2$ under

Assumption (5.1). Therefore,

$$
\begin{aligned}
&\mathbb{E}[I(i_1,...,i_q)^2]\\
=&\frac{1}{\binom{n}{2q}^2\|\Sigma\|_q^{2q}}\sum_{1\le i_1<\cdots<i_q<j_1<\cdots<j_q\le n}\sum_{1\le i_1'<\cdots<i_q'<j_1'<\cdots<j_q'\le n}\sum_{l_1,l_2,l_3,l_4=1}^{q}\mathbb{E}\left[\prod_{k=1}^{p}X_{i_k,l_1}X_{i_k,l_2}X_{i_k',l_3}X_{i_k',l_4}\right]\\
=&\frac{1}{\binom{n}{2q}^2\|\Sigma\|_q^{2q}}\sum_{1\le i_1<\cdots<i_q<j_1<\cdots<j_q\le n}\sum_{1\le i_1'<\cdots<i_q'<j_1'<\cdots<j_q'\le n}\\
&\quad\sum_{l_1,l_2,l_3,l_4=1}^{p}\sum_{m=0}^{q}C_m\mathbb{E}(X_{0,l_1}X_{0,l_2}X_{0,l_3}X_{0,l_4})^m(\Sigma_{l_1,l_2}\Sigma_{l_3,l_4})^{q-m}\\
=&\frac{1}{\binom{n}{2q}^2\|\Sigma\|_q^{2q}}\sum_{1\le i_1<\cdots<i_q<j_1<\cdots<j_q\le n}\sum_{1\le i_1'<\cdots<i_q'<j_1'<\cdots<j_q'\le n}C_0(i_1,...,i_q,i_1',...,i_q')\sum_{l_1,l_2,l_3,l_4=1}^{p}\Sigma_{l_1,l_2}^q\Sigma_{l_3,l_4}^q\\
+&\frac{1}{\binom{n}{2q}^2\|\Sigma\|_q^{2q}}o(n^{4q})\sum_{l_1,l_2,l_3,l_4=1}^{p}\sum_{m=1}^{q}\mathbb{E}(X_{0,l_1}X_{0,l_2}X_{0,l_3}X_{0,l_4})^m(\Sigma_{l_1,l_2}\Sigma_{l_3,l_4})^{q-m}\\
=&1+o(1)\to 1.
\end{aligned}
$$

This completes the proof for (1). To show (2), it suffices to show that $\mathbb{E}[J(t_1,s_1,...,t_q,s_q)^2]\to 0$. Specifically,

$$
\begin{aligned}
\mathbb{E}[J(t_1,s_1,...,t_q,s_q)^2]=&\frac{1}{\binom{n}{2q}^2\|\Sigma\|_q^{2q}}\sum_{1\le i_1<\cdots<i_q<j_1<\cdots<j_q\le n}\sum_{1\le i_1'<\cdots<i_q'<j_1'<\cdots<j_q'\le n}\\
&\sum_{l_1,l_2,l_3,l_4=1}^{p}\mathbb{E}\left[\prod_{k=1}^{q}X_{t_k,l_1}X_{s_k,l_2}X_{t_k',l_3}X_{s_k',l_4}\right],
\end{aligned}
$$

for $t_1,s_1\in\{i_1,j_1\},...,t_q,s_q\in\{i_q,j_q\},t_1',s_1'\in\{i_1',j_1'\},...,t_q',s_q'\in\{i_q',j_q'\}$, and there exists at least one $k=1,...,q$ such that $t_k\neq s_k$ ($t_k'\neq s_k'$). Since the expectation of a product of random variables can be expressed in terms of joint cumulants, we have

$$\mathbb{E}\left[\prod_{k=1}^{q} X_{t_k,l_1} X_{s_k,l_2} X_{t'_k,l_3} X_{s'_k,l_4}\right] = \sum_{\pi} \prod_{B\in\pi} cum(X_{i,l} : (i,l) \in B),$$

where π runs through the list of all partitions of $\{(t_1, l_1), (s_1, l_2), ..., (t'_q, l_3), (s'_q, l_4)\}$ and B runs thorough the list of all blocks of the partition π. Due to the special structure of our statistic, there is a set of partitions $\mathcal{S}$ such that for every $\pi \in \mathcal{S}$, the product of joint cumulants over all $B \in \pi$ is zero. And for each $\pi \in \mathcal{S}^c$ there are nice properties related to the blocks $B \in \pi$. Here we shall illustrate these properties as follows. To be clear, since we are dealing with a double indexed array $X_{i,l}$, we call "i" as the temporal index and "l" as the spatial index. For $\forall\pi \in \mathcal{S}^c$,

