Supplementary Material for: Estimating and testing nonlinear local dependence between two time series

September 10, 2017

Virginia Lacal

Dag Tjøstheim

Department of Mathematics

University of Bergen

Post box 7802

5020 Bergen Norway

A Appendix

First, note that

$$J_b(x,y) = \int K_b\left(v-x, w-y\right) u\left((v,w), \theta_{0,b}(x,y)\right) \cdot \tag{A.1}$$
$$\cdot\, u^t\left((v,w), \theta_{0,b}(x,y)\right) \psi\left((v,w), \theta_{0,b}(x,y)\right) dvdw$$
$$- \int K_b\left(v-x, w-y\right) \nabla u\left((v,w), \theta_{0,b}(x,y)\right) \cdot$$
$$\cdot \left[f(v,w) - \psi\left((v,w), \theta_{0,b}(x,y)\right)\right] dvdw \ ,$$

where $u((x,y),\theta) = \nabla \log(\psi((x,y),\theta))$. We follow the idea in Hjort and Jones (1995) and Tjøstheim and Hufthammer (2013). The matrix $J_b(x,y)$ can be written as:

$$J_b(x,y) = \int K(r)K(s) u\left((x+b_1r, y+b_2s), \theta_{0,b}(x,y)\right) u^t\left((x+b_1r, y+b_2s), \theta_{0,b}(x,y)\right) \cdot$$
$$\cdot\, \psi\left((x+b_1r, y+b_2s), \theta_{0,b}(x,y)\right) drds$$
$$- \int K(r)K(s) \nabla u\left((x+b_1r, y+b_2s), \theta_{0,b}(x,y)\right) \cdot$$
$$\cdot \left[f(x+b_1r, y+b_2s) - \psi\left((x+b_1r, y+b_2s), \theta_{0,b}(x,y)\right)\right] drds,$$

where we set $r = \frac{v-x}{b_1}$ and $s = \frac{w-y}{b_2}$. Taylor expanding u, it follows that

$$J_b(x,y) \sim \int K(r)K(s) D(x,y) w_b(r,s) w_b^t(r,s) D^t(x,y) \cdot$$
$$\cdot\, \psi\left((x+b_1r, y+b_2s), \theta_{0,b}(x,y)\right) drds$$
$$- \int K(r)K(s) \nabla u\left((x+b_1r, y+b_2s), \theta_{0,b}(x,y)\right) \cdot$$
$$\cdot \left[f(x+b_1r, y+b_2s) - \psi\left((x+b_1r, y+b_2s), \theta_{0,b}(x,y)\right)\right] drds,$$

where $w_b(r,s)$ is the 6-dimensional vector defined by $w_b^t(r,s) = (1, b_1r, b_2s, b_1^2r^2, b_1b_2rs, b_2^2s^2)$ and $D(x,y)$ is the 5×6-matrix $D(x,y) = \left(u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{1}{2}\frac{\partial^2 u}{\partial x^2}, \frac{\partial^2 u}{\partial x \partial y}, \frac{1}{2}\frac{\partial^2 u}{\partial y^2}\right)$. Taylor expanding

ψ, ∇u and f only at the first order, we have that

$$
\begin{aligned}
J_b(x,y) \sim\ & D(x,y)\psi\left((x,y),\theta_{0,b}(x,y)\right)\left[\int K(r)K(s)w_b(r,s)w_b^t(r,s)drds\right]D^t(x,y) \\
& -\nabla u\left((x,y),\theta_{0,b}(x,y)\right)\left[f(x,y)-\psi\left((x,y),\theta_{0,b}(x,y)\right)\right]\int K(r)K(s)drds.
\end{aligned}
$$

Computing $\int K(r)K(s)w_b(r,s)w_b^t(r,s)drds$, we obtain the following matrix

$$
H_b = \begin{pmatrix}
1 & 0 & 0 & \alpha b_1^2 & 0 & \alpha b_2^2 \\
0 & \alpha b_1^2 & 0 & 0 & 0 & 0 \\
0 & 0 & \alpha b_2^2 & 0 & 0 & 0 \\
\alpha b_1^2 & 0 & 0 & \beta b_1^4 & 0 & \alpha^2 b_1^2 b_2^2 \\
0 & 0 & 0 & 0 & \alpha^2 b_1^2 b_2^2 & 0 \\
\alpha b_2^2 & 0 & 0 & \alpha^2 b_1^2 b_2^2 & 0 & \beta b_2^4
\end{pmatrix}
$$

where $\alpha = \int K(s)s^2 ds \geq 0$ and $\beta = \int K(s)s^4 ds \geq 0$. The matrix H_b is of rank 5, therefore, $D(x,y)H_bD^t(x,y)$ is of rank 5. Hence,

$$
\begin{aligned}
J_b(x,y) \sim\ & D(x,y)\psi\left((x,y),\theta_{0,b}(x,y)\right)H_bD^t(x,y) \\
& -\nabla u\left((x,y),\theta_{0,b}(x,y)\right)\left[f(x,y)-\psi\left((x,y),\theta_{0,b}(x,y)\right)\right] \doteq J_{b,2}(x,y). \qquad \text{(A.2)}
\end{aligned}
$$

It follows (cf. Tjøstheim and Hufthammer (2013)) that $(b_1b_2)^{-2}J_{b,2}(x,y)$ is non-singular and positive definite as $b_1, b_2 \to 0$. The same idea can be used to find an approximating expression for

$$
\begin{aligned}
M_b(x,y) = b_1b_2 & \int K_b^2(v-x,w-y)u((v,w),\theta_{0,b}(x,y)))\cdot \qquad \text{(A.3)} \\
& \cdot u^t((v,w),\theta_{0,b}(x,y)))f(v,w)dvdw \\
& - b_1b_2\left(\int K_b(v-x,w-y)u((v,w),\theta_{0,b}(x,y)))f(v,w)dvdw\right)\cdot \\
& \cdot\left(\int K_b(v-x,w-y)u((v,w),\theta_{0,b}(x,y)))f(v,w)dvdw\right)^t.
\end{aligned}
$$

B Appendix

In this appendix the proofs of the theorems in Section 2 are given.

Proof of Theorem 2.1. Define $M < \infty$ such that $|L^{(b)}((X_t, Y_t), \theta) - L^{(b)}((X_t, Y_t), \theta')| \leq C_t|\theta - \theta'|$ and $C_t \leq M$ and $\Delta < \infty$ such that $\|L^{(b)}((X_t, Y_t), \theta)\|_4 \leq \Delta$. Given $\delta > 0$, let $\{\eta(\theta_i, \delta) : i = 1, \dots, B\}$ be a finite sub-cover of Θ, where $\eta(\theta_i, \delta) = \{\theta \in \Theta : |\theta - \theta_i| < \delta\}$. Then

$$\sup_{\theta \in \Theta} |L_{n,b}(\theta) - \mu_b| = \max_i \sup_{\theta \in \eta(\theta_i, \delta)} |L_{n,b}(\theta) - \mu_b|. \tag{B.1}$$

Hence, we can write

$$\mathbb{P}\left(\sup_{\theta \in \Theta} |L_{n,b}(\theta) - \mu_b| > \epsilon\right) \leq \sum_{i=1}^{B} \mathbb{P}\left(\sup_{\theta \in \eta(\theta_i, \delta)} |L_{n,b}(\theta) - \mu_b| > \epsilon\right). \tag{B.2}$$

By the Lipschitz continuity of $\left\{L^{(b)}((X_t, Y_t), \theta)\right\}$, if $\theta \in \eta(\theta_i, \delta)$, then

$$|L_{n,b}(\theta) - \mu_b| \leq |L_{n,b}(\theta) - L_{n,b}(\theta_i)| + |L_{n,b}(\theta_i) - \mu_b| \leq \frac{\delta}{n} \sum_{t=1}^{n} C_t + |L_{n,b}(\theta_i) - \mu_b|.$$

By the Markov inequality,

$$\begin{aligned}
\mathbb{P}\left(\sup_{\theta \in \eta(\theta_i, \delta)} |L_{n,b}(\theta) - \mu_b| > \epsilon\right) &\leq \mathbb{P}\left(\frac{\delta}{n} \sum_{t=1}^{n} C_t > \frac{\epsilon}{2}\right) + \mathbb{P}\left(|L_{n,b}(\theta_i) - \mu_b| > \frac{\epsilon}{2}\right) \\
&\leq \frac{2\delta}{n\epsilon} \mathbb{E}\left(\sum_{t=1}^{n} C_t\right) + \frac{4}{\epsilon^2} \mathrm{Var}\left(L_{n,b}(\theta_i)\right) \\
&\leq \frac{2\delta M}{\epsilon} + \frac{4}{\epsilon^2} \mathrm{Var}\left(L_{n,b}(\theta_i)\right).
\end{aligned}$$

By the mixing inequality (Corollary A.2, Hall and Heyde (1980)) and the 4-dominance of $\left\{L^{(b)}((X_t, Y_t), \theta)\right\}$,

$$
\begin{aligned}
\operatorname{Var}\left(L_{n,b}(\theta_i)\right) &= \frac{1}{n^2}\operatorname{Var}\left(\sum_{t=1}^{n} L^{(b)}((X_t,Y_t),\theta_i)\right)\\
&= \frac{1}{n^2}\sum_{t=1}^{n}\sum_{s=1}^{n}\operatorname{Cov}\left(L^{(b)}((X_t,Y_t),\theta_i), L^{(b)}((X_s,Y_s),\theta_i)\right)\\
&\leq \frac{8}{n^2}\sum_{t=1}^{n}\sum_{s=1}^{n}\|L^{(b)}((X_t,Y_t),\theta_i)\|_4\|L^{(b)}((X_s,Y_s),\theta_i)\|_4\alpha^{\frac{|t-s|}{2}}\\
&\leq \frac{8\Delta^2}{n^2}\sum_{t=1}^{n}\sum_{s=1}^{n}\alpha^{\frac{|t-s|}{2}} = \frac{8\Delta^2}{n^2}\left(n+2\sum_{j=1}^{n-1}(n-j)\alpha^{\frac{j}{2}}\right)\\
&\leq \frac{8\Delta^2}{n^2}\left(n+2(n-1)\sum_{j=1}^{n-1}\alpha^{\frac{j}{2}}\right) = \frac{8\Delta^2}{n^2}\left(n+2(n-1)\frac{\alpha^{\frac{1}{2}}-\alpha^{\frac{n}{2}}}{1-\alpha^{\frac{1}{2}}}\right).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\mathbb{P}\left(\sup_{\theta\in\eta(\theta_i,\delta)}|L_{n,b}(\theta)-\mu_b|>\epsilon\right) &\leq \frac{2\delta M}{\epsilon}+\frac{32\Delta^2}{n^2}\left(n+2(n-1)\frac{\alpha^{\frac{1}{2}}-\alpha^{\frac{n}{2}}}{1-\alpha^{\frac{1}{2}}}\right)\\
&< \zeta+\frac{32\Delta^2}{n^2}\left(n+2(n-1)\frac{\alpha^{\frac{1}{2}}-\alpha^{\frac{n}{2}}}{1-\alpha^{\frac{1}{2}}}\right),
\end{aligned}
$$

$\forall n$ sufficiently large, $\forall\zeta>0$, $\delta<\frac{\epsilon\zeta}{2M}$. From this, it follows that, as $n\to\infty$,

$$
\mathbb{P}\left(\sup_{\theta\in\Theta}|L_{n,b}(\theta)-\mu_b|>\epsilon\right)<B\zeta,\ \forall\zeta>0,\ B<\infty,
$$

therefore,

$$
\lim_{n\to\infty}\mathbb{P}\left(\sup_{\theta\in\Theta}\mid L_{n,b}(\theta)-\mu_b\mid>\epsilon\right)=0,\ \forall\epsilon>0.
$$

This means that

$$
L_{n,b}(\theta)-\mathbb{E}\left(L_{n,b}(\theta)\right)\xrightarrow{\mathbb{P}}0 \text{ uniformly on } \Theta. \tag{B.3}
$$

Note that, from the stationarity of the process $\{(X_t,Y_t)\}$, it follows that $\mathbb{E}\left(L_{n,b}(\theta)\right)=\mathbb{E}(L^{(b)}((X_t,Y_t),\theta))$. From the fact that $L_{n,b}$ is continuous on Θ $a.s.-\mathbb{P}$ with maximizer $\theta_{n,b}$, assumption **A2)** and (B.3), it follows that $L_{n,b}(\theta_{0,b})-\mathbb{E}\left(L_{n,b}(\theta_{0,b})\right)\xrightarrow{\mathbb{P}}0$. Moreover, using (B.3) and the definition of $\theta_{n,b}$ and $\theta_{0,b}$ (that is, assumption **A2)**), we have that

$L_{n,b}(\theta_{n,b})-\mathbb{E}\left(L^{(b)}((X_t,Y_t),\theta_{0,b})\right) = \sup_{\theta\in\Theta} L_{n,b}(\theta)-\sup_{\theta\in\Theta}\mathbb{E}\left(L^{(b)}((X_t,Y_t),\theta)\right) \xrightarrow{\mathbb{P}} 0$. This means that $L_{n,b}(\theta_{n,b}) - L_{n,b}(\theta_{0,b}) \xrightarrow{\mathbb{P}} 0$. By the assumption that $\theta_{n,b}$ is a maximizer of $L_{n,b}$, for every $\epsilon > 0$ there exists a $\eta > 0$ such that $\mid L_{n,b}(\theta_{n,b}) - L_{n,b}(\theta) \mid > \eta$ for every θ with $\mid \theta_{n,b} - \theta \mid > \epsilon$. Therefore, if we take $\theta = \theta_{0,b}$, the event $\{\mid \theta_{n,b} - \theta_{0,b} \mid > \epsilon\}$ is contained in the event $\{\mid L_{n,b}(\theta_{n,b}) - L_{n,b}(\theta_{0,b}) \mid > \eta\}$, meaning that for every $\epsilon > 0$, there exists $\eta > 0$, such that

$$\mathbb{P}\left(\mid \theta_{n,b} - \theta_{0,b} \mid > \epsilon\right) \le \mathbb{P}\left(\mid L_{n,b}(\theta_{n,b}) - L_{n,b}(\theta_{0,b}) \mid > \eta\right) \xrightarrow{\mathbb{P}} 0.$$

The last statement of the theorem follows since $\theta_{0,b} \to \theta_0$ by definition. □

