Statistica Sinica: Supplement

A Self-Normalized Approach to Sequential Change-point Detection for Time Series

Ngai Hang Chan[1], Wai Leong Ng[2] and Chun Yip Yau[1]

The Chinese University of Hong Kong[1] *and Hang Seng University of Hong Kong*[2]

Supplementary Material

This supplementary material contains figures for time series plots of realizations of the models in Section 4 and the proofs for the main results in the paper.

S1 Figures for Section 4

This section shows figures for time series plots of realizations of the models in Section 4.

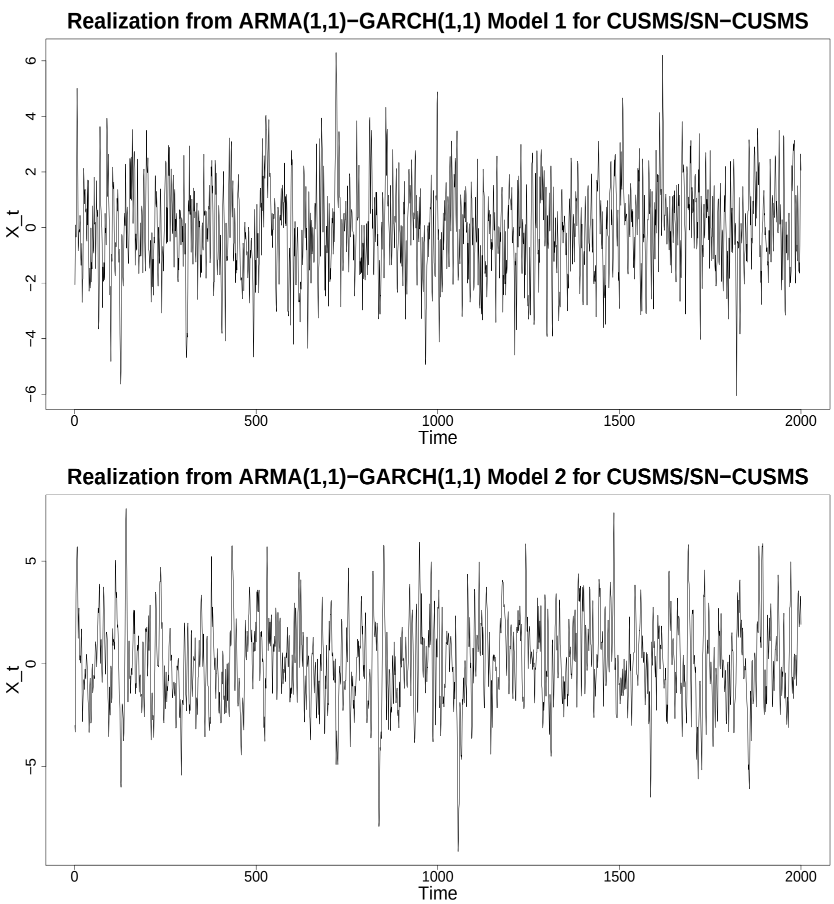

Figure S.1: Realizations for Models 1 and 2 for CUSMS/SN-CUSMS in Section 4.2 with $m = 1000$ and $T = 1$ without any change-point.

