Online Appendix of "LM tests for joint breaks in the dynamics and level of a time series"

Juan J. Dolado, Heiko Rachinger, Carlos Velasco[1]

This version: January 2018.

We provide here three auxiliary lemmas and their proofs, together with the proofs of Theorem 1 to 3, Corollary 1, Propositions 1 to 3 of Dolado, Rachinger and Velasco (2018, DRV henceforth), as well as provide some further simulations for the LM and LMW-type break tests implemented in Regimes 1 and 2.

Lemmata

Lemma A.1. Let $x_{T,t} = r_{T,t}(L)\varepsilon_t$ and $y_{T,j} = s_{T,t}(L)\varepsilon_t$ where $r_{T,t}(L) = \sum_{j=0}^{t} r_{T,j}L^j$ and $s_{T,t}(L) = \sum_{j=0}^{t} s_{T,j}L^j$ denote filters, possibly depending on T, with $|r_{T,j}| \leq K|j|_+^{\varsigma-1}$ and $|s_{T,j}| \leq K|j|_+^{\varsigma-1}$as $T\to\infty$ for $\varsigma < 0.5$, where $|j|_+ = \max\{|j|,1\}$, and ε_r is $iid\left(0,\sigma^2\right)$ with finite fourth moment. Then

$$z_T(\lambda) = T^{-1}\sum_{t=1}^{[T\lambda]}\left(x_{T,t}y_{T,t} - E\left[x_{T,t}y_{T,t}\right]\right) \to_p 0$$

uniformly for $\lambda \in [0,1]$.

Proof. First we show that for every λ, $z_T(\lambda) \to_p 0$ and therefore $z_T(\lambda) \to_d 0$, since $E[z_T(\lambda)] = 0$ and $Var[z_T(\lambda)]$ is equal to

$$\begin{aligned} & T^{-2}\sum_{t=1}^{[T\lambda]}\sum_{t'=1}^{[T\lambda]}\left(E\left[x_{T,t}y_{T,t}x_{T,t'}y_{T,t'}\right] - E\left[x_{T,t}y_{T,t}\right]E\left[x_{T,t'}y_{T,t'}\right]\right) \\ = \; & T^{-2}\sum_{t=1}^{[T\lambda]}\sum_{t'=1}^{[T\lambda]}\left(\begin{array}{c} E\left[x_{T,t}x_{T,t'}\right]E\left[y_{T,t}y_{T,t'}\right] - E\left[x_{T,t}y_{T,t'}\right]E\left[x_{T,t'}y_{T,t}\right] \\ +\mathrm{cum}\left(x_{T,t}, y_{T,t}, x_{T,t'}, y_{T,t'}\right)\end{array}\right). \end{aligned}$$

Then

$$E\left[x_{T,t}y_{T,t'}\right] = \sum_{j=0}^{t}\sum_{j'=0}^{t'} r_{T,j}s_{T,j'}E\left[\varepsilon_{t-j}\varepsilon_{t'-j'}\right] = \sigma^2\sum_{j=0}^{t\wedge t'} r_{T,j}s_{T,|t-t'|+j}$$

and therefore, uniformly in T,

$$\left|E\left[x_{T,t}y_{T,t'}\right]\right| \leq K\sum_{j=0}^{t\wedge t'}|j|_+^{\varsigma-1}\left(|t-t'|_+ + j\right)^{\varsigma-1} \leq K\left(t\wedge t'\right)^{\varsigma}|t-t'|_+^{\varsigma-1},$$

[1]Corresponding author: Carlos Velasco, Department of Economics, Universidad Carlos III de Madrid, Getafe, 28903 Madrid, Spain; E-Mail: carlos.velasco@uc3m.es

where we assume w.l.o.g. $\varsigma > 0$ (because if $\varsigma \leq 0$ we can replace it by an arbitrarily small positive number) and $K > 0$ is a constant that may change from line to line, while by independence of ε_t,

$$\begin{aligned}
\text{cum}\left(x_{T,t}, y_{T,t}, x_{T,t'}, y_{T,t'}\right) &= \sum_{j=0}^{t}\sum_{j'=0}^{t'}\sum_{k=0}^{t}\sum_{k'=0}^{t'} r_{T,j}s_{T,j'}r_{T,k}s_{T,k'}\text{cum}\left(\varepsilon_{t-j}, \varepsilon_{t'-j'}, \varepsilon_{t-k}, \varepsilon_{t'-k'}\right) \\
&= \kappa_4 \sum_{j=0}^{t\wedge t'} r_{T,j}^2 s_{T,|t-t'|+j}^2 = O\left(\sum_{j=0}^{t\wedge t'} |j|_+^{2\varsigma-2}\left(|t-t'|_+ + j\right)^{2\varsigma-2}\right) \\
&= O\left(|t-t'|_+^{2\varsigma-2}\right)
\end{aligned}$$

for $\varsigma < 0.5$, so $\sum_{j=0}^{t\wedge t'} |j|_+^{2\varsigma-2} = O(1)$, where κ_4 is the fourth order cumulant of ε_t and therefore, compiling the two type of contributions,

$$Var\left[z_T(\lambda)\right] = O\left(T^{2\varsigma-2}\sum_{t=1}^{T}\sum_{t'=1}^{T}|t-t'|_+^{2\varsigma-2}\right) = O\left(T^{2\varsigma-1}\right) = o(1).$$