Proof of Theorem 2.2. It is a generalization of Theorem 3 of Tjøstheim and Hufthammer (2013). Define $Q_n(\theta) = -\frac{n}{(b_1b_2)^2}L_{n,b}(\theta)$ and consider the Taylor expansion of $\nabla Q_n(\theta)$

$$0 = \frac{1}{\sqrt{n}}\nabla Q_n(\theta_{n,b}) = \frac{1}{\sqrt{n}}\nabla Q_n(\theta_{0,b}) + \frac{1}{n}\nabla^2 Q_n(\tilde{\theta})\sqrt{n}(\theta_{n,b} - \theta_{0,b})$$

where $\tilde{\theta}$ is determined by the mean value theorem. Therefore,

$$\begin{aligned}
-\frac{(b_1b_2)^{\frac{3}{2}}}{\sqrt{n}}\nabla Q_n(\theta_{0,b}) &= \frac{(b_1b_2)^{\frac{3}{2}}}{n}\nabla^2 Q_n(\tilde{\theta})\sqrt{n}(\theta_{n,b} - \theta_{0,b}) \\
&= \frac{(b_1b_2)^{\frac{3}{2}}}{n}\left[\nabla^2 Q_n(\theta_{0,b}) + \left(\nabla^2 Q_n(\tilde{\theta}) - \nabla^2 Q_n(\theta_{0,b})\right)\right]\sqrt{n}(\theta_{n,b} - \theta_{0,b}).
\end{aligned}$$

If we can prove that

1. $\frac{1}{n}\nabla Q_n(\theta_{0,b}) \to 0$ a.s.;
2. $\frac{1}{n}\nabla^2 Q_n(\theta_{0,b}) \to \tilde{J}$ a.s., where $\tilde{J}$ is a 5×5 positive definite matrix that can be identified with the limit of $\frac{1}{(b_1b_2)^2}J_{n,b}$ as $n \to \infty$ and $b \to 0$;
3. $\lim_{n\to\infty}\limsup_{\delta\to 0}\frac{1}{n\delta}\mid \nabla^2 Q_n(\tilde{\theta}) - \nabla^2 Q_n(\theta_{0,b}) \mid < \infty$;
4. $\mathrm{Var}\left(-\frac{(b_1b_2)^{\frac{3}{2}}}{\sqrt{n}}\nabla Q_n(\theta_{0,b})\right) = \frac{1}{(b_1b_2)^2}M_{n,b}$;

then, using Theorem 4.4 of Masry and Tjøstheim (1995) and Theorem 2.2 of Klimko and Nelson (1978), we have the result. To prove point 1., we need to use Theorem 4.1 of Masry and Tjøstheim (1995) and assumption **A2)**. Indeed, with $u((x,y),\theta) = \nabla \log(\psi((x,y),\theta))$, we have that

$$\begin{aligned}\frac{1}{n}\nabla Q_n(\theta)\mid_{\theta=\theta_{0,b}} &= \frac{1}{(b_1b_2)^2}\int K_b(v-x,w-y)u((v,w),\theta_{0,b})\psi((v,w),\theta_{0,b})dvdw \\ &- \frac{1}{n(b_1b_2)^2}\sum_{t=1}^{n} K_b(X_t-x,Y_t-y)u((X_t,Y_t),\theta_{0,b}) \\ &= O\left(\frac{1}{(b_1b_2)^2}\left(\frac{\log n}{nb_1b_2}\right)^{\frac{1}{2}}\right) = O\left(\left(\frac{\log n}{n(b_1b_2)^5}\right)^{\frac{1}{2}}\right).\end{aligned}$$

and assumption 3, stated in the theorem, can be used. In the same way it is possible to prove point 2.,

$$\begin{aligned}\frac{1}{n}\nabla^2 Q_n(\theta)\mid_{\theta=\theta_{0,b}} &= \frac{1}{(b_1b_2)^2}\int K_b(v-x,w-y)u((v,w),\theta_{0,b})u^t((v,w),\theta_{0,b})\psi((v,w),\theta_{0,b})dvdw \\ &+ \frac{1}{(b_1b_2)^2}\int K_b(v-x,w-y)\nabla u((v,w),\theta_{0,b})\psi((v,w),\theta_{0,b})dvdw \\ &- \frac{1}{n(b_1b_2)^2}\sum_{t=1}^{n} K_b(X_t-x,Y_t-y)\nabla u((X_t,Y_t),\theta_{0,b}) \\ &= \frac{1}{(b_1b_2)^2}J_{b,2} - \frac{1}{n(b_1b_2)^2}\sum_{t=1}^{n} K_b(X_t-x,Y_t-y)\nabla u((X_t,Y_t),\theta_{0,b}) \\ &+ \frac{1}{(b_1b_2)^2}\int K_b(v-x,w-y)\nabla u((v,w),\theta_{0,b})f(v,w)dvdw \to \tilde{J}\ a.s.\ ,\end{aligned}$$

where $J_{b,2}$ is defined in Appendix A. Point 3. follows from the mean value theorem and the fact that the third derivative of $Q_n'(\theta) = (b_1b_2)^2 Q_n(\theta)$ can be bounded by a constant c,

$$\begin{aligned}\frac{1}{n\delta}\left[\nabla^2 Q_n(\tilde{\theta}) - \nabla^2 Q_n(\theta_{0,b})\right] &= \frac{1}{n\delta}\nabla^3 Q_n(\hat{\theta})(\theta_{0,b}-\tilde{\theta}) \\ &= \frac{1}{n\delta(b_1b_2)^2}\nabla^3 Q_n'(\hat{\theta})(\theta_{0,b}-\tilde{\theta}) \leq \frac{c}{n(b_1b_2)^2}\end{aligned}$$

where $\hat{\theta}$ is determined by the mean value theorem and $\mid \theta_{0,b} - \tilde{\theta} \mid < \delta$. Finally, point 4. is a straightforward consequence of the definition of $\nabla Q_n(\theta_{0,b})$. □

C Appendix

In this appendix the proof of the theorems in Section 3 are given.

Proof of Theorem 3.1. The proof is just an application of the continuous mapping theorem. □

Proof of Theorem 3.2. The proof of this theorem is essentially the same as for of Theorem 3.1 in Lacal and Tjøstheim (2016). The asymptotic normality for b fixed is proved as in that theorem. We only need to evaluate the variance of $\int A_b(x,y)dG_n(x,y)$, where $G_n(x,y) = \sqrt{n}(F_n(x,y) - F(x,y))$. Since $\mathbb{E}\left(\int A_b(x,y)dG_n(x,y)\right) = 0$,

$$\begin{aligned}
\mathrm{Var}\left(\int A_b(x,y)dG_n(x,y)\right) &= \mathbb{E}\left(\int A_b(x,y)A_b(v,w)dG_n(x,y)dG_n(v,w)\right) \\
&= \frac{1}{n}\sum_r\sum_s \mathbb{E}\left(A_b\left(X_r,Y_r\right)A_b\left(X_s,Y_s\right)\right) \\
&- \sum_r \mathbb{E}\left(\int A_b\left(X_r,Y_r\right)A_b(v,w)dF(v,w)\right) \\
&- \sum_s \mathbb{E}\left(\int A_b(x,y)A_b\left(X_s,Y_s\right)dF(x,y)\right) \\
&+ n\int A_b(x,y)A_b(v,w)dF(x,y)dF(v,w) \\
&= E_1 + E_2 + E_3 + E_4
\end{aligned}$$

The contribution of terms with $r \neq s$ to E_1 is

$$\begin{aligned}
\frac{1}{n}\sum_{r\neq s}\int A_b(x,y)A_b(v,w)dF^{(r-s)}(x,y,v,w) = \sum_{k=1}^{n}\frac{n-k}{n}\int A_b(x,y)A_b(v,w)\cdot \\
\cdot\left[dF^{(k)}(x,y,v,w) + dF^{(-k)}(x,y,v,w)\right],
\end{aligned}$$

whereas, terms with $r = s$ contribute with $\int A_b(x,y)A_b(x,y)dF(x,y)$. Moreover, $E_2 =$

$E_3 = -E_4$, so that,

$$
\begin{aligned}
\operatorname{Var}\left(\int A_b(x,y)dG_n(x,y)\right) &= \int A_b^2(x,y)dF(x,y) \\
&- n\int A_b(x,y)A_b(v,y)dF(x,y)dF(v,w) \\
&+ \int A_b(x,y)A_b(v,w)\sum_{k=1}^{n}\frac{n-k}{n}\left[dF^{(k)}(x,y,v,w) + dF^{(-k)}(x,y,v,w)\right] \\
&+ \int A_b(x,y)A_b(v,w)dF(x,y)dF(v,w) \\
&- \int A_b(x,y)A_b(v,w)dF(x,y)dF(v,w) \\
&= \int A_b^2(x,y)dF(x,y) \\
&- \int A_b(x,y)A_b(v,w)dF(x,y)dF(v,w) \\
&+ \int A_b(x,y)A_b(v,w)\sum_{k=1}^{n}\frac{n-k}{n}\left[dF^{(k)}(x,y,v,w) + dF^{(-k)}(x,y,v,w)\right] \\
&+ \int A_b(x,y)A_b(v,w)\sum_{k=1}^{n}\frac{1-n}{n}dF(x,y)dF(v,w).
\end{aligned}
$$

Focusing on the last two terms, if we add and subtract $\sum_{k=1}^{n}\frac{2(n-k)}{n}F(x,y)F(v,w)$, we have:

$$
\begin{aligned}
&\int A_b(x,y)A_b(v,w)\sum_{k=1}^{n}\left[\frac{n-k}{n}\left(dF^{(k)}(x,y,v,w) + dF^{(k)}(x,y,v,w)\right) + \frac{1-n}{n}dF(x,y)dF(v,w)\right] \\
&= \int A_b(x,y)A_b(v,w)\sum_{k=1}^{n}\frac{n-k}{n}\left[dF^{(k)}(x,y,v,w) - dF(x,y)dF(v,w)\right] \\
&+ \int A_b(x,y)A_b(v,w)\sum_{k=1}^{n}\frac{n-k}{n}\left[dF^{(-k)}(x,y,v,w) - dF(x,y)dF(v,w)\right] \ .
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\operatorname{Var}\left(\int A_b(x,y)dG_n(x,y)\right) &= \int A_b^2(x,y)dF(x,y) - \int A_b(x,y)A_b(v,w)dF(x,y)dF(v,w) \\
&+ \int A_b(x,y)A_b(v,w)\sum_{k=1}^{n}\frac{n-k}{n}\left[dF^{(k)}(x,y,v,w) - dF(x,y)dF(v,w)\right] \\
&+ \int A_b(x,y)A_b(v,w)\sum_{k=1}^{n}\frac{n-k}{n}\left[dF^{(-k)}(x,y,v,w) - dF(x,y)dF(v,w)\right]
\end{aligned}
$$

To check whether the variance converges, we need to prove the convergence of the sum in the last two integrals. Using Assumption **A1** and Corollary A.2 of Hall and Heyde (1980), and choosing arbitrary $1 < q < p$ such that $\frac{1}{p} + \frac{1}{q} < 1$,

$$
\begin{aligned}
&\Big| \sum_{k=1}^{n} \frac{n-k}{n} \int A_b(x,y) A_b(v,w) \left[f^{(k)}(x,y,v,w) - f(x,y) f(v,w) \right] dx dy dv dw \Big| \\
&\leq \sum_{k=1}^{n} \frac{n-k}{n} \mid \mathbb{E}\left(A_b(X_s, Y_s) A_b(X_{s-k}, Y_{s-k}) \right) - \mathbb{E}\left(A_b(X_s, Y_s) \right) \mathbb{E}\left(A_b(X_{s-k}, Y_{s-k}) \right) \mid \\
&\leq 8 \sum_{k=1}^{n} \frac{n-k}{n} \|A_b\|_p \|A_b\|_q \alpha_k^{1-\frac{1}{p}-\frac{1}{q}} = 8 \|A_b\|_p \|A_b\|_q \, O\left(\frac{(n-1)}{n} \sum_{k=1}^{n} \left(\alpha^{1-\frac{1}{p}-\frac{1}{q}} \right)^k \right) \\
&\xrightarrow{n \to \infty} 8 \frac{\alpha^{1-\frac{1}{p}-\frac{1}{q}}}{1 - \alpha^{1-\frac{1}{p}-\frac{1}{q}}} \|A_b\|_p \|A_b\|_q < \infty,
\end{aligned}
$$

where $f^{(k)}$ is the density function of (X_t, Y_t, X_s, Y_s) with $k = t - s$. The same is true for $k = s - t$. For $n \to \infty$ and $b \to 0$, we need assumption 3 of Theorem 2.2 and use the approach of Joe (1989) and let $b \to 0$ in the expressions for b fixed. Finally, the asymptotic normality of $\int A_b(x,y) dG_n(x,y)$ follows from Francq and Zakoïan (2005) since $p > 2$, Assumption **A1** holds and

$$
\lim_{n \to \infty} \operatorname{Var}\left(\int A_b(x,y) dG_n(x,y) \right) < \infty.
$$

The last part of the theorem follows from the above proof, from Theorem 2.1, and Proposition 6.3.9 of Brockwell and Davis (2006). □

Proof of Corollary 3.1. This follows in a straightforward fashion from the method of proof of Theorem 3.2. □

D Appendix

In this appendix the asymptotic theory for the validity of the bootstrap is given with some auxiliary preliminary results proved in Appendix E.

Let $(\Lambda, \mathcal{G}, \mathbb{P}^*_\omega)$ be the probability space where $\{X^*_t\}$ and $\{Y^*_t\}$, the bootstraps of $\{X_t\}$

and $\{Y_t\}$, respectively, are defined. In particular, for the estimation part, the two times series are bootstrapped together ($\{(X_t, Y_t\}$), while, for the independence testing part, they are bootstrapped separately ($\{X_t\}$ and $\{Y_t\}$). As we stated in Section 2 for $\theta_{0,b}$ and for $\theta_{n,b}$ we assume that the point (x, y), at which the local log-likelihood is estimated, is fixed, so that, $\theta^*_{n,b} = \theta^*_{n,b}(x, y)$, which is the 5-dimensional vector of parameter estimates obtained with the bootstrap.