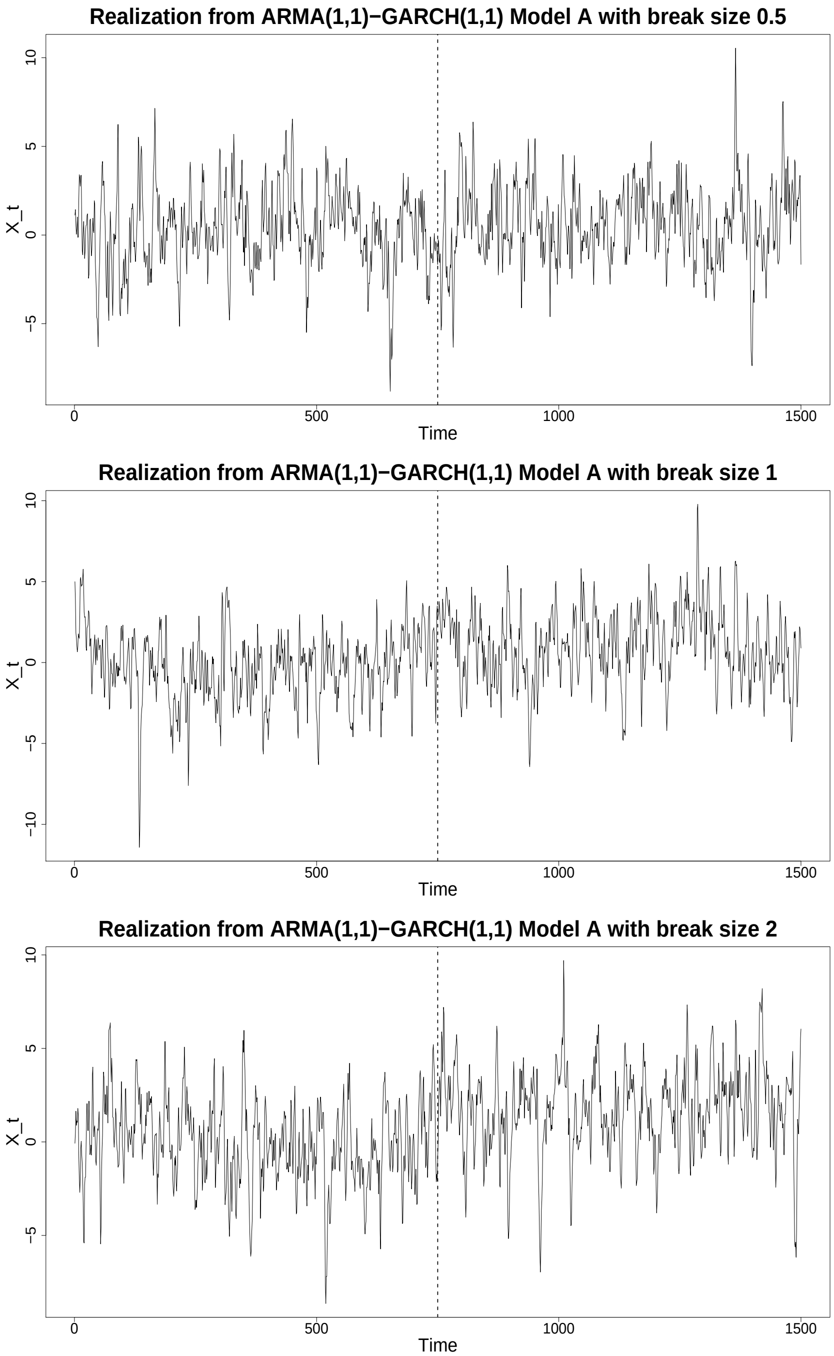

Figure S.2: Realizations for Model $\mathcal{A}$ in Section 4.3 with break size $\Delta = 0.5$, 1 and 2, $m = 500$, $T = 2$, $k^* = 250$. The change-points are represented by the vertical dash lines.

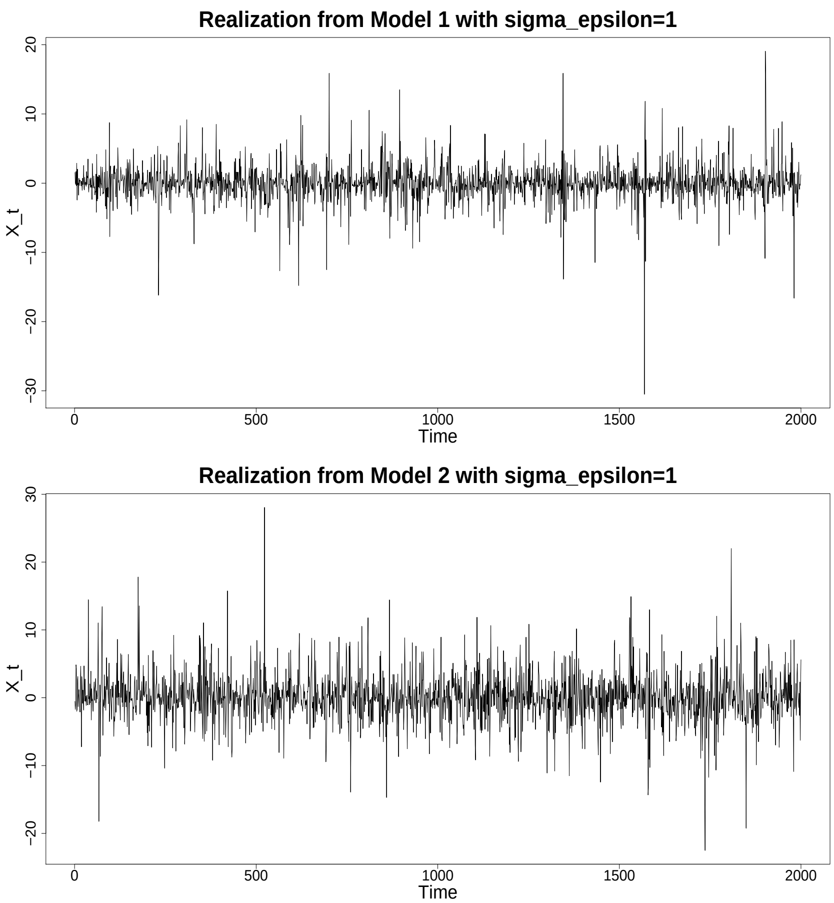

Figure S.3: Realizations of Models 1 and 2 with $\sigma_\epsilon = 1$ in Section 4.5 with $m = 1000$ and $T = 1$ without any change-point.