1. **The size of every block $B \in \pi$ cannot exceed 4.** Since $i_1, ..., i_q, j_1, ..., j_q$ are all distinct, and $i'_1, ..., i'_q, j'_1, ..., j'_q$ are all distinct, it is impossible to have any three indices in $\{i_1, ..., i_q, j_1, ..., j_q, i'_1, ..., i'_q, j'_1, ..., j'_q\}$ that are equal. And any joint cumulants of order greater than or equal to 5 will include at least three indices and they cannot be all equal.

2. **There are no blocks with size 1.** This is because the cumulant of a single random variable with mean zero is also zero.

3. **Every $B \in \pi$ must contain only one distinct temporal index.** Otherwise $\prod_{B\in\pi} cum(X_{i,l} : (i,l) \in B) = 0$.

The above properties imply that for $\forall\pi \in \mathcal{S}^c$ and $\forall B \in \pi$, $cum(X_{i,l} : (i,l) \in B)$

has to be one of the followings: $cum(X_{0,l_1}, X_{0,l_2}, X_{0,l_3}, X_{0,l_4})$, $cum(X_{0,l_1}, X_{0,l_2}, X_{0,l_3})$, $cum(X_{0,l_1}, X_{0,l_2}, X_{0,l_4})$, $cum(X_{0,l_1}, X_{0,l_3}, X_{0,l_4})$, $cum(X_{0,l_2}, X_{0,l_3}, X_{0,l_4})$, Σ_{l_1,l_2}, Σ_{l_1,l_3}, Σ_{l_1,l_4}, Σ_{l_2,l_3}, Σ_{l_2,l_4}, Σ_{l_3,l_4}.

If we assume $l_1 \le l_2 \le l_3 \le l_4$, it can be shown that

$$\prod_{B\in\pi} cum(X_{i,l} : (i,l) \in B) \le C(1 \vee (l_2 - l_1))^{-r}(1 \vee (l_4 - l_3))^{-r}, \tag{S1.2}$$

for some generic positive constant C and any partition π. To see this, we notice that at least one $k = 1, ..., q$, say k_0, such that $t_{k_0} \neq s_{k_0}$ and $t'_{k_0} \neq s'_{k_0}$. For every $\pi \in \mathcal{S}^c$ there exists $B_1, B_2 \in \pi$ such that $(t_{k_0}, l_1) \in B_1$ and $(s'_{k_0}, l_4) \in B_2$. Based on the third property above, all other elements in B_1 must have the same temporal index as t_{k_0}. And because of the first property above, all i_k, j_k for $k \neq k_0$ and s_{k_0} are different from t_{k_0}. This implies that the spatial indices for all other elements in B_1 have to be either l_3 or l_4, not l_1 and l_2. For the same reason, the spatial indices for all other elements in B_2 can only be either l_1 or l_2. Therefore,

$$cum(X_{i,l} : (i,l) \in B_1) \in \{cum(X_{0,l_1}, X_{0,l_3}, X_{0,l_4}), \Sigma_{l_1,l_3}, \Sigma_{l_1,l_4}\},$$

and

$$cum(X_{i,l} : (i,l) \in B_2) \in \{cum(X_{0,l_1}, X_{0,l_2}, X_{0,l_4}), \Sigma_{l_1,l_4}, \Sigma_{l_2,l_4}\}.$$

Under Assumption (5.2), $cum(X_{i,l} : (i,l) \in B_1) \le C(1 \vee (l_2 - l_1))^{-r}$ and $cum(X_{i,l} : (i,l) \in B_2) \le C(1 \vee (l_4 - l_3))^{-r}$. And the joint cumulants are uniformly bounded above for those $B \in \pi \setminus \{B_1, B_2\}$. Thus Equation S1.2 is proved.