First, we need some expressions for calculating the mean value and the variance of a time series under the bootstrap and referring to the probability space $(\Lambda, \mathcal{G}, \mathbb{P}^*_\omega)$. For fixed b, consider a stationary time series $\left\{Z_t^{(b)}(\theta)\right\}$, $\theta \in \Theta$ and define $\bar{Z}_n(\theta) = \bar{Z}_{n,b}(\theta) = \frac{1}{n}\sum_{t=1}^n Z_t^{(b)}(\theta)$, $\bar{Z}_{\alpha,n}(\theta) = \bar{Z}_{\alpha,n,b}(\theta) = \sum_{t=1}^n \alpha_n(t) Z_t^{(b)}(\theta)$, where $\alpha_n(t)$ is defined below. Moreover, $\mathbb{E}^*$ is the mean value under the measure $\mathbb{P}^*_\omega$. From Politis and Romano (1994) and Gonçalves and White (2004), we know that, for the stationary bootstrap,

$$\mathbb{E}^*\left(\bar{Z}^*_n(\theta)\right) = \bar{Z}_n(\theta) \tag{D.1}$$

and, again, using Var to denote the covariance matrix,

$$\begin{aligned}
\mathrm{Var}^*\left(\bar{Z}^*_n(\theta)\right) = & \frac{1}{n^2}\sum_{s=1}^{n}(Z_s^{(b)}(\theta) - \bar{Z}_n(\theta))(Z_s^{(b)}(\theta) - \bar{Z}_n(\theta))^t \\
& + \frac{1}{n^2}\sum_{\tau=1}^{n-1}\gamma_n(\tau)\sum_{s=1}^{n-\tau}\Big[(Z_s^{(b)}(\theta) - \bar{Z}_n(\theta))(Z_{s+\tau}^{(b)}(\theta) - \bar{Z}_n(\theta))^t \\
& +(Z_{s+\tau}^{(b)}(\theta) - \bar{Z}_n(\theta))(Z_s^{(b)}(\theta) - \bar{Z}_n(\theta))^t\Big]
\end{aligned} \tag{D.2}$$

and where

$$\gamma_n(\tau) = \left(1 - \frac{\tau}{n}\right)\left(1 - \frac{1}{l_n}\right)^\tau + \frac{\tau}{n}\left(1 - \frac{1}{l_n}\right)^{n-\tau}, \quad \sum_{\tau=1}^{n-1}\gamma_n(\tau) \le l_n \tag{D.3}$$

and for the moving block bootstrap,

$$\mathbb{E}^*\left(\bar{Z}^*_n(\theta)\right) = \sum_{t=1}^{n}\alpha_n(t) Z_t^{(b)} = \bar{Z}_n(\theta) + O_p\left(\frac{l_n}{n}\right) \tag{D.4}$$

$$\operatorname{Var}^*\left(\bar{Z}_n^*(\theta)\right) = \frac{1}{n}\sum_{s=1}^{n}\alpha_n(s)(Z_s^{(b)}(\theta)-\bar{Z}_{\alpha,n}(\theta))(Z_s^{(b)}(\theta)-\bar{Z}_{\alpha,n}(\theta))^t \tag{D.5}$$
$$+\frac{1}{n}\sum_{\tau=1}^{l_n-1}\left(1-\frac{\tau}{l_n}\right)\sum_{s=1}^{n-|\tau|}\beta_n(t,\tau)\Big[(Z_s^{(b)}(\theta)-\bar{Z}_{\alpha,n}(\theta))(Z_{s+\tau}^{(b)}(\theta)-\bar{Z}_{\alpha,n}(\theta))^t$$
$$+(Z_{s+\tau}^{(b)}(\theta)-\bar{Z}_{\alpha,n}(\theta))(Z_s^{(b)}(\theta)-\bar{Z}_{\alpha,n}(\theta))^t\Big]$$

where

$$\alpha_n(t) = \begin{cases} \frac{t}{l_n(n-l_n+1)}, & 1\le t\le l_n-1 \\ \frac{1}{n-l_n+1}, & l\le t\le n-l_n+1 \\ \frac{n-t+1}{l_n(n-l_n+1)}, & n-l_n+2\le t\le n \end{cases}, \quad \sum_{t=1}^{n}\alpha_n(t)=1 \tag{D.6}$$

$$\beta_n(t,\tau) = \begin{cases} \frac{t}{(l_n-|\tau|)(n-l_n+1)}, & 1\le t\le l_n-\mid\tau\mid-1 \\ \frac{1}{n-l_n+1}, & l_n-\mid\tau\mid\le t\le n-l_n+1 \\ \frac{n-t-|\tau|+1}{(l_n-|\tau|)(n-l_n+1)}, & n-l_n+2\le t\le n-\mid\tau\mid \end{cases}, \quad \sum_{t=1}^{n-|\tau|}\beta_n(t,\tau)=1. \tag{D.7}$$

Remark D.1. It is easy to see that $|\gamma_n(t)|\le 2$, $\alpha_n(t)\le\frac{1}{n-l_n+1}$ and $\beta_n(t,\tau)\le\frac{1}{n-l_n+1}$ for every $t,\tau=1,\dots,n$.

In the next theorems, $Z_t^{(b)}$ will be a scalar related to the log-likelihood, i.e. $Z_t^{(b)}(\theta)=L^{(b)}((X_t,Y_t),\theta)$, or a vector being the gradient, i.e. $Z_t^{(b)}(\theta)=\nabla L^{(b)}((X_t,Y_t),\theta)$, or a matrix being the Hessian matrix of the log-likelihood, i.e. $Z_t^{(b)}(\theta)=\nabla^2L^{(b)}((X_t,Y_t),\theta)$.

To the assumptions **A1)-A3)** of Section 2.1, we need to add the assumptions **A4** and **A5** below.

A4) $(\Lambda,\mathcal{G})$ is a measurable space, $(\Lambda,\mathcal{G},\mathbb{P}_\omega^*)$ is a complete probability space, for all $\omega\in\Omega$, and $\left\{L_{n,b}^*:\Lambda\times\Omega\times\Theta\longrightarrow\bar{\mathbb{R}}\right\}$ is a sequence of random functions such that $L_{n,b}^*(\theta)=L_{n,b}((\underline{X}_n^*(\lambda,\omega),\underline{Y}_n^*(\lambda,\omega)),\theta)$, where $X_t^*(\lambda,\omega)=X_{\tau_t^X(\lambda)}(\omega)$, $Y_t^*(\lambda,\omega)=Y_{\tau_t^Y(\lambda)}(\omega)$, $\tau_t^X,\tau_t^Y:\Lambda\longrightarrow\mathbb{N}$, $\omega\in\Omega$, $\lambda\in\Lambda$, τ_t^X and τ_t^Y are vectors of random indexes, representing the bootstrap operation. If the two time series are bootstrapped together, then $\tau_t^X=\tau_t^Y$. Moreover, the block length l_n is such that $l_n=o(\sqrt{n})$ as $n\to\infty$.

A5) For $\theta \in \Theta$ and for every $t = 1, \ldots, n$, $\left\{L^{(b)}((X_t, Y_t), \theta)\right\}$ is Lipschitz continuous on Θ, $\left\{\nabla L^{(b)}((X_t, Y_t), \theta)\right\}$ is 6-dominated on Θ uniformly in t, n, and $\left\{\nabla^2 L^{(b)}((X_t, Y_t), \theta)\right\}$ is Lipschitz continuous on Θ and $2-$dominated on Θ uniformly in t, n.

Assumption **A4** sets the stage for proving the validity of the bootstrap. It defines the probability space of the bootstrapped samples $\{X_t^*\}$ and $\{Y_t^*\}$. Assumption **A5** is needed in Theorem D.2, and it is fulfilled in the situation we consider. Indeed, the kernel function has a compact support and the term $\log(\psi((X_t, Y_t), \theta))$ is differentiable, therefore $\left\{L^{(b)}((X_t, Y_t), \theta)\right\}$ and $\left\{\nabla^2 L^{(b)}((X_t, Y_t), \theta)\right\}$ are Lipschitz continuous. Further, since the kernel function has a compact support and the continuous term $\log(\psi((X_t, Y_t), \theta))$ has a maximum in the compact set Θ, so $\left\{\nabla L^{(b)}((X_t, Y_t), \theta)\right\}$ and $\left\{\nabla^2 L^{(b)}((X_t, Y_t), \theta)\right\}$ are bounded by a constant. This implies that the two processes are, respectively, 6-dominated and 2-dominated.

We will use the following notations, taken from Gonçalves and White (2002, 2004), for the convergence of variables in the probability space $(\Lambda, \mathcal{G}, \mathbb{P}_\omega^*)$. First, we write $Y_n^* \xrightarrow{\mathbb{P}_\omega^*, \mathbb{P}} 0$, if for any $\epsilon, \delta > 0$,

$$\lim_{n \to \infty} \mathbb{P}\left(\omega : \mathbb{P}_\omega^*\left(\lambda : \mid Y_n^*(\lambda, \omega) \mid > \epsilon\right) > \delta\right) = 0.$$

Further, we write $Y_n^* \Rightarrow^{d_{\mathbb{P}_\omega^*}} N(0,1)$ prob- $\mathbb{P}$, if for every subsequence $\{n'\}$, there exists a further subsequence $\{n''\}$ such that $Y_{n''}^* \Rightarrow^{d_{\mathbb{P}_\omega^*}} N(0,1)$ a.s. (see Gonçalves and White (2004), page 210). This definition is based on the fact that convergence in probability implies almost sure convergence for such kinds of subsequences (see Theorem 20.5 of Billingsley (2012)).

D.1 Estimation

We need to show that the bootstrap is valid and to do that we need to prove the consistency (Theorem D.1) and the asymptotic normality (Theorem D.2) of the parameter estimates after the bootstrap.

Theorem D.1 *Let assumptions **A1)**, **A2)** and **A4)** hold, and let* $\left\{L^{(b)}((X_t, Y_t), \theta)\right\}$,

$\theta \in \Theta$, be Lipschitz continuous on Θ and 4-dominated on Θ uniformly in t, n.. Moreover, assume that, as $b = b_n \to 0$, $\sigma_b^2 \doteq Var(L^{(b)}(X_t, Y_t), \theta) \to \sigma^2 < \infty$ and $\mu_b \doteq \mathbb{E}(L^{(b)}(X_t, Y_t), \theta)) \to \mu < \infty$, for all $t = 1, \ldots, n$ and $\theta \in \Theta$. Then $\theta^*_{n,b} - \theta_{n,b} \xrightarrow{\mathbb{P}^*_\omega, \mathbb{P}} 0$.

Proof of Theorem D.1. Using the same arguments as in Theorem 2.1 with (D.1) and (D.4) instead of the stationary condition and Lemmas E.1 and E.2 in Appendix E, it is seen that $\theta^*_{n,b} - \theta_{0,b} \xrightarrow{\mathbb{P}^*_\omega, \mathbb{P}} 0$. Therefore,

$$\begin{aligned}\mathbb{P}\left(\mathbb{P}^*_\omega\left(\mid \theta^*_{n,b} - \theta_{n,b} \mid > \delta\right) > \xi\right) \leq \mathbb{P}\left(\mathbb{P}^*_\omega\left(\mid \theta^*_{n,b} - \theta_{0,b} \mid > \frac{\delta}{2}\right) > \frac{\xi}{2}\right) \\ + \mathbb{P}\left(\mid \theta_{n,b} - \theta_{0,b} \mid > \frac{\xi}{2}\right) \to 0,\end{aligned}$$

as $n \to \infty$ and $\forall \delta, \xi > 0$. □

Theorem D.2 *Let the assumptions of Theorem 2.2 and* ***A1) - A5)*** *hold, then* $\sqrt{b_1 b_2} M_{n,b}^{-\frac{1}{2}} J_{n,b} \sqrt{n}\left(\theta^*_{n,b} - \theta_{n,b}\right) \Rightarrow^{d_{\mathbb{P}^*_\omega}} N(0, I_5)$ *prob* $- \mathbb{P}$.

Proof of Theorem D.2. By **A4** and Lemma A.1 of Gonçalves and White (2004), X^*_n is $\mathcal{G}-$measurable, $\forall \omega \in \Omega$. Therefore, $L^*_{n,b}$ is $\mathcal{G}-$measurable, $\forall (\omega, \theta) \in \Omega \times \Theta$. By definition, $L^*_{n,b}$ is continuously differentiable of order 2 on Θ $a.s. - \mathbb{P}$. Taylor expanding $\nabla L^*_{n,b}(\theta)$, we have that

$$\begin{aligned}\nabla L^*_{n,b}(\theta_{n,b}) &= \nabla L^*_{n,b}(\theta^*_{n,b}) + \nabla^2 L^*_{n,b}(\theta^*_{n,b})\left(\theta_{n,b} - \theta^*_{n,b}\right) + o_{p^*}(1) \\ &= \left[\nabla^2 L^*_{n,b}(\theta^*_{n,b}) - \nabla^2 L^*_{n,b}(\theta_{0,b})\right]\left(\theta_{n,b} - \theta^*_{n,b}\right) \\ &+ \nabla^2 L^*_{n,b}(\theta_{0,b})\left(\theta_{n,b} - \theta^*_{n,b}\right) + o_{p^*}(1),\end{aligned}$$

where $o_{p^*}(1)$ denotes the small order under $\mathbb{P}^*_\omega$. To prove this theorem, we first need to ensure that $\left[\nabla^2 L^*_{n,b}(\theta^*_{n,b}) - \nabla^2 L^*_{n,b}(\theta_{0,b})\right]\left(\theta_{n,b} - \theta^*_{n,b}\right) \xrightarrow{\mathbb{P}^*_\omega, \mathbb{P}} 0$ and second calculate the asymptotic distribution of $\nabla L^*_{n,b}(\theta_{n,b})$. For the first part, we have that

$$\nabla^2 L^*_{n,b}(\theta^*_{n,b}) - \nabla^2 L^*_{n,b}(\theta_{0,b}) = \left[\nabla^2 L^*_{n,b}(\theta^*_{n,b}) - \nabla^2 L_{n,b}(\theta^*_{n,b})\right] + \left[\nabla^2 L_{n,b}(\theta^*_{n,b}) - \nabla^2 L_{n,b}(\theta_{0,b})\right] + \left[\nabla^2 L_{n,b}(\theta_{0,b}) - \nabla^2 L^*_{n,b}(\theta_{0,b})\right] \doteq E_1 + E_2 + E_3.$$

Applying Lemma E.1 and Lemma E.2 to each component of $Z_t^{(b)}(\theta) = \nabla^2 L^{(b)}((X_t, Y_t), \theta)$, it follows that $\nabla^2 L^*_{n,b}(\theta) - \nabla^2 L_{n,b}(\theta) \xrightarrow{\mathbb{P}^*_\omega, \mathbb{P}} 0$ uniformly on Θ. We have that, for suitable matrix and vector norms $|\cdot|$,

$$| E_1 | \leq \sup_{\theta \in \Theta} | \nabla^2 L^*_{n,b}(\theta) - \nabla^2 L_{n,b}(\theta) | \xrightarrow{\mathbb{P}^*_\omega, \mathbb{P}} 0,$$

$$\begin{aligned} | E_2 | &\leq \frac{1}{n} \sum_{t=1}^{n} | \nabla^2 L^{(b)}((X_t, Y_t), \theta^*_{n,b}) - \nabla^2 L^{(b)}((X_t, Y_t), \theta_{0,b}) | \\ &\leq \frac{1}{n} \sum_{t=1}^{n} C_t \mid \theta^*_{n,b} - \theta_{0,b} | \leq M \mid \theta^*_{n,b} - \theta_{0,b} | \xrightarrow{\mathbb{P}^*_\omega, \mathbb{P}} 0 \end{aligned}$$

and

$$| E_3 | \leq \sup_{\theta \in \Theta} | \nabla^2 L_{n,b}(\theta) - \nabla^2 L^*_{n,b}(\theta) | \xrightarrow{\mathbb{P}^*_\omega, \mathbb{P}} 0,$$

where M is a sufficiently large constant such that $C_t \leq M$, for every $t = 1, \ldots, n$. Moreover, we have that

$$\begin{aligned} \sqrt{n} \left(\nabla L^*_{n,b}(\theta_{n,b}) - \nabla L_{n,b}(\theta_{n,b})\right) &= \sqrt{n} \left(\nabla L^*_{n,b}(\theta_{0,b}) - \nabla L_{n,b}(\theta_{0,b})\right) \\ &\quad - \sqrt{n} \left(\nabla L_{n,b}(\theta_{n,b}) - \nabla L_{n,b}(\theta_{0,b})\right) \\ &\quad + \sqrt{n} \left(\nabla L^*_{n,b}(\theta_{n,b}) - \nabla L^*_{n,b}(\theta_{0,b})\right) \doteq E_1 + E_2 + E_3. \end{aligned}$$