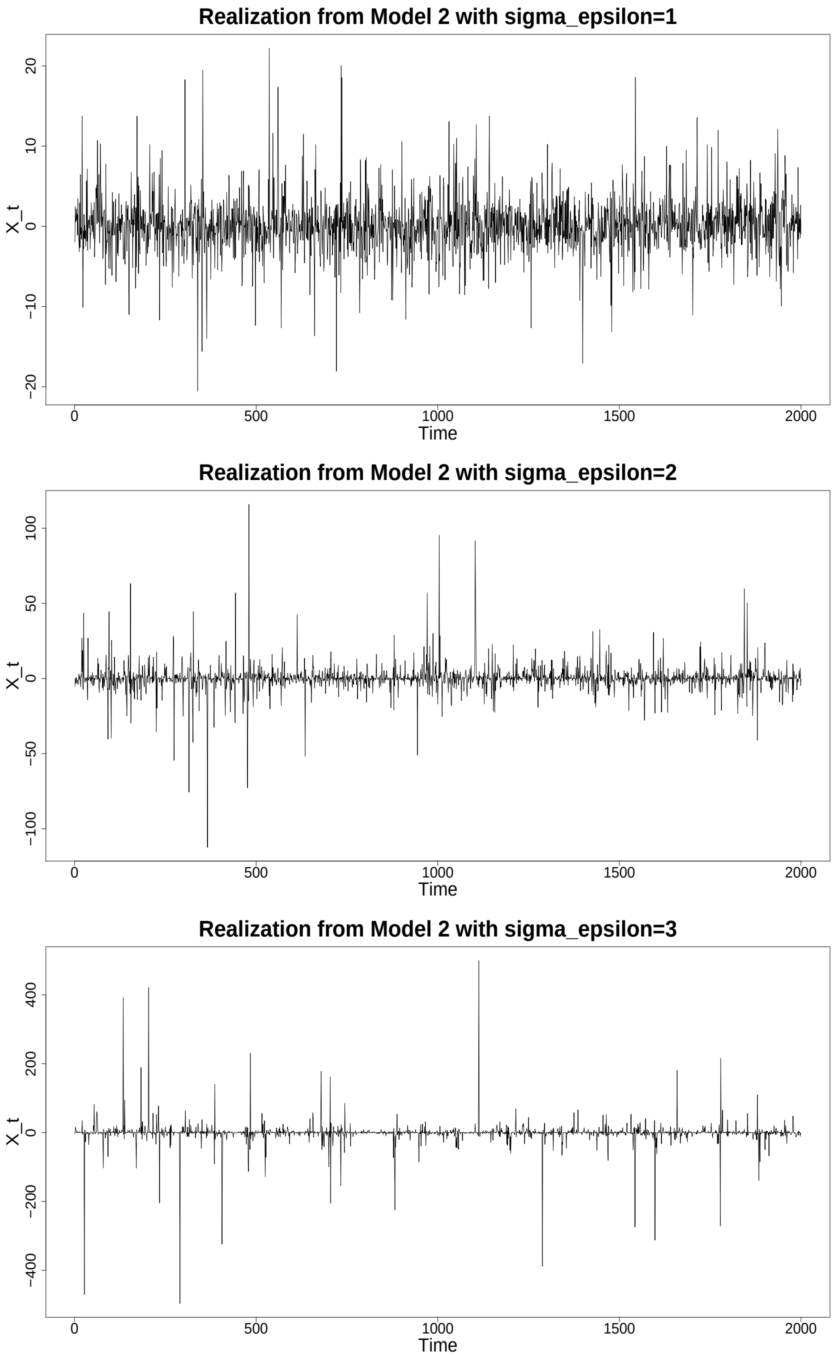

Figure S.4: Realizations of Model 2 with $\sigma_\epsilon = 1$, 2 and 3 in Section 4.5 with $m = 1000$ and $T = 1$ without any change-point.