To show uniformity in λ we show weak convergence of $z_T(\lambda)$ to the zero function in $\mathcal{D}([0,1])$, the space of càdlàg functions on $[0,1]$ with Skorokhod J_1-metric, by showing that $z_T(\lambda)$ is tight because for all T large enough and $0 \leq \lambda_1 < \lambda_2 \leq 1$,

$$E\left[|z_T(\lambda_2) - z_T(\lambda_1)|^\gamma\right] \leq C\,|\lambda_2 - \lambda_1|^\alpha \tag{1}$$

for some constants $\gamma > 0$, $\alpha > 1$ and $C > 0$ independent of T, Billingsley (1968). Note that we can consider only $|\lambda_2 - \lambda_1| \geq T^{-1}$ because otherwise $z_T(\lambda_2) - z_T(\lambda_1) = 0$.

First, for $\gamma = 2$,

$$E\left[\left(z_T(\lambda_2) - z_T(\lambda_1)\right)^2\right] = T^{-2}\sum_{t=1+[T\lambda_1]}^{[T\lambda_2]}\sum_{t'=1+[T\lambda_1]}^{[T\lambda_2]}\left(E\left[x_{T,t}y_{T,t}x_{T,t'}y_{T,t'}\right] - E\left[x_{T,t}y_{T,t}\right]E\left[x_{T,t'}y_{T,t'}\right]\right)$$

and therefore, for $\varsigma < 0.5$,

$$\begin{aligned}
E\left[\left(z_T(\lambda_2) - z_T(\lambda_1)\right)^2\right] &\leq KT^{2\varsigma-2}\sum_{t=1+[T\lambda_1]}^{[T\lambda_2]}\sum_{t'=1+[T\lambda_1]}^{[T\lambda_2]}|t-t'|_+^{2\varsigma-2} \\
&\leq KT^{2\varsigma-2}\sum_{j=0}^{[T(\lambda_2-\lambda_1)]}\left(T(\lambda_2-\lambda_1) - j\right)|j|_+^{2\varsigma-2} \\
&\leq KT^{2\varsigma-1}(\lambda_2-\lambda_1)\sum_{j=0}^{[T(\lambda_2-\lambda_1)]}|j|_+^{2\varsigma-2} \\
&\leq K(\lambda_2-\lambda_1)^{2-2\varsigma}
\end{aligned}$$

because $2\varsigma - 2 < 0$, so (1) holds with $\alpha = 2 - 2\varsigma > 1$, and the lemma follows. $\square$

Lemma A.2. Let $x_{T,t} = r_{T,t}(L)\,\varepsilon_t$ where $r_{T,t}(L) = \sum_{j=0}^{t} r_{T,j}L^j$ denotes a filter, possibly depending on T, with $|r_{T,j}| \leq K\,|j|_+^{\varsigma-1}$ as $T \to \infty$ for $\varsigma < 0.5$, where $|j|_+ = \max\{|j|, 1\}$, and ε_r is $iid\left(0, \sigma^2\right)$ with finite fourth moment. Let $\xi_t \sim Ct^{-\delta}$ as $t \to \infty$ for $\delta \in (-0.5, 0.5)$. Then

$$z_T(\lambda) = T^{\delta-1}\sum_{t=1}^{[T\lambda]}\xi_t x_{T,t} \to_p 0$$

uniformly for $\lambda \in \Lambda = [\epsilon, 1-\epsilon], \epsilon > 0$.

Proof. As in Lemma A.1, first we show that for every λ, $z_T(\lambda) \to_p 0$ and therefore $z_T(\lambda) \to_d 0$, since $E[z_T(\lambda)] = 0$ and $Var[z_T(\lambda)]$ is equal to

$$T^{2\delta-2} \sum_{t=1}^{[T\lambda]} \sum_{t'=1}^{[T\lambda]} \xi_t \xi_{t'} E[x_{T,t} x_{T,t'}] = \sigma^2 T^{2\delta-2} \sum_{t=1}^{[T\lambda]} \sum_{t'=1}^{[T\lambda]} \sum_{j=0}^{t \wedge t'} \xi_t \xi_{t'} r_{T,j} r_{T,|t-t'|+j},$$