Taylor expanding E_2 and E_3, we have, in probability,

$$\sqrt{b_1 b_2} M_{n,b}^{-\frac{1}{2}} (E_2 + E_3) \sim \sqrt{b_1 b_2} M_{n,b}^{-\frac{1}{2}} \left(\nabla^2 L^*_{n,b}(\theta_{0,b}) - \nabla^2 L_{n,b}(\theta_{0,b})\right) \sqrt{n}(\theta_{n,b} - \theta_{0,b}),$$

but, from Theorem 2.2, $\sqrt{b_1 b_2} M_{n,b}^{-\frac{1}{2}} J_{n,b} \sqrt{n}\,(\theta_{n,b} - \theta_{0,b}) \Rightarrow N(0, I_5)$, which implies, together with **A3**, that $\sqrt{n}\sqrt{b_1 b_2}(\theta_{n,b} - \theta_{0,b}) = O_p(1)$. Moreover, applying Lemma E.1 to each element of $Z_t^{(b)}(\theta) = \nabla^2 L^{(b)}((X_t, Y_t), \theta)$, it follows that:

$$\nabla^2 L_{n,b}^*(\theta) - \nabla^2 L_{n,b}(\theta) \Rightarrow^{d_{\mathbb{P}_\omega^*}} 0 \text{ prob} - \mathbb{P},\ \forall \theta \in \Theta\ .$$

Then, from the results above and from **A3**, it follows that

$$\sqrt{b_1 b_2} M_{n,b}^{-\frac{1}{2}} (E_2 + E_3) \Rightarrow^{d_{\mathbb{P}_\omega^*}} 0 \text{ prob} - \mathbb{P}\ .$$

By the definition of θ_n, $\sqrt{n} \nabla L_{n,b}(\theta_{n,b}) = 0$, therefore,

$$\begin{aligned} &\sqrt{n}\sqrt{b_1 b_2} M_{n,b}^{-\frac{1}{2}} \nabla L_{n,b}^*(\theta_{n,b}) - \sqrt{b_1 b_2} M_{n,b}^{-\frac{1}{2}} E_1 \\ &= \sqrt{b_1 b_2} M_{n,b}^{-\frac{1}{2}} (E_2 + E_3) \Rightarrow^{d_{\mathbb{P}_\omega^*}} 0 \text{ prob} - \mathbb{P}. \end{aligned}$$

Applying (E.5) of Theorem E.1 of Appendix E with $Z_t^{(b)}(\theta) = \nabla L^{(b)}((X_t, Y_t), \theta)$ and $V_n = (b_1 b_2)^{-1} M_{n,b}$, we have that $\sqrt{b_1 b_2} M_{n,b}^{-\frac{1}{2}} E_1 \Rightarrow^{d_{\mathbb{P}_\omega^*}} N(0, I_5)$ prob- $\mathbb{P}$. Therefore,

$$\sqrt{n}\sqrt{b_1 b_2} M_{n,b}^{-\frac{1}{2}} \nabla L_{n,b}^*(\theta_{n,b}) \Rightarrow^{d_{\mathbb{P}_\omega^*}} N(0, I_5) \text{ prob} - \mathbb{P}\ .$$

Now, putting everything together, using Theorem D.1 with $Z_t^{(b)}(\theta) = \nabla^2 L_t^{(b)}(\theta)$ and Lemma E.1

$$\begin{aligned} \nabla L_{n,b}^*(\theta_{n,b}) &= \nabla^2 L_{n,b}^*(\theta_{0,b}) \left(\theta_{n,b} - \theta_{n,b}^*\right) + o_{p^*}(1) \\ &= \nabla^2 L_{n,b}(\theta_{0,b}) \left(\theta_{n,b} - \theta_{n,b}^*\right) + o_{p^*}(1) \end{aligned}$$

Using the Markov inequality, the mixing inequality (Corollary A.2, Hall and Heyde (1980)) and the 4-dominance of $\left\{\nabla^2 L_t^{(b)}(\theta)\right\}$, one can show, using the same technique as in (E.3), that $\nabla^2 L_{n,b}(\theta_{0,b}) - \mathbb{E}\left(\nabla^2 L_{n,b}(\theta_{0,b})\right) \xrightarrow{\mathbb{P}} 0$. This leads to

$$\begin{aligned}\nabla L^*_{n,b}(\theta_{n,b}) &= \mathbb{E}\left(\nabla^2 L_{n,b}(\theta_{0,b})\right)\left(\theta_{n,b} - \theta^*_{n,b}\right) + o_p(1) + o_{p^*}(1)\\ &= -J_{n,b}\left(\theta_{n,b} - \theta^*_{n,b}\right) + o_p(1) + o_{p^*}(1).\end{aligned}$$

Therefore,

$$\sqrt{b_1 b_2} M_{n,b}^{-\frac{1}{2}} J_{n,b} \sqrt{n}\left(\theta^*_{n,b} - \theta_{n,b}\right) \Rightarrow^{d_{\mathbb{P}^*_\omega}} N(0, I_5) \text{ prob} - \mathbb{P}.$$

$\square$

Like in Section 2.1, Theorems D.1 and D.2 still hold if we consider the case of $\{(X_t, Y_{t-k})\}$ instead of $\{(X_t, Y_t)\}$, because to prove them it is just sufficient to substitute θ with $\theta^{(k)}$. Also in this case, the matrices $M_{n,b}$ and $J_{n,b}$ will depend on the lag k. Moreover, the same asymptotic results hold with essentially the same proofs for the bootstrapped time series $X_t^{n,*}$ and $Y_t^{n,*}$, where, $X_t^{n,*}$ and $Y_t^{n,*}$ are originated from the transformed series $\{X_t^n, Y_t^n\}$.

D.2 Testing

To ensure the validity of the bootstrap, we need to check the consistency (see Theorem D.3 below) and the asymptotic normality (see Theorem D.4) of the test statistic $T^*_{n,b}$. We can do that for both the stationary and the moving block bootstrap.

Theorem D.3 *Under the assumptions of Theorem D.1 and assuming that h is continuous, it follows that $T^*_{n,b} - T_{n,b} \xrightarrow{\mathbb{P}^*_\omega, \mathbb{P}} 0$.*

Proof of Theorem D.3. The proof is just an application of the continuous mapping theorem to the result obtained in Theorem D.1. $\square$

Theorem D.4 *Under the assumptions of Corollary 3.1 and Theorem 2.1, and **A4**, it follows that*

$$\sqrt{n}\left[C_n\left(A_b\right)\right]^{-\frac{1}{2}}\left(T^*_{n,b} - T_{n,b}\right) \Rightarrow^{d_{\mathbb{P}^*_\omega}} N\left(0,1\right) \;\; prob - \mathbb{P}.$$

Proof of Theorem D.4. For the stationary bootstrap, the proof of this theorem is the same one as for Theorem 4.4 part **B** in Lacal and Tjøstheim (2016). Instead, for the moving block bootstrap, we need to check whether $\mathbb{E}^*\left(-I^*((x,y),\theta_{n,b}(x,y))\right) \xrightarrow{\mathbb{P}} -I((x,y),\theta_{0,b}(x,y))$ still holds, where $I((x,y),\theta) = \dfrac{\partial^2}{\partial\theta^2}L_{n,b}((x,y),\theta)$. This is true because, using the same reasoning of Theorem 4.4 part **B** in Lacal and Tjøstheim (2016) and (D.4),

$$\mathbb{E}^*\left(-I^*((x,y),\theta_{n,b}(x,y))\right) = -I((x,y),\theta_{n,b}(x,y)) + O_p\left(\frac{l_n}{n}\right) \sim -I((x,y),\theta_{0,b}(x,y)),$$

where the last approximation holds in probability. □

The situation with a standardized and normalized series $\{(X_t^n, Y_t^n)\}$ can be treated in essentially the same way, with modification mentioned in the remark after Corollary 3.1, under the null hypothesis of independence between $\{X_t\}$ and $\{Y_t\}$, and with the test functional $T_{n,b}$ having the same asymptotic distribution. Finally, a bootstrap version of the test functionals in Remark 3.2 can be constructed.

E Appendix

The proofs of the lemmas are inspired by corresponding lemmas in Gonçalves and White (2002, 2004). We prove Lemmas E.1 and E.2 in the situation in which $\left\{Z_t^{(b)}\right\}$ is a scalar, but the results still hold with notational changes when $\left\{Z_t^{(b)}\right\}$ is a vector or a matrix. Indeed, as is well known (Brockwell and Davis (2006)), the convergence in probability of a vector or a matrix holds when convergence in probability of their components holds.

Lemma E.1 *Let $\left\{Z_t^{(b)}(\theta)\right\}$, $\theta \in \Theta$, be a stationary and 2-dominated on Θ uniformly in t, n, $\forall t = 1, \ldots, n$, $l_n = o(n)$ and $l_n \to \infty$. Moreover, assume that, as $b = b_n \to 0$, $\sigma_b^2 \doteq Var(Z_t^{(b)}) \to \sigma^2 < \infty$ and $\mu_b \doteq \mathbb{E}(Z_t^{(b)}) \to \mu < \infty$. Then, $\forall\theta \in \Theta$, Θ a compact set,*

$$\bar{Z}_n^*(\theta) - \bar{Z}_n(\theta) \xrightarrow{\mathbb{P}_\omega^*, \mathbb{P}} 0$$

both for the stationary and for the moving block bootstrap.

Proof. Define $\Delta < \infty$ such that, by the 2-dominance, $\|Z_t^{(b)}\|_2 < \Delta$. This is reasonable since $\|Z_t^{(b)}\|_2 = \left(\mathbb{E}\left(|Z_t^{(b)}|^2\right)\right)^{\frac{1}{2}} = (\sigma_b^2 + \mu_b^2)^{\frac{1}{2}} \to (\sigma^2 + \mu^2)^{\frac{1}{2}}$ as $b \to 0$.

$$\bar{Z}_n^*(\theta) - \bar{Z}_n(\theta) = \left[\bar{Z}_n^*(\theta) - \mathbb{E}^*\left(\bar{Z}_n^*(\theta)\right)\right] + \left[\mathbb{E}^*\left(\bar{Z}_n^*(\theta)\right) - \bar{Z}_n(\theta)\right] = E_1 + E_2$$

For the second term, using (D.1) and (D.4), we have that:

(i) for the stationary bootstrap, $E_2 = 0$;

(ii) for the moving block bootstrap, $E_2 = O_p\left(\frac{l_n}{n}\right)$.

Therefore, in both cases, $E_2 \xrightarrow{\mathbb{P}} 0$. Applying the Markov inequality twice, we have:

$$\mathbb{P}\left(\mathbb{P}_\omega^*\left(\mid E_1 \mid > \delta\right) > \xi\right) \leq \mathbb{P}\left(\mathrm{Var}^*\left(\bar{Z}_n^*(\theta)\right) > \delta^2\xi\right) \leq \frac{1}{\delta^2\xi}\|\mathrm{Var}^*\left(\bar{Z}_n^*(\theta)\right)\|_1.$$

By (D.2), (D.5), Remark D.1, the Minkowski inequality, the Hölder inequality and the 2-dominance property of $\left\{Z_t^{(b)}(\theta)\right\}$,

(i) for the stationary bootstrap,

$$\begin{aligned}
\|\mathrm{Var}^*\left(\bar{Z}_n^*(\theta)\right)\|_1 \leq{}& \frac{1}{n^2}\sum_{s=1}^{n}\|Z_s^{(b)}(\theta) - \bar{Z}_n(\theta)\|_2\|Z_s^{(b)}(\theta) - \bar{Z}_n(\theta)\|_2 \\
&+ \frac{1}{n^2}\sum_{\tau=1}^{n-1}\gamma_n(\tau)\sum_{s=1}^{n-\tau}\Big[\|Z_s^{(b)}(\theta) - \bar{Z}_n(\theta)\|_2\|Z_{s+\tau}^{(b)}(\theta) - \bar{Z}_n(\theta)\|_2 \\
&+\|Z_{s+\tau}^{(b)}(\theta) - \bar{Z}_n(\theta)\|_2\|Z_s^{(b)}(\theta) - \bar{Z}_n(\theta)\|_2\Big] \\
\leq{}& \frac{1}{n^2}n4\Delta^2 + \frac{1}{n^2}\sum_{\tau=1}^{n-1}\gamma_n(\tau)8\Delta^2(n-\tau) \leq \frac{4\Delta^2}{n}\left(1 + \frac{2}{n}(n-1)l_n\right).
\end{aligned}$$

(*ii*) for the moving block bootstrap,

$$\begin{aligned}
\|\mathrm{Var}^*\left(\bar{Z}_n^*(\theta)\right)\|_1 \le & \frac{1}{n}\sum_{s=1}^{n}\alpha_n(s)\|Z_s^{(b)}(\theta)-\bar{Z}_{\alpha,n}(\theta)\|_2\|Z_s^{(b)}(\theta)-\bar{Z}_{\alpha,n}(\theta)\|_2 \\
& + \frac{1}{n}\sum_{\tau=1}^{l_n-1}\left(1-\frac{\tau}{l_n}\right)\sum_{s=1}^{n-\tau}\beta_n(t,\tau)\Big[\|Z_s^{(b)}(\theta)-\bar{Z}_{\alpha,n}(\theta)\|_2\|Z_{s+\tau}^{(b)}(\theta)-\bar{Z}_{\alpha,n}(\theta)\|_2 \\
& + \|Z_{s+\tau}^{(b)}(\theta)-\bar{Z}_{\alpha,n}(\theta)\|_2\|Z_s^{(b)}(\theta)-\bar{Z}_{\alpha,n}(\theta)\|_2\Big] \\
& \le \frac{4\Delta^2}{n}\sum_{s=1}^{n}\alpha_n(s) + \frac{8\Delta^2}{n}\sum_{\tau=1}^{l_n-1}\left(1-\frac{\tau}{l_n}\right)\sum_{s=1}^{n-\tau}\beta_n(t,\tau) \\
& = \frac{4\Delta^2}{n}\left(1+2\sum_{\tau=1}^{l_n-1}\left(1-\frac{\tau}{l_n}\right)\right) = \frac{2\Delta^2 l_n}{n}.
\end{aligned}$$

Therefore,

$$\mathbb{P}\left(\mathbb{P}_\omega^*\left(\mid \bar{Z}_n^*(\theta)-\bar{Z}_n(\theta)\mid > \delta\right) > \xi\right) = O\left(\frac{l_n}{n}\right).$$

□

Lemma E.2 *Let* $\left\{Z_t^{(b)}(\theta)\right\}$, $\theta \in \Theta$, *be Lipschitz continuous on* Θ, *that is* $\mid Z_t^{(b)}(\theta) - Z_t^{(b)}(\theta') \mid \le C_t \mid \theta-\theta' \mid$ *a.s.* $-\ \mathbb{P}$, $\forall\theta,\theta' \in \Theta$ *with a sufficiently large constant* M *such that* $C_t \le M$. *Moreover, assume that* $\forall\theta\in\Theta$, Θ *a compact set,* $\bar{Z}_n^*(\theta)-\bar{Z}_n(\theta) \xrightarrow{\mathbb{P}_\omega^*,\mathbb{P}} 0$, *that is Lemma E.1 holds, with* $l_n = o(n)$ *and* $l_n \to \infty$. *Then,* $\forall\delta,\ \xi>0$,

$$\lim_{n\to\infty}\mathbb{P}\left(\mathbb{P}_\omega^*\left(\sup_{\theta\in\Theta}\mid \bar{Z}_n^*(\theta)-\bar{Z}_n(\theta)\mid > \delta\right) > \xi\right) = 0$$

both for the stationary and for the moving block bootstrap.