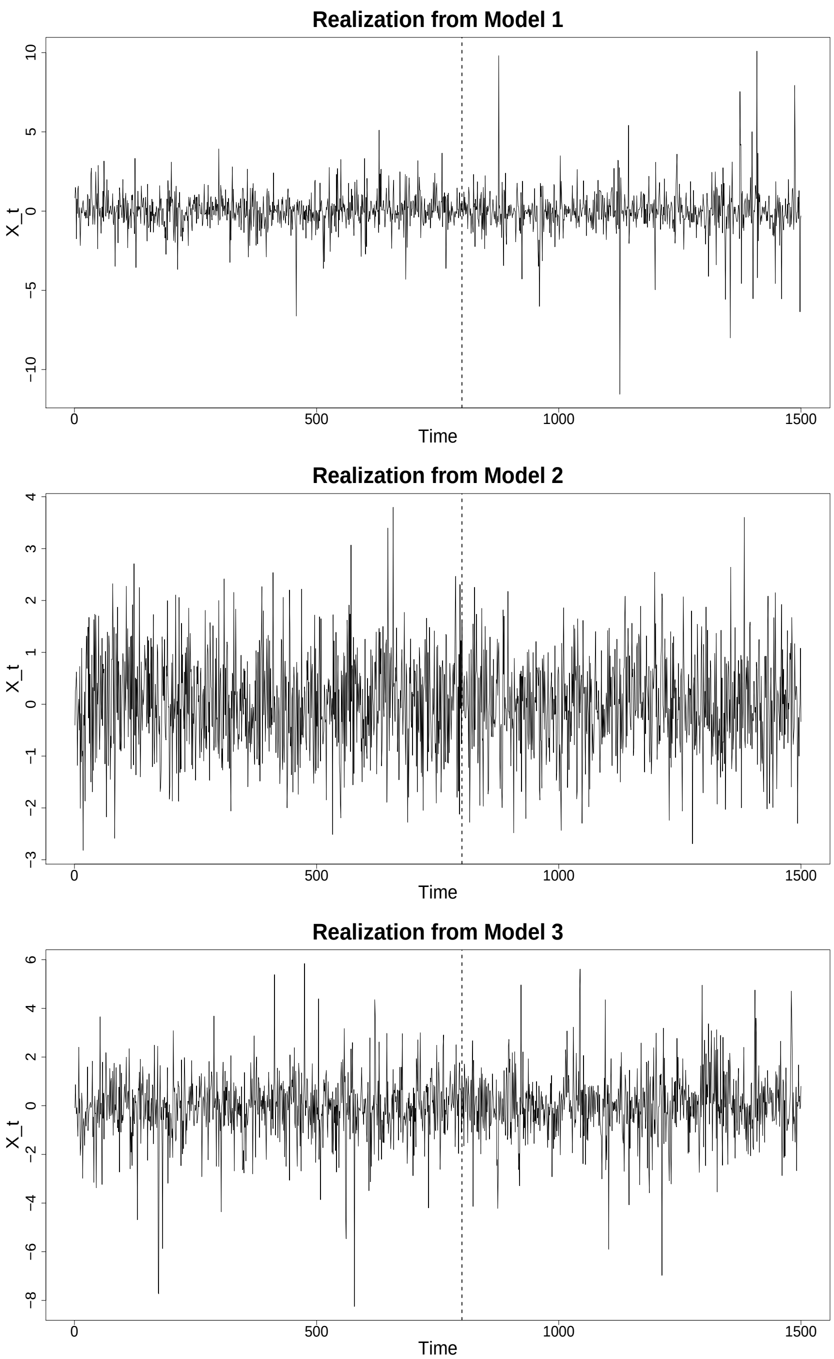

Figure S.5: Realizations for Models 1, 2 and 3 in Section 4.6 with $m = 750$, $T = 1$ and $k^* = 50$. The change-points are represented by the vertical dash lines.

S2 Proofs

S2.1 Proof in Section 3.3

For simplicity, denote $L_j(\boldsymbol{\theta}) = L(\mathbf{X}_j, \boldsymbol{\theta})$ and $L_j^*(\boldsymbol{\theta}) = L(\mathbf{X}_j^*, \boldsymbol{\theta})$. Also, let $\|\cdot\|$ be the maximum norm, i.e., for $x = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$, $\|x\| = \max_{i\in\{1,2,\ldots,d\}} |x_i|$.

Proof of Theorem 1(a). By Lemma 1 and the continuous mapping theorem, it suffices to show that

$$\sup_{1\le s\le m} \left\| \frac{\sum_{j=1}^{s} L_j(\hat{\boldsymbol{\theta}}_m) - \left(\sum_{j=1}^{s} L_j(\boldsymbol{\theta}_0) - \frac{s}{m}\sum_{j=1}^{m} L_j(\boldsymbol{\theta}_0)\right)}{m^{\frac{1}{2}}} \right\| \xrightarrow{p} 0\,. \quad \text{(S2.1)}$$

In the following, we consider

$$\sum_{j=1}^{s} L_j(\hat{\boldsymbol{\theta}}_m) = \left(\sum_{j=1}^{s} L_{j1}(\hat{\boldsymbol{\theta}}_m), \sum_{j=1}^{s} L_{j2}(\hat{\boldsymbol{\theta}}_m), \ldots, \sum_{j=1}^{s} L_{jd}(\hat{\boldsymbol{\theta}}_m)\right)$$

and using mean value theorem coordinate-wise, for each $i = 1, 2, \ldots, d$ and for all $s = 1, 2, \ldots, m$, we have

$$\left| \frac{\sum_{j=1}^{s} L_{ji}(\hat{\boldsymbol{\theta}}_m) - \left[\sum_{j=1}^{s} L_{ji}(\boldsymbol{\theta}_0) + \sum_{j=1}^{s} L'_{ji}(\boldsymbol{\theta}^*_{msi})(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_0)\right]}{m^{\frac{1}{2}}} \right| = 0\,, \quad \text{(S2.2)}$$

where $\boldsymbol{\theta}^*_{msi}$ is between $\hat{\boldsymbol{\theta}}_m$ and $\boldsymbol{\theta}_0$.