Furthermore, define $Ind(t_1, s_1, ..., t_q, s_q, t'_1, s'_1, ..., t'_q, s'_q)$ as the indicator function corresponding to the event that for every $k = 1, 2, .., q$ that $t_k \neq s_k$, there exists $k' = 1, ..., q$ such that $t_k = t'_{k'}$ or $t_k = s'_{k'}$, then $\mathbb{E}\left[\prod_{k=1}^{p} X_{t_k,l_1} X_{s_k,l_2} X_{t'_k,l_3} X_{s'_k,l_4}\right] \neq 0$ only if

$$Ind(t_1, s_1, ..., t_q, s_q, t'_1, s'_1, ..., t'_q, s'_q) = 1.$$

It is also easy to see that

$$\sum_{1 \leq i_1 < \cdots < i_q < j_1 < \cdots < j_q \leq n} \sum_{1 \leq i'_1 < \cdots < i'_q < j'_1 < \cdots < j'_q \leq n} Ind(t_1, s_1, ..., t_q, s_q, t'_1, s'_1, ..., t'_q, s'_q) = o(n^{4q}).$$

Combining all the results above, we have

$$\begin{aligned}
&\mathbb{E}[J(t_1, s_1, ..., t_q, s_q)^2] \\
=&\frac{1}{\binom{n}{2q}^2 \|\Sigma\|_q^{2q}} \sum_{1 \leq i_1 < \cdots < i_q < j_1 < \cdots < j_q \leq n} \sum_{1 \leq i'_1 < \cdots < i'_q < j'_1 < \cdots < j'_q \leq n} \sum_{l_1, l_2, l_3, l_4 = 1}^{p} \mathbb{E}\left[\prod_{k=1}^{q} X_{t_k,l_1} X_{s_k,l_2} X_{t'_k,l_3} X_{s'_k,l_4}\right], \\
\leq&\frac{C}{\binom{n}{2q}^2 \|\Sigma\|_q^{2q}} \sum_{1 \leq i_1 < \cdots < i_q < j_1 < \cdots < j_q \leq n} \sum_{1 \leq i'_1 < \cdots < i'_q < j'_1 < \cdots < j'_q \leq n} \\
&\qquad \sum_{1 \leq l_1 \leq l_2 \leq l_3 \leq l_4 \leq p} Ind(t_1, s_1, ..., t_q, s_q, t'_1, s'_1, ..., t'_q, s'_q)(1 \vee (l_2 - l_1))^{-r}(1 \vee (l_4 - l_3))^{-r} \\
\leq&\frac{o(n^{4q})}{\binom{n}{2q}^2 \|\Sigma\|_q^{2q}} \left(\sum_{1 \leq l_1 \leq l_2 \leq p} (1 \vee (l_2 - l_1))^{-r}\right)^2 \lesssim \frac{p^2}{\|\Sigma\|_q^{2q}} o(1) = o(1) \to 0,
\end{aligned}$$

where the last equality is because $\|\Sigma\|_q^{2q} = (\sum_{i,j=1}^{p} \Sigma_{i,j}^q)^2 \gtrsim p^2$.

This completes the proof of (2), as well as the whole proof.

$\square$

Table 1 shows additional simulation results for the size of the proposed monitoring statistics for $n = 200$. The size distortion problem has improved for almost all settings.

Table 1: Size of different monitoring procedures

$(n, p) = (200, 200)$	T1			T2			T3		
size $\alpha = 0.1$	L2	L6	Comb	L2	L6	Comb	L2	L6	Comb
$\rho = 0.2$	0.104	0.072	0.074	0.097	0.072	0.073	0.102	0.071	0.073
$\rho = 0.5$	0.105	0.064	0.091	0.107	0.064	0.085	0.104	0.065	0.087
$\rho = 0.8$	0.127	0.037	0.089	0.133	0.038	0.099	0.131	0.039	0.099

Bibliography

Wang, R., S. Volgushev, and X. Shao (2019). Inference for change points in high dimensional data. *arXiv preprint arXiv:1905.08446*.