Proof. The idea of this proof is equal to the one of Lemma C.2 of Lacal and Tjøstheim (2017). The only thing that we need to check is whether the following expression still holds:

$$\mathbb{P}\left(\mathbb{P}_\omega^*\left(\sup_{\theta\in\eta(\theta_i,\epsilon)}\mid \bar{Z}_n^*(\theta)-\bar{Z}_n^*(\theta_i)\mid > \frac{\delta}{3}\right) > \frac{\xi}{3}\right) < \frac{\zeta}{3},$$

$\forall n$ sufficiently large and $\forall\zeta>0$, where $\eta(\theta_i,\epsilon) = \{\theta\in\Theta : |\theta-\theta_i| < \epsilon\}$ with $\epsilon>0$. By the

Markov inequality and the Fatou's lemma for series,

$$\mathbb{P}^*_\omega\left(\sup_{\theta\in\eta(\theta_i,\epsilon)} \mid \bar{Z}^*_n(\theta)-\bar{Z}^*_n(\theta_i) \mid > \frac{\delta}{3}\right) \leq \frac{3}{\delta}\mathbb{E}^*\left(\sup_{\theta\in\eta(\theta_i,\epsilon)} \mid \bar{Z}^*_n(\theta)-\bar{Z}^*_n(\theta_i) \mid\right)$$
$$\leq \frac{3}{\delta n}\sum_{t=1}^{n}\mathbb{E}^*\left(\sup_{\theta\in\eta(\theta_i,\epsilon)} \mid Z^{*(b)}_t(\theta)-Z^{*(b)}_t(\theta_i) \mid\right).$$

For the stationary bootstrap, using (D.1), we have that

$$\begin{aligned}\mathbb{P}^*_\omega\left(\sup_{\theta\in\eta(\theta_i,\epsilon)} \mid \bar{Z}^*_n(\theta)-\bar{Z}^*_n(\theta_i) \mid > \frac{\delta}{3}\right) &\leq \frac{3}{\delta n}\sum_{t=1}^{n}\mathbb{E}^*\left(\sup_{\theta\in\eta(\theta_i,\epsilon)} \mid Z^{*(b)}_t(\theta)-Z^{*(b)}_t(\theta_i) \mid\right)\\ &\leq \frac{3}{\delta n}\sum_{t=1}^{n}\sup_{\theta\in\eta(\theta_i,\epsilon)} \mid Z^{(b)}_t(\theta)-Z^{(b)}_t(\theta_i) \mid\\ &\leq \frac{3\epsilon}{\delta n}\sum_{t=1}^{n}C_t \leq \frac{3\epsilon}{\delta(n-l_n+1)}\sum_{t=1}^{n}C_t\end{aligned}$$

and for the moving block bootstrap, using Remark D.1, we have that

$$\begin{aligned}\mathbb{P}^*_\omega\left(\sup_{\theta\in\eta(\theta_i,\epsilon)} \mid \bar{Z}^*_n(\theta)-\bar{Z}^*_n(\theta_i) \mid > \frac{\delta}{3}\right) &\leq \frac{3}{\delta n}\sum_{t=1}^{n}\mathbb{E}^*\left(\sup_{\theta\in\eta(\theta_i,\epsilon)} \mid Z^{*(b)}_t(\theta)-Z^{*(b)}_t(\theta_i) \mid\right)\\ &\leq \frac{3}{\delta}\sum_{t=1}^{n}\alpha_n(t)\sup_{\theta\in\eta(\theta_i,\epsilon)} \mid Z^{(b)}_t(\theta)-Z^{(b)}_t(\theta_i) \mid + O_p\left(\frac{l_n}{n}\right)\\ &\leq \frac{3\epsilon}{\delta(n-l_n+1)}\sum_{t=1}^{n}C_t + O_p\left(\frac{l_n}{n}\right).\end{aligned}$$

Therefore, again using the Markov inequality for both, asymptotically,

$$\begin{aligned}\mathbb{P}\left(\mathbb{P}^*_\omega\left(\sup_{\theta\in\eta(\theta_i,\epsilon)} \mid \bar{Z}^*_n(\theta)-\bar{Z}^*_n(\theta_i) \mid > \frac{\delta}{3}\right) > \frac{\xi}{3}\right) &\leq \mathbb{P}\left(\frac{3\epsilon}{\delta(n-l_n+1)}\sum_{t=1}^{n}C_t > \frac{\xi}{3}\right)\\ &\leq \frac{9\epsilon}{\xi\delta(n-l_n+1)}\mathbb{E}\left(\sum_{t=1}^{n}C_t\right) \leq \frac{9\epsilon M}{\xi\delta} < \frac{\zeta}{3},\end{aligned}$$

$\forall n$ sufficiently large, $l_n = o(n)$, $\forall\zeta > 0$ and $\epsilon < \frac{\zeta\delta}{9M}$. □

Lemmas E.3 and E.4 are needed in Theorem E.1, which is used to prove the asymptotic normality of $Z^{(b)}_t = \nabla L^{(b)}_t(\theta)$.

Lemma E.3 *Let* $\left\{Z^{(b)}_t\right\}$ *be a univariate stationary time series with second moments,* 6-

dominated on Θ uniformly in t, n, $\forall t = 1, \ldots, n$, with $l_n = o(\sqrt{n})$ and $l_n \to \infty$ and α-mixing with $\alpha_k = O(\alpha^k)$, $\alpha \in (0,1)$. Moreover, assume that, as $b = b_n \to 0$, $\sigma_b^2 \doteq Var(Z_t^{(b)}) \to \sigma^2 < \infty$ and $\mu_b \doteq \mathbb{E}(Z_t^{(b)}) \to \mu < \infty$. Then, using the notations of Gonçalves and White (2002), $\hat{\sigma}_n^2 - \sigma_n^2 \xrightarrow{\mathbb{P}} 0$, where $\hat{\sigma}_n^2 = \hat{\sigma}_{n,b}^2 = Var^*\left(\sqrt{n}\bar{Z}_n^*\right)$ and $\sigma_n^2 = \sigma_{n,b}^2 = Var\left(\sqrt{n}\bar{Z}_n\right)$, both for the stationary and for the moving block bootstrap.

Proof. First we state the proof in the situation of the stationary bootstrap and, then, of the moving block bootstrap. Define $\hat{R}(t,\tau) = Z_t^{(b)} Z_{t+\tau}^{(b)}$, $R(t,\tau) = \mathbb{E}\left(Z_t^{(b)} Z_{t+\tau}^{(b)}\right)$ and $\Delta < \infty$ such that $\|Z_t^{(b)}\|_6 \leq \Delta$. Then, we have that,

$$\begin{aligned}\sigma_n^2 &= \frac{1}{n}\mathrm{Var}\left(\sum_{t=1}^n Z_t^{(b)}\right) = \frac{1}{n}\sum_{t=1}^n \mathrm{Var}\left(Z_t^{(b)}\right) + \frac{2}{n}\sum_{\tau=1}^{n-1}\sum_{t=1}^{n-\tau} Cov\left(Z_t^{(b)}, Z_{t+\tau}^{(b)}\right) \\ &= \sigma_b^2 + \frac{2}{n}\sum_{\tau=1}^{n-1}\sum_{t=1}^{n-\tau}\left(R(t,\tau) - \mu_b^2\right).\end{aligned} \tag{E.1}$$

The proof consists of two steps:

1. $\tilde{\sigma}_n^2 - \sigma_n^2 \xrightarrow{\mathbb{P}} 0$;
2. $\hat{\sigma}_n^2 - \tilde{\sigma}_n^2 \xrightarrow{\mathbb{P}} 0$;

where, $\tilde{\sigma}_n^2$ is equal to $\hat{\sigma}_n^2$, except that it replaces $\bar{Z}_n$ (for the stationary bootstrap) and $\bar{Z}_{\alpha,n}$ (for the moving block bootstrap with α defined in Section 3) with μ_b. First, we will prove these two steps for the stationary bootstrap (S) and then, for the moving block bootstrap (M).

Step 1 (S): Since

$$\tilde{\sigma}_n^2 = \frac{1}{n}\sum_{t=1}^n \left(Z_t^{(b)} - \mu_b\right)^2 + \frac{2}{n}\sum_{\tau=1}^{n-1}\sum_{t=1}^{n-\tau}\gamma_n(\tau)\left(Z_t^{(b)} - \mu_b\right)\left(Z_{t+\tau}^{(b)} - \mu_b\right) \tag{E.2}$$

with γ_n defined in (D.3), we have that

$$\begin{aligned}\tilde{\sigma}_n^2 - \sigma_n^2 = &\left[\frac{1}{n}\sum_{t=1}^n\left(Z_t^{(b)} - \mu_b\right)^2 - \sigma_b^2\right] + \frac{2}{n}\sum_{\tau=1}^{n-1}\sum_{t=1}^{n-\tau}\Big[\gamma_n(\tau)\hat{R}(t,\tau) - R(t,\tau) \\ &+\gamma_n(\tau)\mu_b^2 - \gamma_n(\tau)\mu_b\left(Z_t^{(b)} + Z_{t+\tau}^{(b)}\right) + \mu_b^2\Big] = E_1 + E_2.\end{aligned}$$

By the Markov inequality, the mixing inequality (Corollary A.2, Hall and Heyde (1980)) and the fact that $\mathbb{E}\left(E_1\right) = 0$,

$$
\begin{aligned}
\mathbb{P}\left(\mid E_1 \mid > \epsilon\right) &\leq \frac{1}{\epsilon^2}\mathrm{Var}\left(\frac{1}{n}\sum_{t=1}^{n}\left(Z_t^{(b)} - \mu_b\right)^2\right) \\
&= \frac{1}{\epsilon^2 n^2}\sum_{t=1}^{n}\sum_{s=1}^{n}\mathrm{Cov}\left(\left(Z_t^{(b)} - \mu_b\right)^2, \left(Z_s^{(b)} - \mu_b\right)^2\right) \\
&\leq \frac{1}{\epsilon^2 n^2}\sum_{t=1}^{n}\sum_{s=1}^{n} 8\|(Z_t^{(b)} - \mu_b)^2\|_3\|(Z_s^{(b)} - \mu_b)^2\|_3 \alpha^{\frac{|t-s|}{3}} \\
&= \frac{8}{\epsilon^2 n^2}\sum_{t=1}^{n}\sum_{s=1}^{n}\|Z_t^{(b)} - \mu_b\|_6^2\|Z_s^{(b)} - \mu_b\|_6^2 \alpha^{\frac{|t-s|}{3}} \\
&\leq \frac{8(\Delta + |\mu_b|)^4}{\epsilon^2 n^2}\sum_{t=1}^{n}\sum_{s=1}^{n}\alpha^{\frac{|t-s|}{3}}
\end{aligned}
$$

because, by the Minkowski inequality and the 6-dominance of $\left\{Z_t^{(b)}\right\}$, we have that $\|Z_t^{(b)} - \mu_b\|_6^2 \leq (\|Z_t^{(b)}\|_6 + |\mu_b|)^2 \leq (\Delta + |\mu_b|)^2$. Therefore,

$$
\begin{aligned}
\mathbb{P}\left(\mid E_1 \mid > \epsilon\right) &\leq \frac{8(\Delta + |\mu_b|)^4}{\epsilon^2 n^2}\sum_{t=1}^{n}\sum_{s=1}^{n}\alpha^{\frac{|t-s|}{3}} \\
&= \frac{8(\Delta + |\mu_b|)^4}{\epsilon^2 n^2}\left(n + 2\sum_{j=1}^{n-1}(n-j)\alpha^{\frac{j}{3}}\right) \\
&\leq \frac{8(\Delta + |\mu_b|)^4}{\epsilon^2 n^2}\left(n + 2(n-1)\sum_{j=1}^{n-1}\alpha^{\frac{j}{3}}\right) \\
&= \frac{8(\Delta + |\mu_b|)^4}{\epsilon^2 n^2}\left(n + 2(n-1)\frac{\alpha^{\frac{1}{3}} - \alpha^{\frac{n}{3}}}{1 - \alpha^{\frac{1}{3}}}\right) \to 0,
\end{aligned} \tag{E.3}
$$

as $n \to \infty$, since $\alpha \in (0,1)$. To prove that the absolute value of $E_2 \xrightarrow{\mathbb{P}} 0$, we show that the bias and the variance of E_2 go to 0 as $n \to \infty$.

$$
\begin{aligned}
\mid \mathbb{E}\left(E_2\right) \mid &\leq \frac{2}{n}\sum_{\tau=1}^{n-1}\sum_{t=1}^{n-\tau} \mid \left(\gamma_n(\tau) - 1\right) R(t,\tau) - \mu_b^2\left(\gamma_n(\tau) - 1\right) \mid \\
&\leq \frac{2}{n}\sum_{\tau=1}^{n-1}\sum_{t=1}^{n-\tau} \mid \gamma_n(\tau) - 1 \mid\mid R(t,\tau) - \mu_b^2 \mid \to 0,
\end{aligned}
$$

by Remark D.1, the dominated convergence theorem for series, the mixing inequality (Corollary A.2, Hall and Heyde (1980)), since $\frac{1}{n} \mid \gamma_n(\tau) - 1 \mid \cdot \mid R(t,\tau) - \mu_b^2 \mid \to 0$ and, by the fact that $\left\{Z_t^{(b)}\right\}$ is 6-dominated,

$$\sum_{\tau \geq 1} \frac{1}{n} \mid \gamma_n(\tau) - 1 \mid \sum_{t=1}^{n-\tau} \mid R(t,\tau) - \mu_b^2 \mid \leq \sum_{\tau \geq 1} \frac{1}{n} 3(n-\tau) 8\Delta^2 \alpha^{\frac{\tau}{2}} \leq 24\Delta^2 \sum_{\tau \geq 1} \alpha^{\frac{\tau}{2}} < \infty.$$