Also, by Assumption $\mathcal{A}$.5 and Lemma 1(a) and the uniform law of large numbers, we have for all $i = 1, 2, \ldots, d$,

$$\begin{aligned}
&\sup_{1\le s\le m}\left|\frac{\left[\sum_{j=1}^{s} L'_{ji}(\boldsymbol{\theta}^*_{msi}) - s\mathbb{E}(L'_{ji}(\boldsymbol{\theta}_0))\right](\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_0)}{m^{\frac{1}{2}}}\right| \\
\le\;& \sup_{1\le s\le m}\left|\left[\frac{\sum_{j=1}^{s} L'_{ji}(\boldsymbol{\theta}^*_{msi}) - s\mathbb{E}(L'_{ji}(\boldsymbol{\theta}_0))}{m}\right]\sqrt{m}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_0)\right| \\
=\;& |\sqrt{m}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_0)| \sup_{1\le s\le m}\left|\frac{\sum_{j=1}^{s} L'_{ji}(\boldsymbol{\theta}^*_{msi}) - s\mathbb{E}(L'_{ji}(\boldsymbol{\theta}_0))}{m}\right| \\
=\;& |O_p(1)| \sup_{1\le s\le m}\left|\frac{\sum_{j=1}^{s} L'_{ji}(\boldsymbol{\theta}^*_{msi}) - s\mathbb{E}(L'_{ji}(\boldsymbol{\theta}_0))}{m}\right| \xrightarrow{p} 0\,, \qquad \text{(S2.3)}
\end{aligned}$$

Using mean value theorem coordinate-wise, for each $i = 1, 2, \ldots, d$ and particularly for $s = m$, we have

$$\left|\frac{\sum_{j=1}^{m} L_{ji}(\hat{\boldsymbol{\theta}}_m) - \left[\sum_{j=1}^{m} L_{ji}(\boldsymbol{\theta}_0) + \sum_{j=1}^{m} L'_{ji}(\boldsymbol{\theta}^*_{mmi})(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_0)\right]}{m}\right| = 0\,,$$

where $\boldsymbol{\theta}^*_{mmi}$ is between $\hat{\boldsymbol{\theta}}_m$ and $\boldsymbol{\theta}_0$. Note that $\sum_{j=1}^{m} L_{ji}(\hat{\boldsymbol{\theta}}_m) = 0$ by definition. By the uniform law of large numbers and the positive definiteness of $\mathbb{E}(L'_j(\boldsymbol{\theta}_0))$, solving the system of linear equations yields

$$\mathbb{E}(L'_j(\boldsymbol{\theta}_0))(1 + o_p(1))(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_0) = -\frac{1}{m}\sum_{j=1}^{m} L_j(\boldsymbol{\theta}_0)\,.$$

Hence, we have

$$\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_0 = \left[-\frac{1}{m}\mathbb{E}(L'_j(\boldsymbol{\theta}_0))^{-1}\sum_{j=1}^{m} L_j(\boldsymbol{\theta}_0)\right](1 + o_p(1))\,. \qquad \text{(S2.4)}$$

Combining (S2.2), (S2.3) and (S2.4), we have

$$\sup_{1\le s\le m}\left\|\frac{\sum_{j=1}^{s} L_j(\hat{\boldsymbol{\theta}}_m) - \left(\sum_{j=1}^{s} L_j(\boldsymbol{\theta}_0) - \frac{s}{m}\sum_{j=1}^{m} L_j(\boldsymbol{\theta}_0)\right) + o_p(1)\left(\frac{s}{m}\sum_{j=1}^{m} L_j(\boldsymbol{\theta}_0)\right)}{m^{\frac{1}{2}}}\right\| \xrightarrow{p} 0\,. \qquad \text{(S2.5)}$$

Since

$$\sup_{1\le s\le m}\left\|\frac{\frac{s}{m}\sum_{j=1}^{m}L_j(\boldsymbol{\theta}_0)}{m^{\frac{1}{2}}}\right\| = \sup_{1\le s\le m}\left|\frac{s}{m}\right|\left\|\frac{\sum_{j=1}^{m}L_j(\boldsymbol{\theta}_0)}{m^{\frac{1}{2}}}\right\| = O_p(1)\,,$$

together with (S2.5), we have (S2.1).

By Lemma 1(b), we have for any $r \in [0,1]$ that,

$$\frac{\sum_{j=1}^{\lfloor mr\rfloor}L_j(\hat{\boldsymbol{\theta}}_m)}{\sqrt{m}} \xrightarrow{\mathcal{D}[0,1]} \mathbf{M}(\boldsymbol{\theta}_0)^{\frac{1}{2}}\left[\mathbb{B}_d(r) - r\mathbb{B}_d(1)\right]\,,$$

and thus the results follow from the continuous mapping theorem.

□

Proof of Theorem 1(b). For $T < \infty$, similar to the proof of Theorem 1(a), using mean value theorem on each coordinate i, i.e., for each $i = 1, 2, \ldots, d$ and for all $k = 1, 2, \ldots, mT$, we have

$$\left|\frac{\sum_{j=m+1}^{m+k}L_{ji}(\hat{\boldsymbol{\theta}}_m) - \left[\sum_{j=m+1}^{m+k}L_{ji}(\boldsymbol{\theta}_0) + \sum_{j=m+1}^{m+k}L'_{ji}(\boldsymbol{\theta}^*_{mki})(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_0)\right]}{m^{\frac{1}{2}}\left(1+\frac{k}{m}\right)}\right| = 0\,, \tag{S2.6}$$

where $\boldsymbol{\theta}^*_{mki}$ is between $\hat{\boldsymbol{\theta}}_m$ and $\boldsymbol{\theta}_0$.

Also, by Assumption $\mathcal{A}.5$, Lemma 1(a) and the uniform law of large numbers, we have that, for all $i = 1, 2, \dots, d$,

$$\begin{aligned}
&\sup_{1\le k\le mT}\left|\frac{\left[\sum_{j=m+1}^{m+k} L'_{ji}(\boldsymbol{\theta}^*_{mki}) - k\mathbb{E}(L'_{ji}(\boldsymbol{\theta}_0))\right](\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_0)}{m^{\frac{1}{2}}\left(1+\frac{k}{m}\right)}\right| \\
\le\ & \sup_{1\le k\le mT}\left|\left[\frac{\sum_{j=m+1}^{m+k} L'_{ji}(\boldsymbol{\theta}^*_{mki}) - k\mathbb{E}(L'_{ji}(\boldsymbol{\theta}_0))}{m+k}\right]\sqrt{m}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_0)\right| \\
=\ & |\sqrt{m}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_0)| \sup_{1\le k\le mT}\left|\frac{\sum_{j=m+1}^{m+k} L'_{ji}(\boldsymbol{\theta}^*_{mki}) - k\mathbb{E}(L'_{ji}(\boldsymbol{\theta}_0))}{m+k}\right| \\
=\ & |O_p(1)| \sup_{1\le k\le mT}\left|\frac{\sum_{j=m+1}^{m+k} L'_{ji}(\boldsymbol{\theta}^*_{mki}) - k\mathbb{E}(L'_{ji}(\boldsymbol{\theta}_0))}{m+k}\right| \xrightarrow{p} 0\,. \qquad \text{(S2.7)}
\end{aligned}$$

Combining (S2.4), (S2.6) and (S2.7), we have

$$\sup_{1\le k\le mT}\left\|\frac{\sum_{j=m+1}^{m+k} L_j(\hat{\boldsymbol{\theta}}_m) - \left(\sum_{j=m+1}^{m+k} L_j(\boldsymbol{\theta}_0) - \frac{k}{m}\sum_{j=1}^{m} L_j(\boldsymbol{\theta}_0)\right) + o_p(1)\left(\frac{k}{m}\sum_{j=1}^{m} L_j(\boldsymbol{\theta}_0)\right)}{m^{\frac{1}{2}}\left(1+\frac{k}{m}\right)}\right\| \xrightarrow{p} 0\,.$$

Since

$$\sup_{1\le k\le mT}\left\|\frac{\frac{k}{m}\sum_{j=1}^{m} L_j(\boldsymbol{\theta}_0)}{m^{\frac{1}{2}}\left(1+\frac{k}{m}\right)}\right\| = \sup_{1\le k\le mT}\left|\frac{\frac{k}{m}}{1+\frac{k}{m}}\right|\left\|\frac{\sum_{j=1}^{m} L_j(\boldsymbol{\theta}_0)}{m^{\frac{1}{2}}}\right\| = O_p(1)\,,$$

we have

$$\sup_{1\le k\le mT}\left\|\frac{\sum_{j=m+1}^{m+k} L_j(\hat{\boldsymbol{\theta}}_m) - \left(\sum_{j=m+1}^{m+k} L_j(\boldsymbol{\theta}_0) - \frac{k}{m}\sum_{j=1}^{m} L_j(\boldsymbol{\theta}_0)\right)}{m^{\frac{1}{2}}\left(1+\frac{k}{m}\right)}\right\| \xrightarrow{p} 0\,.$$

By Lemma 1(b), we have that for any $s \in [0, T]$,

$$\frac{S_m(\lfloor ms\rfloor, \hat{\boldsymbol{\theta}}_m)}{m^{\frac{1}{2}}(1+\frac{\lfloor ms\rfloor}{m})} = \frac{\sum_{j=m+1}^{m+\lfloor ms\rfloor} L_j(\hat{\boldsymbol{\theta}}_m)}{m^{\frac{1}{2}}(1+\frac{\lfloor ms\rfloor}{m})} \xrightarrow{\mathcal{D}[0,T]} \frac{\mathbf{M}(\boldsymbol{\theta}_0)^{\frac{1}{2}}\left[\mathbb{B}_d(1+s) - (1+s)\mathbb{B}_d(1)\right]}{1+s}\,. \qquad \text{(S2.8)}$$

Note that $\{\mathbb{B}_d(1+s) - (1+s)\mathbb{B}_d(1)\}_{s\in[0,T]}$ is independent of $\{\mathbb{B}_d(r) -$