By the Cauchy-Schwarz inequality,

$$\begin{aligned}
\mathrm{Var}\left(E_2\right) &= \frac{4}{n^2} \mathrm{Var}\left(\sum_{\tau=1}^{n-1} \sum_{t=1}^{n-\tau} \left[\gamma_n(\tau)\hat{R}(t,\tau) - \mu_b \gamma_n(\tau)\left(Z_t^{(b)} + Z_{t+\tau}^{(b)}\right)\right]\right) \\
&= \frac{4}{n^2} \sum_{\tau=1}^{n-1} \sum_{\lambda=1}^{n-1} \gamma_n(\tau)\gamma_n(\lambda) \mathrm{Cov}\left(\sum_{t=1}^{n-\tau} \left[\hat{R}(t,\tau) - \mu_b\left(Z_t^{(b)} + Z_{t+\tau}^{(b)}\right)\right],\right. \\
&\left. \sum_{s=1}^{n-\lambda} \left[\hat{R}(s,\lambda) - \mu_b\left(Z_s^{(b)} + Z_{s+\lambda}^{(b)}\right)\right]\right) \\
&\leq \frac{4}{n^2} \sum_{\tau=1}^{n-1} \sum_{\lambda=1}^{n-1} \gamma_n(\tau)\gamma_n(\lambda) \sqrt{\mathrm{Var}\left(\sum_{t=1}^{n-\tau} \left[\hat{R}(t,\tau) - \mu_b\left(Z_t^{(b)} + Z_{t+\tau}^{(b)}\right)\right]\right)} \cdot \\
&\cdot \sqrt{\mathrm{Var}\left(\sum_{s=1}^{n-\lambda} \left[\hat{R}(s,\lambda) - \mu_b\left(Z_s^{(b)} + Z_{s+\lambda}^{(b)}\right)\right]\right)}.
\end{aligned}$$

By the mixing inequality (Corollary A.2, Hall and Heyde (1980)), we have that

$$\begin{aligned}
&\mathrm{Var}\left(\sum_{t=1}^{n-\tau} \left[\hat{R}(t,\tau) - \mu_b\left(Z_t^{(b)} + Z_{t+\tau}^{(b)}\right)\right]\right) = \mathrm{Var}\left(\sum_{t=1}^{n-\tau} \left[Z_t^{(b)} Z_{t+\tau}^{(b)} - \mu_b\left(Z_t^{(b)} + Z_{t+\tau}^{(b)}\right)\right]\right) \\
&= \sum_{t=1}^{n-\tau} \mathrm{Var}\left(Z_t^{(b)} Z_{t+\tau}^{(b)} - \mu_b\left(Z_t^{(b)} + Z_{t+\tau}^{(b)}\right)\right) \\
&+ 2\sum_{t=1}^{n-\tau} \sum_{s=t+1}^{n-\tau} \mathrm{Cov}\left(Z_t^{(b)} Z_{t+\tau}^{(b)} - \mu_b\left(Z_t^{(b)} + Z_{t+\tau}^{(b)}\right), Z_s^{(b)} Z_{s+\tau}^{(b)} - \mu_b\left(Z_s^{(b)} + Z_{s+\tau}^{(b)}\right)\right) \\
&\leq \sum_{t=1}^{n-\tau} \|Z_t^{(b)} Z_{t+\tau}^{(b)} - \mu_b(Z_t^{(b)} + Z_{t+\tau}^{(b)})\|_2^2 \\
&+ 2\sum_{t=1}^{n-\tau} \sum_{s=t+1}^{n-\tau} 8\|Z_t^{(b)} Z_{t+\tau}^{(b)} - \mu_b(Z_t^{(b)} + Z_{t+\tau}^{(b)})\|_3 \|Z_s^{(b)} Z_{s+\tau}^{(b)} - \mu_b(Z_s^{(b)} + Z_{s+\tau}^{(b)})\|_3 \alpha^{\frac{|t-s|}{3}} \\
&\leq (n-\tau)(\Delta^2 + 2|\mu_b|\Delta)^2 + 16(\Delta^2 + 2|\mu_b|\Delta)^2 \sum_{t=1}^{n-\tau} \sum_{s=t+1}^{n-\tau} \alpha^{\frac{|t-s|}{3}},
\end{aligned}$$

because, by the Minkowski inequality, the generalization of the Hölder inequality and the 6-dominance of $\left\{Z_t^{(b)}\right\}$,

$$
\begin{aligned}
\|Z_t^{(b)} Z_{t+\tau}^{(b)} - \mu_b(Z_t^{(b)} + Z_{t+\tau}^{(b)})\|_3 &\leq \|Z_t^{(b)} Z_{t+\tau}^{(b)}\|_3 + |\mu_b|(\|Z_t^{(b)}\|_3 + \|Z_{t+\tau}^{(b)}\|_3) \\
&\leq \|Z_t^{(b)}\|_6 \|Z_{t+\tau}^{(b)}\|_6 + |\mu_b|(\|Z_t^{(b)}\|_3 + \|Z_{t+\tau}^{(b)}\|_3) \leq \Delta^2 + 2|\mu_b|\Delta.
\end{aligned}
$$

Therefore, using the reasoning of (E.3),

$$
\begin{aligned}
\operatorname{Var}\left(\sum_{t=1}^{n-\tau}\left[\hat{R}(t,\tau) - \mu_b\left(Z_t^{(b)} + Z_{t+\tau}^{(b)}\right)\right]\right) &\leq (\Delta^2 + 2|\mu_b|\Delta)^2\left((n-\tau) + 16\sum_{t=1}^{n-\tau}\sum_{s=t+1}^{n-\tau}\alpha^{\frac{|t-s|}{3}}\right) \\
&\leq (\Delta^2 + 2|\mu_b|\Delta)^2 (n-1)\left(1 + 16\frac{\alpha^{\frac{1}{3}} - \alpha^{\frac{n}{3}}}{1-\alpha^{\frac{1}{3}}}\right).
\end{aligned}
$$

Putting everything together and using the definition of γ_n, we have that,

$$
\begin{aligned}
\operatorname{Var}(E_2) \leq \frac{4}{n^2}\sum_{\tau=1}^{n-1}\sum_{\lambda=1}^{n-1}\gamma_n(\tau)\gamma_n(\lambda) &\sqrt{\operatorname{Var}\left(\sum_{t=1}^{n-\tau}\left[\hat{R}(t,\tau) - \mu_b\left(Z_t^{(b)} + Z_{t+\tau}^{(b)}\right)\right]\right)}\cdot \\
&\cdot\sqrt{\operatorname{Var}\left(\sum_{s=1}^{n-\lambda}\left[\hat{R}(s,\lambda) - \mu_b\left(Z_s^{(b)} + Z_{s+\lambda}^{(b)}\right)\right]\right)} \\
&\leq \frac{4}{n^2}(\Delta^2 + 2|\mu_b|\Delta)^2(n-1)\left(1 + 16\frac{\alpha^{\frac{1}{2}} - \alpha^{\frac{n}{2}}}{1-\alpha^{\frac{1}{2}}}\right)\sum_{\tau=1}^{n-1}\sum_{\lambda=1}^{n-1}\gamma_n(\tau)\gamma_n(\lambda) \\
&\leq \frac{4}{n^2}(\Delta^2 + 2|\mu_b|\Delta)^2(n-1)\left(1 + 16\frac{\alpha^{\frac{1}{2}} - \alpha^{\frac{n}{2}}}{1-\alpha^{\frac{1}{2}}}\right) l_n^2 \to 0,
\end{aligned}
$$

since $l_n = o(\sqrt{n})$ and $\alpha \in (0,1)$.

Step 2 (S): By (D.2) we know that,

$$
\begin{aligned}
\hat{\sigma}_n^2 &= \frac{1}{n}\sum_{t=1}^{n}\left(Z_t^{(b)}-\bar{Z}_n\right)^2+\frac{2}{n}\sum_{\tau=1}^{n-1}\sum_{t=1}^{n-\tau}\gamma_n(\tau)\left(Z_t^{(b)}-\bar{Z}_n\right)\left(Z_{t+\tau}^{(b)}-\bar{Z}_n\right)\\
&= \frac{1}{n}\sum_{t=1}^{n}\left(Z_t^{(b)}\right)^2-\bar{Z}_n^2+\frac{2}{n}\sum_{\tau=1}^{n-1}\sum_{t=1}^{n-\tau}\gamma_n(\tau)Z_t^{(b)}Z_{t+\tau}^{(b)}\\
&+\frac{2}{n}\sum_{\tau=1}^{n-1}\sum_{t=1}^{n-\tau}\gamma_n(\tau)\bar{Z}_n\left(\bar{Z}_n-Z_t^{(b)}-Z_{t+\tau}^{(b)}\right)\\
&= \frac{1}{n}\sum_{t=1}^{n}\left(Z_t^{(b)}\right)^2-\bar{Z}_n^2+\frac{2}{n}\sum_{\tau=1}^{n-1}\sum_{t=1}^{n-\tau}\gamma_n(\tau)Z_t^{(b)}Z_{t+\tau}^{(b)}-2\bar{Z}_n^2\sum_{\tau=1}^{n-1}\left(1-\frac{\tau}{n}\right)\left(1-\frac{1}{l_n}\right)^{\tau}
\end{aligned}
$$

because,

$$
\begin{aligned}
&\frac{2}{n}\sum_{\tau=1}^{n-1}\sum_{t=1}^{n-\tau}\gamma_n(\tau)\bar{Z}_n\left(\bar{Z}_n-Z_t^{(b)}-Z_{t+\tau}^{(b)}\right)=\frac{2}{n}\bar{Z}_n\left[\sum_{\tau=1}^{n-1}\left(\frac{n-\tau}{n}\right)\left(1-\frac{1}{l_n}\right)^{\tau}\right.\\
&\cdot\left.\left(\frac{n-\tau}{n}\sum_{t=1}^{n}Z_t^{(b)}-\sum_{t=1}^{n-\tau}Z_t^{(b)}-\sum_{t=1}^{n-\tau}Z_{t+\tau}^{(b)}\right)\right]+\frac{2}{n}\bar{Z}_n\left[\sum_{\tau=1}^{n-1}\left(\frac{\tau}{n}\right)\left(1-\frac{1}{l_n}\right)^{n-\tau}\right.\\
&\cdot\left.\left(\frac{n-\tau}{n}\sum_{t=1}^{n}Z_t^{(b)}-\sum_{t=1}^{n-\tau}Z_t^{(b)}-\sum_{t=1}^{n-\tau}Z_{t+\tau}^{(b)}\right)\right]=\frac{2}{n}\bar{Z}_n\sum_{\tau=1}^{n-1}\left(\frac{n-\tau}{n}\right)\left(1-\frac{1}{l_n}\right)^{\tau}\\
&\cdot\left(\frac{n-\tau}{n}\sum_{t=1}^{n}Z_t^{(b)}-\sum_{t=1}^{n-\tau}Z_t^{(b)}-\sum_{t=1}^{n-\tau}Z_{t+\tau}^{(b)}+\frac{\tau}{n}\sum_{t=1}^{n}Z_t^{(b)}-\sum_{t=1}^{\tau}Z_t^{(b)}-\sum_{t=1}^{\tau}Z_{t+n-\tau}^{(b)}\right)\\
&=-2\bar{Z}_n^2\sum_{\tau=1}^{n-1}\left(1-\frac{\tau}{n}\right)\left(1-\frac{1}{l_n}\right)^{\tau}.
\end{aligned}
$$

In the same way,

$$
\begin{aligned}
\tilde{\sigma}_n^2 &= \frac{1}{n}\sum_{t=1}^{n}\left(Z_t^{(b)}\right)^2+\mu_b^2-2\mu_b\bar{Z}_n+\frac{2}{n}\sum_{\tau=1}^{n-1}\sum_{t=1}^{n-\tau}\gamma_n(\tau)Z_t^{(b)}Z_{t+\tau}^{(b)}\\
&+2\mu_b\left(\mu_b-2\bar{Z}_n\right)\sum_{\tau=1}^{n-1}\left(1-\frac{\tau}{n}\right)\left(1-\frac{1}{l_n}\right)^{\tau}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\hat{\sigma}_n^2 - \tilde{\sigma}_n^2 &= -\bar{Z}_n^2 - 2\bar{Z}_n^2 \sum_{\tau=1}^{n-1} \left(1 - \frac{\tau}{n}\right)\left(1 - \frac{1}{l_n}\right)^{\tau} - \mu_b^2 + 2\mu_b \bar{Z}_n \\
&\quad - 2\mu_b \sum_{\tau=1}^{n-1} \left(1 - \frac{\tau}{n}\right)\left(1 - \frac{1}{l_n}\right)^{\tau} \left(\mu_b - 2\bar{Z}_n\right) \\
&= -\left(\bar{Z}_n - \mu_b\right)^2 - 2\left(\bar{Z}_n - \mu_b\right)^2 \sum_{\tau=1}^{n-1} \left(1 - \frac{\tau}{n}\right)\left(1 - \frac{1}{l_n}\right)^{\tau} \xrightarrow{\mathbb{P}} 0,
\end{aligned}
$$

because $\left(\bar{Z}_n - \mu_b\right)^2 = O_p\left(\frac{1}{n}\right) = o_p(1)$ and $\left(\bar{Z}_n - \mu_b\right)^2 \sum_{\tau=1}^{n-1}\left(1 - \frac{\tau}{n}\right)\left(1 - \frac{1}{l_n}\right)^{\tau} = O_p\left(\frac{l_n}{n}\right) = o_p(1)$, since $l_n = o(\sqrt{n})$.

The proof for the moving block bootstrap follows the same technique as the one for the stationary bootstrap.