$r\mathbb{B}_d(1)\}_{r\in[0,1]}$. Hence, by Theorem 1(a) and (S2.8),

$$\begin{aligned}&\sup_{1\le k\le mT}\frac{S_m(k,\hat{\boldsymbol{\theta}}_m)'D_m(\hat{\boldsymbol{\theta}}_m)^{-1}S_m(k,\hat{\boldsymbol{\theta}}_m)}{m\left(1+\frac{k}{m}\right)^2}\\ \xrightarrow{d}\quad&\sup_{0\le s\le T}\frac{\left[\mathbb{B}_d(1+s)-(1+s)\mathbb{B}_d(1)\right]'\mathbf{V}^{-1}\left[\mathbb{B}_d(1+s)-(1+s)\mathbb{B}_d(1)\right]}{(1+s)^2}\,.\end{aligned}$$

Since

$$P(T_m\le mT|H_0)=P\left(\sup_{1\le k\le mT}\frac{S_m(k,\hat{\boldsymbol{\theta}}_m)'D_m(\hat{\boldsymbol{\theta}}_m)^{-1}S_m(k,\hat{\boldsymbol{\theta}}_m)}{m\left(1+\frac{k}{m}\right)^2}>c\right),$$

taking limit on both sides yields (3.1). □

Proof of Theorem 1(c). For $T=\infty$, by similar arguments as above, we can show that

$$\sup_{1\le k<\infty}\left\|\frac{\sum_{j=m+1}^{m+k}L_j(\hat{\boldsymbol{\theta}}_m)-\left(\sum_{j=m+1}^{m+k}L_j(\boldsymbol{\theta}_0)-\frac{k}{m}\sum_{j=1}^{m}L_j(\boldsymbol{\theta}_0)\right)}{m^{\frac{1}{2}}\left(1+\frac{k}{m}\right)}\right\|\xrightarrow{p}0\,.$$

Thus, it suffices to show that

$$\sup_{1\le k<\infty}\left\|\frac{\sum_{j=m+1}^{m+k}L_j(\boldsymbol{\theta}_0)-\frac{k}{m}\sum_{j=1}^{m}L_j(\boldsymbol{\theta}_0)}{m^{\frac{1}{2}}\left(1+\frac{k}{m}\right)}\right\|\xrightarrow{d}\sup_{0\le s<\infty}\left\|\frac{\mathbf{M}(\boldsymbol{\theta}_0)^{\frac{1}{2}}\left[\mathbb{B}_d(1+s)-(1+s)\mathbb{B}_d(1)\right]}{(1+s)}\right\|.$$

Hence, it in turn suffices to show that

$$\sup_{mT\le k<\infty}\left\|\frac{\sum_{j=m+1}^{m+k}L_j(\boldsymbol{\theta}_0)}{m^{\frac{1}{2}}\left(1+\frac{k}{m}\right)}\right\|\xrightarrow{p}0\,,\qquad\text{(S2.9)}$$

and

$$\sup_{T\le s<\infty}\left\|\frac{\mathbf{M}(\boldsymbol{\theta}_0)^{\frac{1}{2}}\mathbb{B}_d(1+s)}{(1+s)}\right\|\xrightarrow{p}0\,.\qquad\text{(S2.10)}$$

For (S2.9), by the additional ρ-mixing conditions, we have the ρ-mixing Hájek-Rényi inequality, see Theorem 1 of Wan (2013). Specifically, for some constant c^*, for each $i = 1, 2, \ldots, d$ and any $\epsilon > 0$, we have

$$P\left(\max_{mT\le k\le n}\left|\frac{\sum_{j=m+1}^{m+k} L_{ji}(\boldsymbol{\theta}_0)}{m^{\frac{1}{2}}(1+\frac{k}{m})}\right| \ge \epsilon\right) \le \frac{c^*}{\epsilon^2}\left(\sum_{j=m+1}^{m+mT}\frac{Var(L_{ji}(\boldsymbol{\theta}_0))}{[m^{\frac{1}{2}}(1+T)]^2}\right)+4\sum_{j=m+mT+1}^{n}\frac{Var(L_{ji}(\boldsymbol{\theta}_0))}{[m^{\frac{1}{2}}(1+\frac{j}{m})]^2}\,. \tag{S2.11}$$

Next, taking $\lim_{T\to\infty}\limsup_{m\to\infty}\lim_{n\to\infty}$ on both sides of (S2.11), we have for $\epsilon > 0$ that,

$$\lim_{T\to\infty}\limsup_{m\to\infty}\lim_{n\to\infty} P\left(\sup_{mT\le k\le n}\left\|\frac{\sum_{j=m+1}^{m+k} L_{j}(\boldsymbol{\theta}_0)}{m^{\frac{1}{2}}(1+\frac{k}{m})}\right\| \ge \epsilon\right) = 0\,.$$

which yields (S2.9).

For (S2.10), by the law of iterated logarithm, we also have

$$\sup_{T\le s<\infty}\left\|\frac{\mathbf{M}(\boldsymbol{\theta}_0)^{\frac{1}{2}}\mathbb{B}_d(1+s)}{(1+s)}\right\| \xrightarrow{a.s.} 0\,,$$

as $T\to\infty$. Thus, we have (S2.10) and hence we have

$$\begin{aligned}&\lim_{m\to\infty} P(T_m<\infty|H_0)\\ =\;& P\left(\sup_{0\le s<\infty}\frac{[\mathbb{B}_d(1+s)-(1+s)\mathbb{B}_d(1)]'\mathbf{V}^{-1}[\mathbb{B}_d(1+s)-(1+s)\mathbb{B}_d(1)]}{(1+s)^2} > c\right)\,.\end{aligned}$$

Note that $\mathbf{V}$ is a functional of $\{\mathbb{B}_d(r)\}_{r\in[0,1)}$. Hence, by the independent increment of the standard Brownian motion, $\{\mathbb{B}_d(1+s)-(1+s)\mathbb{B}_d(1)\}_{s\in[0,\infty)}$ is independent of $\mathbf{V}$. By the proof of Theorem 1 in Hušková and Koubková (2005), we have

$$\{\mathbb{B}_d(1+s)-(1+s)\mathbb{B}_d(1)\} \stackrel{d}{=} \left\{(1+s)\mathbb{B}_d^*\left(\frac{s}{1+s}\right)\right\}\,,$$

where $\mathbb{B}_d^*(\cdot)$ is independent of $\mathbf{V}$. Hence, we have

$$\begin{aligned} & P\left(\sup_{0\le s<\infty} \frac{[\mathbb{B}_d(1+s)-(1+s)\mathbb{B}_d(1)]'\mathbf{V}^{-1}[\mathbb{B}_d(1+s)-(1+s)\mathbb{B}_d(1)]}{(1+s)^2} > c\right) \\ = & P\left(\sup_{0\le s<\infty} \frac{[(1+s)\mathbb{B}_d^*(\frac{s}{1+s})]'\mathbf{V}^{-1}[(1+s)\mathbb{B}_d^*(\frac{s}{1+s})]}{(1+s)^2} > c\right) \\ = & P\left(\sup_{0\le s<\infty} \mathbb{B}_d^*\left(\frac{s}{1+s}\right)'\mathbf{V}^{-1}\mathbb{B}_d^*\left(\frac{s}{1+s}\right) > c\right) \\ = & P\left(\sup_{0\le u<1} \mathbb{B}_d^*(u)'\mathbf{V}^{-1}\mathbb{B}_d^*(u) > c\right), \end{aligned}$$

and the results of Theorem 1(c) follow.

□

S2.2 Proof in Section 3.4

Proof of Theorem 2. By Assumptions $\mathcal{B}.1$, $\mathcal{B}.2$ and the uniform law of large numbers, we have

$$\frac{1}{m+mT-t^*+1}\sum_{t=t^*}^{m+mT} L_j^*(\hat{\boldsymbol{\theta}}_m) = \mathbb{E}(L_j^*(\boldsymbol{\theta}_0)) + o_p(1)\,.$$

Hence, for coordinates $i = 1, 2, \ldots, d$ in which $\mathbb{E}(L_{ji}^*(\boldsymbol{\theta}_0)) \neq 0$, we have

$$\begin{aligned} \frac{S_m(mT,\hat{\boldsymbol{\theta}}_m)}{m^{\frac{1}{2}}(1+\frac{mT}{m})} &= m^{-\frac{1}{2}}(1+T)^{-1}\left(\sum_{t=m+1}^{t^*-1} L_j(\hat{\boldsymbol{\theta}}_m) + \sum_{t=t^*}^{m+mT} L_j^*(\hat{\boldsymbol{\theta}}_m)\right) \\ &= O_p(1) + \frac{m+mT-t^*+1}{m^{\frac{1}{2}}(1+T)}\big(\mathbb{E}(L_j^*(\boldsymbol{\theta}_0)) + o_p(1)\big) \\ &= O_p(\sqrt{m})\,. \end{aligned}$$

Also, we have

$$\begin{aligned}
\sup_{1\le k\le mT} \frac{S_m(k,\hat{\boldsymbol{\theta}}_m)' D_m(\hat{\boldsymbol{\theta}}_m)^{-1} S_m(k,\hat{\boldsymbol{\theta}}_m)}{m\left(1+\frac{k}{m}\right)^2} &\ge \left(\frac{S_m(mT,\hat{\boldsymbol{\theta}}_m)}{m^{\frac{1}{2}}(1+\frac{mT}{m})}\right)' D_m(\hat{\boldsymbol{\theta}}_m)^{-1} \left(\frac{S_m(mT,\hat{\boldsymbol{\theta}}_m)}{m^{\frac{1}{2}}(1+\frac{mT}{m})}\right) \\
&\to \infty\,,
\end{aligned}$$

as $m \to \infty$. As a result,

$$\begin{aligned}
\lim_{m\to\infty} P(T_m \le mT|H_1) &= \lim_{m\to\infty} P\left(\sup_{1\le k\le mT} \frac{S_m(k,\hat{\boldsymbol{\theta}}_m)' D_m(\hat{\boldsymbol{\theta}}_m)^{-1} S_m(k,\hat{\boldsymbol{\theta}}_m)}{m\left(1+\frac{k}{m}\right)^2} > c \middle| H_1\right) \\
&\to 1\,.
\end{aligned}$$

Similar arguments can be applied for the case of open-end procedure. Hence, the proof is complete. □