Step 1 (M): Since

$$
\tilde{\sigma}_n^2 = \sum_{t=1}^{n} \alpha_n(t)\left(Z_t^{(b)} - \mu_b\right)^2 + 2\sum_{\tau=1}^{l_n-1}\left(1 - \frac{\tau}{l_n}\right)\sum_{t=1}^{n-\tau} \beta_n(t,\tau)\left(Z_t^{(b)} - \mu_b\right)\left(Z_{t+\tau}^{(b)} - \mu_b\right) \tag{E.4}
$$

with α_n and β_n defined in (D.6) and (D.7), we have that

$$
\begin{aligned}
\tilde{\sigma}_n^2 - \sigma_n^2 &= \left[\sum_{t=1}^{n} \alpha_n(t)\left(Z_t^{(b)} - \mu_b\right)^2 - \sigma_b^2\right] \\
&\quad + 2\left[\sum_{\tau=1}^{l_n-1}\left(1 - \frac{\tau}{l_n}\right)\sum_{t=1}^{n-\tau} \beta_n(t,\tau)\left(\hat{R}(t,\tau) + \mu_b^2 - \mu_b\left(Z_t^{(b)} + Z_{t+\tau}^{(b)}\right)\right)\right. \\
&\quad \left. - \frac{1}{n}\sum_{\tau=1}^{n-1}\sum_{t=1}^{n-\tau}\left(R(t,\tau) - \mu_b^2\right)\right] = E_1 + E_2.
\end{aligned}
$$

By the Markov inequality, the Minkowski inequality, Remark D.1, the 6-dominance of $\left\{Z_t^{(b)}\right\}$, the mixing inequality (Corollary A.2, Hall and Heyde (1980)) and the fact that $\mathbb{E}\left(E_1\right) = 0$,

$$
\begin{aligned}
\mathbb{P}\left(\mid E_1 \mid > \epsilon\right) &\leq \frac{1}{\epsilon^2}\text{Var}\left(\sum_{t=1}^{n}\alpha_n(t)\left(Z_t^{(b)}-\mu_b\right)^2\right)\\
&= \frac{1}{\epsilon^2}\sum_{t=1}^{n}\sum_{s=1}^{n}\alpha_n(t)\alpha_n(s)\text{Cov}\left(\left(Z_t^{(b)}-\mu_b\right)^2,\left(Z_s^{(b)}-\mu_b\right)^2\right)\\
&\leq \frac{8}{\epsilon^2}\sum_{t=1}^{n}\sum_{s=1}^{n}\alpha_n(t)\alpha_n(s)\|(Z_t^{(b)}-\mu_b)^2\|_3\|(Z_s^{(b)}-\mu_b)^2\|_3\alpha^{\frac{|t-s|}{2}}\\
&= \frac{8}{\epsilon^2}\sum_{t=1}^{n}\sum_{s=1}^{n}\alpha_n(t)\alpha_n(s)\|Z_t^{(b)}-\mu_b\|_6^2\|Z_s^{(b)}-\mu_b\|_6^2\alpha^{\frac{|t-s|}{2}}\\
&\leq \frac{8(\Delta+|\mu_b|)^4}{\epsilon^2}\sum_{t=1}^{n}\sum_{s=1}^{n}\alpha_n(t)\alpha_n(s)\alpha^{\frac{|t-s|}{2}}\\
&\leq \frac{8(\Delta+|\mu_b|)^4}{\epsilon^2(n-l_n+1)^2}\sum_{t=1}^{n}\sum_{s=1}^{n}\alpha^{\frac{|t-s|}{2}} \to 0,
\end{aligned}
$$

because of the same reasoning as for (E.3), $l_n = o(\sqrt{n})$ and $\alpha \in (0,1)$. To prove that $E_2 \xrightarrow{\mathbb{P}} 0$, we show that the bias and the variance of E_2 go to 0 as $n \to \infty$.

By Remark D.1,

$$
\begin{aligned}
\mid \mathbb{E}\left(E_2\right) \mid &= 2 \mid \sum_{\tau=1}^{l_n-1}\left(1-\frac{\tau}{l_n}\right)\sum_{t=1}^{n-\tau}\beta_n(t,\tau)\left(R(t,\tau)-\mu_b^2\right) - \frac{1}{n}\sum_{\tau=1}^{n-1}\sum_{t=1}^{n-\tau}\left(R(t,\tau)-\mu_b^2\right) \mid\\
&\leq 2\sum_{\tau=1}^{l_n-1}\frac{\tau}{l_n}\sum_{t=1}^{n-\tau}\mid \beta_n(t,\tau)\mid\mid R(t,\tau)-\mu_b^2\mid + \frac{2}{n}\sum_{\tau=l_n}^{n-1}\sum_{t=1}^{n-\tau}\mid R(t,\tau)-\mu_b^2\mid\\
&+ 2\sum_{\tau=1}^{l_n-1}\sum_{t=1}^{n-\tau}\mid \beta_n(t,\tau)-\frac{1}{n}\mid\mid R(t,\tau)-\mu_b^2\mid\\
&\leq \frac{2}{n-l_n+1}\sum_{\tau=1}^{l_n-1}\frac{\tau}{l_n}\sum_{t=1}^{n-\tau}\mid R(t,\tau)-\mu_b^2\mid + \frac{2}{n}\sum_{\tau=l_n}^{n-1}\sum_{t=1}^{n-\tau}\mid R(t,\tau)-\mu_b^2\mid\\
&+ 2\sum_{\tau=1}^{l_n-1}\mid \frac{1}{n-l_n+1}-\frac{1}{n}\mid\sum_{t=1}^{n-\tau}\mid R(t,\tau)-\mu_b^2\mid\\
&\leq 2\sum_{\tau=1}^{n-1}\sum_{t=1}^{n-\tau}\max\left\{\frac{\tau}{l_n(n-l_n+1)},\frac{1}{n}\right\}\mid R(t,\tau)-\mu_b^2\mid\\
&+ \frac{2(l_n-1)}{n(n-l_n+1)}\sum_{\tau=1}^{l_n-1}\sum_{t=1}^{n-\tau}\mid R(t,\tau)-\mu_b^2\mid\\
&\leq \frac{2n}{n-l_n+1}\sum_{\tau=1}^{n-1}\frac{\tau}{nl_n}\sum_{t=1}^{n-\tau}\mid R(t,\tau)-\mu_b^2\mid + \frac{2(l_n-1)}{n(n-l_n+1)}\sum_{\tau=1}^{l_n-1}\sum_{t=1}^{n-\tau}\mid R(t,\tau)-\mu_b^2\mid
\end{aligned}
$$

The first term is going to 0, because of the dominated convergence theorem for series, since $\frac{\tau}{nl_n} \mid R(t,\tau)-\mu_b^2 \mid \to 0$ and, by the mixing inequality (Corollary A.2, Hall and Heyde (1980)) and the fact that $\left\{Z_t^{(b)}\right\}$ is 6-dominated,

$$\sum_{\tau\geq 1}\frac{\tau}{nl_n}\sum_{t=1}^{n-\tau} \mid R(t,\tau)-\mu_b^2 \mid \leq \sum_{\tau\geq 1}\frac{\tau(n-\tau)8\Delta^2\alpha^{\frac{\tau}{2}}}{nl_n} \leq 8\Delta^2\sum_{\tau\geq 1}\frac{\tau}{l_n}\alpha^{\frac{\tau}{2}} < \infty.$$

Also the second term is going to 0, because, by the mixing inequality (Corollary A.2, Hall and Heyde (1980)) and the fact that $\left\{Z_t^{(b)}\right\}$ is 6-dominated,

$$\frac{2(l_n-1)}{n(n-l_n+1)}\sum_{\tau\geq 1}\sum_{t=1}^{n-\tau} \mid R(t,\tau)-\mu_b^2 \mid \leq \frac{(l_n-1)16\Delta^2(n-1)}{n(n-l_n+1)}\sum_{\tau\geq 1}\alpha^{\frac{\tau}{2}} \to 0.$$

Therefore, $\mathbb{E}\left(\mid E_2 \mid\right)\to 0$. By the Cauchy-Schwarz inequality,

$$\begin{aligned}
\mathrm{Var}\left(E_2\right) &= 4\mathrm{Var}\left(\sum_{\tau=1}^{l_n-1}\left(1-\frac{\tau}{l_n}\right)\sum_{t=1}^{n-\tau}\beta_n(t,\tau)\left(\hat{R}(t,\tau)-\mu_b\left(Z_t^{(b)}+Z_{t+\tau}^{(b)}\right)\right)\right)\\
&= 4\sum_{\tau=1}^{l_n-1}\sum_{\lambda=1}^{l_n-1}\left(1-\frac{\tau}{l_n}\right)\left(1-\frac{\lambda}{l_n}\right)\mathrm{Cov}\left(\sum_{t=1}^{n-\tau}\beta_n(t,\tau)\left[\hat{R}(t,\tau)-\mu_b\left(Z_t^{(b)}+Z_{t+\tau}^{(b)}\right)\right],\right.\\
&\left.\sum_{s=1}^{n-\lambda}\beta_n(s,\lambda)\left[\hat{R}(s,\lambda)-\mu_b\left(Z_s^{(b)}+Z_{s+\lambda}^{(b)}\right)\right]\right)\\
&\leq 4\sum_{\tau=1}^{l_n-1}\sum_{\lambda=1}^{l_n-1}\left(1-\frac{\tau}{l_n}\right)\left(1-\frac{\lambda}{l_n}\right)\sqrt{\mathrm{Var}\left(\sum_{t=1}^{n-\tau}\beta_n(t,\tau)\left[\hat{R}(t,\tau)-\mu_b\left(Z_t^{(b)}+Z_{t+\tau}^{(b)}\right)\right]\right)}\cdot\\
&\cdot\sqrt{\mathrm{Var}\left(\sum_{s=1}^{n-\lambda}\beta_n(s,\lambda)\left[\hat{R}(s,\lambda)-\mu_b\left(Z_s^{(b)}+Z_{s+\lambda}^{(b)}\right)\right]\right)}.
\end{aligned}$$

By the Minkowski inequality, the generalization of the Hölder inequality, the mixing inequality (Corollary A.2, Hall and Heyde (1980)), Remark D.1, the 6-dominance of $\left\{Z_t^{(b)}\right\}$ and the same reasoning of (E.3),

$$
\begin{aligned}
&\mathrm{Var}\left(\sum_{t=1}^{n-\tau}\beta_n(t,\tau)\left[\hat{R}(t,\tau)-\mu_b\left(Z_t^{(b)}+Z_{t+\tau}^{(b)}\right)\right]\right)=\mathrm{Var}\left(\sum_{t=1}^{n-\tau}\beta_n(t,\tau)\left[Z_t^{(b)}Z_{t+\tau}^{(b)}-\mu_b\left(Z_t^{(b)}+Z_{t+\tau}^{(b)}\right)\right]\right)\\
&=\sum_{t=1}^{n-\tau}\beta_n^2(t,\tau)\mathrm{Var}\left(Z_t^{(b)}Z_{t+\tau}^{(b)}-\mu_b\left(Z_t^{(b)}+Z_{t+\tau}^{(b)}\right)\right)\\
&+2\sum_{t=1}^{n-\tau}\sum_{s=t+1}^{n-\tau}\beta_n(t,\tau)\beta_n(s,\tau)\mathrm{Cov}\left(Z_t^{(b)}Z_{t+\tau}^{(b)}-\mu_b\left(Z_t^{(b)}+Z_{t+\tau}^{(b)}\right),Z_s^{(b)}Z_{s+\tau}^{(b)}-\mu_b\left(Z_s^{(b)}+Z_{s+\tau}^{(b)}\right)\right)\\
&\le\sum_{t=1}^{n-\tau}\beta_n^2(t,\tau)\|Z_t^{(b)}Z_{t+\tau}^{(b)}-\mu_b(Z_t^{(b)}+Z_{t+\tau}^{(b)})\|_2^2\\
&+2\sum_{t=1}^{n-\tau}\sum_{s=t+1}^{n-\tau}\beta_n(t,\tau)\beta_n(s,\tau)8\|Z_t^{(b)}Z_{t+\tau}^{(b)}-\mu_b(Z_t^{(b)}+Z_{t+\tau}^{(b)})\|_3\|Z_s^{(b)}Z_{s+\tau}^{(b)}-\mu_b(Z_s^{(b)}+Z_{s+\tau}^{(b)})\|_3\alpha^{\frac{|t-s|}{3}}\\
&\le\frac{(\Delta^2+2|\mu_b|\Delta)^2}{(n-l_n+1)^2}\left((n-\tau)+16\sum_{t=1}^{n-\tau}\sum_{s=t+1}^{n-\tau}\alpha^{\frac{|t-s|}{3}}\right)\\
&=\frac{(\Delta^2+2|\mu_b|\Delta)^2(n-1)}{(n-l_n+1)^2}\left(1+16\frac{\alpha^{\frac{1}{3}}-\alpha^{\frac{n}{3}}}{1-\alpha^{\frac{1}{3}}}\right).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\mathrm{Var}\left(E_2\right)&\le 4\sum_{\tau=1}^{l_n-1}\sum_{\lambda=1}^{l_n-1}\left(1-\frac{\tau}{l_n}\right)\left(1-\frac{\lambda}{l_n}\right)\sqrt{\mathrm{Var}\left(\sum_{t=1}^{n-\tau}\left[\hat{R}(t,\tau)-\mu_b\left(Z_t^{(b)}+Z_{t+\tau}^{(b)}\right)\right]\right)}\cdot\\
&\cdot\sqrt{\mathrm{Var}\left(\sum_{s=1}^{n-\lambda}\left[\hat{R}(s,\lambda)-\mu_b\left(Z_s^{(b)}+Z_{s+\lambda}^{(b)}\right)\right]\right)}\\
&\le\frac{4(\Delta^2+2|\mu_b|\Delta)^2(n-1)}{(n-l_n+1)^2}\left(1+16\frac{\alpha^{\frac{1}{2}}-\alpha^{\frac{n}{2}}}{1-\alpha^{\frac{1}{2}}}\right)\sum_{\tau=1}^{l_n-1}\sum_{\lambda=1}^{l_n-1}\left(1-\frac{\tau}{l_n}\right)\left(1-\frac{\lambda}{l_n}\right)\\
&=\frac{4(\Delta^2+2|\mu_b|\Delta)^2(n-1)}{(n-l_n+1)^2}\left(1+16\frac{\alpha^{\frac{1}{2}}-\alpha^{\frac{n}{2}}}{1-\alpha^{\frac{1}{2}}}\right)\left(\frac{l_n-1}{2}\right)^2\to 0,
\end{aligned}
$$

since $l_n=o(\sqrt{n})$ and $\alpha\in(0,1)$.

Step 2 (M): By (D.5) we know that,

$$
\begin{aligned}
\hat{\sigma}_n^2 &= \sum_{t=1}^{n} \alpha_n(t) \left(Z_t^{(b)} - \bar{Z}_{\alpha,n}\right)^2 + 2 \sum_{\tau=1}^{l_n-1} \left(1 - \frac{\tau}{l_n}\right) \sum_{t=1}^{n-\tau} \beta_n(t,\tau) \left(Z_t^{(b)} - \bar{Z}_{\alpha,n}\right) \left(Z_{t+\tau}^{(b)} - \bar{Z}_{\alpha,n}\right) \\
&= \sum_{t=1}^{n} \alpha_n(t) \left(Z_t^{(b)}\right)^2 - \bar{Z}_{\alpha,n}^2 + 2 \sum_{\tau=1}^{l_n-1} \left(1 - \frac{\tau}{l_n}\right) \sum_{t=1}^{n-\tau} \beta_n(t,\tau) Z_t^{(b)} Z_{t+\tau}^{(b)} \\
&+ 2 \sum_{\tau=1}^{l_n-1} \left(1 - \frac{\tau}{l_n}\right) \sum_{t=1}^{n-\tau} \beta_n(t,\tau) \left(\bar{Z}_{\alpha,n}^2 - \bar{Z}_{\alpha,n} \left(Z_t^{(b)} + Z_{t+\tau}^{(b)}\right)\right)
\end{aligned}
$$

In the same way,

$$
\begin{aligned}
\tilde{\sigma}_n^2 &= \sum_{t=1}^{n} \alpha_n(t) \left(Z_t^{(b)}\right)^2 + \mu_b^2 - 2\mu_b \bar{Z}_{\alpha,n} + 2 \sum_{\tau=1}^{l_n-1} \left(1 - \frac{\tau}{l_n}\right) \sum_{t=1}^{n-\tau} \beta_n(t,\tau) Z_t^{(b)} Z_{t+\tau}^{(b)} \\
&+ 2 \sum_{\tau=1}^{l_n-1} \left(1 - \frac{\tau}{l_n}\right) \sum_{t=1}^{n-\tau} \beta_n(t,\tau) \left(\mu_b^2 - \mu_b \left(Z_t^{(b)} + Z_{t+\tau}^{(b)}\right)\right).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\hat{\sigma}_n^2 - \tilde{\sigma}_n^2 = -\bar{Z}_{\alpha,n}^2 - \mu_b^2 + 2\mu_b \bar{Z}_{\alpha,n} + 2\sum_{\tau=1}^{l_n-1}\left(1-\frac{\tau}{l_n}\right)\sum_{t=1}^{n-\tau}\beta_n(t,\tau)\left[\bar{Z}_{\alpha,n}^2 - \bar{Z}_{\alpha,n}\left(Z_t^{(b)} + Z_{t+\tau}^{(b)}\right)\right. \\
\left. -\mu_b^2 + \mu_b\left(Z_t^{(b)} + Z_{t+\tau}^{(b)}\right)\right] = -\left(\bar{Z}_{\alpha,n} - \mu_b\right)^2 + 2\sum_{\tau=1}^{l_n-1}\left(1-\frac{\tau}{l_n}\right) \\
\cdot \sum_{t=1}^{n-\tau}\beta_n(t,\tau)\left[\bar{Z}_{\alpha,n}^2 - \left(\bar{Z}_{\alpha,n} - \mu_b\right)\left(Z_t^{(b)} + Z_{t+\tau}^{(b)}\right) - \mu_b^2\right] \\
= -\left(\bar{Z}_{\alpha,n} - \mu_b\right)^2 + 2\left(\bar{Z}_{\alpha,n}^2 - \mu_b^2\right)\sum_{\tau=1}^{l_n-1}\left(1-\frac{\tau}{l_n}\right)\sum_{t=1}^{n-\tau}\beta_n(t,\tau) \\
- 2\left(\bar{Z}_{\alpha,n} - \mu_b\right)\sum_{\tau=1}^{l_n-1}\left(1-\frac{\tau}{l_n}\right)\sum_{t=1}^{n-\tau}\beta_n(t,\tau)\left[\left(Z_t^{(b)} - \mu_b\right) + \left(Z_{t+\tau}^{(b)} - \mu_b\right)\right] \\
- 4\mu_b\left(\bar{Z}_{\alpha,n} - \mu_b\right)\sum_{\tau=1}^{l_n-1}\left(1-\frac{\tau}{l_n}\right)\sum_{t=1}^{n-\tau}\beta_n(t,\tau) \\
= -\left(\bar{Z}_{\alpha,n} - \mu_b\right)^2 + \left(\bar{Z}_{\alpha,n}^2 - \mu_b^2\right)(l_n - 1) - 2\mu_b\left(\bar{Z}_{\alpha,n} - \mu_b\right)(l_n - 1) \\
- 2\left(\bar{Z}_{\alpha,n} - \mu_b\right)\sum_{\tau=1}^{l_n-1}\left(1-\frac{\tau}{l_n}\right)\sum_{t=1}^{n-\tau}\beta_n(t,\tau)\left[\left(Z_t^{(b)} - \mu_b\right) + \left(Z_{t+\tau}^{(b)} - \mu_b\right)\right] \\
= \left(\bar{Z}_{\alpha,n}^2 - \mu_b^2\right)(l_n - 2) - 2\left(\bar{Z}_{\alpha,n} - \mu_b\right)\sum_{\tau=1}^{l_n-1}\left(1-\frac{\tau}{l_n}\right) \\
\cdot \sum_{t=1}^{n-\tau}\beta_n(t,\tau)\left[\left(Z_t^{(b)} - \mu_b\right) + \left(Z_{t+\tau}^{(b)} - \mu_b\right)\right] \xrightarrow{\mathbb{P}} 0,
\end{aligned}
$$

because $l_n = o(\sqrt{n})$. Actually, by Remark D.1,

$$
\begin{aligned}
\mathbb{P}\left(\mid \bar{Z}_{\alpha,n} - \mu_b \mid^2 \mid l_n - 2 \mid > \epsilon\right) \le \frac{\mid l_n - 2 \mid}{\epsilon}\mathrm{Var}\left(\bar{Z}_{\alpha,n}\right) \le \frac{\mid l_n - 2 \mid}{\epsilon}\sum_{t=1}^{n}\alpha_n^2(t)\sigma_b^2 \\
\le \frac{\mid l_n - 2 \mid \sigma_b^2 n}{\epsilon(n - l_n + 1)^2} \to 0,
\end{aligned}
$$

$$
\mathbb{P}\left(\mid \bar{Z}_{\alpha,n} - \mu_b \mid > \epsilon\right) \le \frac{1}{\epsilon}\mathrm{Var}\left(\bar{Z}_{\alpha,n}\right) \le \frac{\sigma_b^2 n}{\epsilon(n - l_n + 1)^2} \to 0
$$

and

$$\begin{aligned}
&\mathbb{P}\left(\mid \sum_{\tau=1}^{l_n-1}\left(1-\frac{\tau}{l_n}\right)\sum_{t=1}^{n-\tau}\beta_n(t,\tau)\left[\left(Z_t^{(b)}-\mu_b\right)+\left(Z_{t+\tau}^{(b)}-\mu_b\right)\right]\mid > \epsilon\right)\\
&\leq \frac{1}{\epsilon}\sum_{\tau=1}^{l_n-1}\left(1-\frac{\tau}{l_n}\right)\sum_{t=1}^{n-\tau}\beta_n(t,\tau)\left[\mathbb{E}\left(\mid Z_t^{(b)}-\mu_b\mid\right)+\mathbb{E}\left(\mid Z_{t+\tau}^{(b)}-\mu_b\mid\right)\right]\\
&\leq \frac{2}{\epsilon(n-l_n+1)}\sum_{\tau=1}^{l_n-1}\left(1-\frac{\tau}{l_n}\right)(\Delta+\mid\mu_b\mid)\to 0.
\end{aligned}$$

□

Lemma E.4 *Let* $\left\{Z_t^{(b)}\right\}$ *be a* $d\times 1$ *stationary time series with second moments,* 6*-dominated on* Θ *uniformly in* t,n, $\forall t=1,\ldots,n$, *with* $l_n=o(\sqrt{n})$, $l_n\to\infty$, $n\to\infty$ *and* α*-mixing with* $\alpha_k=O(\alpha^k)$, $\alpha\in(0,1)$. *Moreover, assume that, as* $b\to 0$, $\Sigma_b \doteq Var(Z_t^{(b)})\to\Sigma<\infty$ *and* $\mu_b\doteq\mathbb{E}(Z_t^{(b)})\to\mu<\infty$. *Then* $\hat{V}_n-V_n\xrightarrow{\mathbb{P}}0$, *where* $\hat{V}_n=Var^*\left(\sqrt{n}\bar{Z}_n^*\right)$ *and* $V_n=Var\left(\sqrt{n}\bar{Z}_n\right)$ *are covariance matrices, both for the stationary and for the moving block bootstrap.*

Proof. Define $W_i=\lambda^t Z_i^{(b)}$, $\lambda\in\mathbb{R}^d$ such that $\lambda^t\lambda=1$. If we apply Lemma E.3 to the time series $\{W_i\}$, we have that $\lambda^t\hat{V}_n\lambda-\lambda^t V_n\lambda\xrightarrow{\mathbb{P}}0$ which implies $\hat{V}_n-V_n\xrightarrow{\mathbb{P}}0$. □

To prove the asymptotic normality of the parameter estimates after the bootstrap, we will need the following theorem from Gonçalves and White (2002), adapted to our situation. The only change to be made in the proof is that, in our case, to ensure the consistency of the variance after the bootstrap we use Lemma E.4.

Theorem E.1 *Let* $\left\{Z_t^{(b)}\right\}$ *be a* 5×1 *stationary time series with second moments,* 6*-dominated on* Θ *uniformly in* t,n, $\forall t=1,\ldots,n$, *with* $l_n=o(\sqrt{n})$ *and* $l_n\to\infty$ *and* α*-mixing with* $\alpha_k=O(\alpha^k)$, $\alpha\in(0,1)$. *Moreover, assume that, as* $b\to 0$, $\Sigma_b\doteq Var(Z_t^{(b)})\to\Sigma<\infty$ *and* $\mu_b\doteq\mathbb{E}(Z_t^{(b)})\to\mu<\infty$, $\hat{V}_n\doteq Var^*\left(\sqrt{n}\bar{Z}_n^*\right)$ *and* $V_n\doteq Var\left(\sqrt{n}\bar{Z}_n\right)$ *are covariance* *matrices, and, in particular V_n is positive definite uniformly in n. Then, $V_n = O(1)$,*

$$V_n^{-\frac{1}{2}} \sqrt{n} \left(\bar{Z}_n^* - \bar{Z}_n \right) \Rightarrow^{d_{\mathbb{P}_\omega^*}} N(0, I_5) \ prob - \mathbb{P} \tag{E.5}$$

and, $\forall \epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P} \left(\sup_{x\in\mathbb{R}^d} | \ \mathbb{P}_\omega^* \left(\sqrt{n} \left(\bar{Z}_n^* - \bar{Z}_n \right) \leq x \right) - \mathbb{P} \left(\sqrt{n} \left(\bar{Z}_n - \mu_b \right) \leq x \right) | > \epsilon \right) = 0$$

both for the stationary and for the moving block bootstrap.

F Appendix

For the estimation of the parameters and for the test of independence, we need to calculate the bandwidth and to do this there are several algorithms available. For our purposes, we are using the one based on the likelihood cross-validation technique as was done in Lacal and Tjøstheim (2017). This consists in two steps. First, we calculate the leave-one-out estimate $\theta_{n,b}^{-i}(x, y)$ of $\theta(x, y)$, that is, the maximum point of

$$\frac{1}{n-1} \sum_{j\neq i} K_b \left(X_j - x, Y_j - y \right) \log \left(\psi \left((X_j, Y_j), \theta_b(x, y) \right) \right)$$
$$- \int K_b \left(v - x, w - y \right) \psi \left((v, w), \theta_b(x, y) \right) dvdw$$

Then, we take the bandwidth $b = (b_1, b_2)$ such that the following expression is maximized

$$CV(b) = \frac{1}{n} \sum_i \log \left(\psi \left((X_i, Y_i), \theta_{n,b}^{-i}(x, y) \right) \right) \ .$$

As was done in Lacal and Tjøstheim (2017), to avoid extreme bandwidths, we implement lower and upper bounds for the bandwidth (e.g. expressed as a fraction of the standard deviation). As expected, such bounds are less needed in the standardized case, because the bandwidth is often close to 1 for the sample sizes considered by us. For more details see Lacal and Tjøstheim (2017).

For the choice of the block length for the bootstrap we use the R-function **b.star** of the R-package **np** (for more details see Politis and White (2004) and Patton et al. (2009)). We are aware of the fact that the algorithm proposed by Politis and White (2004) uses the global Pearson serial correlation to calculate the block length. Therefore, it may not be quite appropriate for our purposes, and we have consequently introduced a lower bound on the chosen block length. Note that there is a recent paper by Nordman and Lahiri (2014), where they compare asymptotically different procedures for selecting the block length. It may seem that this paper too, to some extent, is based on autocorrelation concepts.

References

Billingsley, P. (2012), Probability and Measure (Anniversary Edition), *John Wiley & sons*, New York.

Brockwell, P.J., Davis, R.A. (2006), Time Series: Theory and Methods, *Springer*, New York.

Gonçalves, S., White, H. (2002), The bootstrap of the mean for dependent heterogeneous arrays, *Econometric Theory*, 18, 1367-1384.

Gonçalves, S., White, H. (2004), Maximum likelihood and the bootstrap for nonlinear dynamic models, *Journal of Econometrics*, 119, 199-219.

Hall, P., Heyde, C.C. (1980), Martingale Limit Theory and its Application, *Academic Press*, New York-London.

Hjort, N., Jones, M.C. (1996), Locally parametric nonparametric density estimation, *Annals of Statistics*, 24, 1619-1647.

Joe, H. (1989), Estimation of entropy and other functionals of a multivariate density, *Annals of the Institute of Statistical Mathematics*, 41, 683-697.

Klimko, L.A., Nelson, P.I. (1978), On conditional least squares estimation for stochastic processes, *The Annals of Statistics*, 6, 3, 629-642.

Lacal, V., Tjøstheim, D. (2017), Local Gaussian autocorrelation and tests of serial independence, *Journal of Time Series Analysis*, 38, 51-71.

Masry, E., Tjøstheim, D. (1995), Nonparametric estimation and identification of nonlinear ARCH time series: strong convergence and asymptotic normality, *Econometric Theory*, 11, 2, 258-289.

Nordman, D. J., Lahiri, S. N. (2014), Convergence rates of empirical block length selectors for block bootstrap, *Bernoulli*, 20, 2, 958-978.

Patton, A., Politis, D., White, H. (2009), CORRECTION TO-"Automatic block-length selection for the dependent bootstrap", *Econometric Reviews*, 28, 4, 372-375.

Politis, D., Romano, J. (1994), The stationary bootstrap, *Journal of American Statistics Association*, 89, 1303-1313.

Politis, D., White, H. (2004), Automatic block-length selection for the dependent bootstrap, *Econometric Reviews*, 23, 1, 53-70.

Tjøstheim, D., Hufthammer, K. (2013), Local Gaussian correlation: a new measure of dependence, *Journal of Econometrics*, 172, 33-48.