so that, assuming $\varsigma > 0$ w.l.o.g., if $\delta < 0$,

$$\begin{aligned} Var[z_T(\lambda)] &\leq KT^{\varsigma+2\delta-2} \sum_{t=1}^{[T\lambda]} \sum_{t'=1}^{[T\lambda]} |\xi_t| |\xi_{t'}| |t-t'|_+^{\varsigma-1} \\ &\leq KT^{\varsigma+2\delta-2} \sum_{t=1}^{[T\lambda]} \sum_{t'=1}^{[T\lambda]} t^{-\delta} t'^{-\delta} |t-t'|_+^{\varsigma-1} \\ &\leq KT^{\varsigma-2} \sum_{t=1}^{[T\lambda]} \sum_{t'=1}^{[T\lambda]} |t-t'|_+^{\varsigma-1} \leq KT^{2\varsigma-1} \end{aligned}$$

while if $\delta \geq 0$

$$\begin{aligned} Var[z_T(\lambda)] &\leq KT^{\varsigma+2\delta-2} \sum_{t=1}^{[T\lambda]} t^{-2\delta} \sum_{t'=t}^{[T\lambda]} |t-t'|_+^{\varsigma-1} \\ &\leq KT^{\varsigma+2\delta-2} T^{1-2\delta} T^{\varsigma} \leq KT^{2\varsigma-1} \end{aligned}$$

because $-2\delta > -1$ and $\varsigma - 1 > -1$, so $Var[z_T(\lambda)] \to 0$ as $T \to \infty$ because $\varsigma < 0.5$.

To show uniformity in λ we show again weak convergence of $z_T(\lambda)$ to the zero function in $\mathcal{D}(\Lambda)$, by showing that $z_T(\lambda)$ is tight because for $0 < \epsilon \leq \lambda_1 < \lambda_2 \leq 1-\epsilon < 1$,

$$\begin{aligned} E\left[(z_T(\lambda_2) - z_T(\lambda_1))^2\right] &= T^{2\delta-2} \sum_{t=1+[T\lambda_1]}^{[T\lambda_2]} \sum_{t'=1+[T\lambda_1]}^{[T\lambda_2]} \xi_t \xi_{t'} E[x_{T,t} x_{T,t'}] \\ &\leq KT^{\varsigma-2} \sum_{t=1+[T\lambda_1]}^{[T\lambda_2]} \sum_{t'=1+[T\lambda_1]}^{[T\lambda_2]} |t/T|^{-\delta} |t'/T|^{-\delta} |t-t'|_+^{\varsigma-1} \\ &\leq KT^{\varsigma-1} |\lambda_2 - \lambda_1| \sum_{j=0}^{[T(\lambda_2-\lambda_1)]} |j|_+^{\varsigma-1} \\ &\leq K |\lambda_2 - \lambda_1|^2 \end{aligned}$$

because $\max_{[T\lambda_1] < t \leq [T\lambda_2]} |t/T|^{-\delta} \leq K$, so (1) holds for $\gamma = 2$ and $\alpha = 2$, and the lemma follows. □

Lemma A.3. Let $x_{T,t} = r_{T,t}(L) \varepsilon_t$ where $r_{T,t}(L) = \sum_{j=0}^{t} r_{T,j} L^j$ denotes a filter, possibly depending on T, with $r_{T,0} = 0$ and $|r_{T,j}| \leq K |j|^{\varsigma-1}$ as $T \to \infty$ for $\varsigma < 0.5$, and ε_r is $iid\left(0, \sigma^2\right)$ with finite eight moments. Then

$$z_T(\lambda) = T^{-1/2} \sum_{t=1}^{[T\lambda]} x_{T,t} \varepsilon_t$$

is tight in $\mathcal{D}([0,1])$.

Proof. We check condition (1) for $\gamma = 4$. For that, and denoting by μ_3 and μ_4 the third and fourth moments of ε_t, respectively,

$$\begin{aligned} E\left[\left(z_T(\lambda_2) - z_T(\lambda_1)\right)^4\right] &= T^{-2} \sum_{t_1=1+[T\lambda_1]}^{[T\lambda_2]} \sum_{t_2=1+[T\lambda_1]}^{[T\lambda_2]} \sum_{t_3=1+[T\lambda_1]}^{[T\lambda_2]} \sum_{t_4=1+[T\lambda_1]}^{[T\lambda_2]} E\left[x_{T,t_1}\varepsilon_{t_1} x_{T,t_2}\varepsilon_{t_2} x_{T,t_3}\varepsilon_{t_3} x_{T,t_4}\varepsilon_{t_4}\right] \\ &= 6\sigma^2 T^{-2} \sum_{t_1=1+[T\lambda_1]}^{[T\lambda_2]} \sum_{t_2=1+[T\lambda_1]}^{[T\lambda_2]} \sum_{t_3=1+[T\lambda_1]}^{[T\lambda_2]} E\left[x_{T,t_1}\varepsilon_{t_1} x_{T,t_2}\varepsilon_{t_2} x_{T,t_3}^2\right] 1_{\{t_3 > \max(t_1,t_2)\}} \quad (2) \\ &\quad + 4\mu_3 T^{-2} \sum_{t_1=1+[T\lambda_1]}^{[T\lambda_2]} \sum_{t_2=1+[T\lambda_1]}^{[T\lambda_2]} E\left[x_{T,t_1}\varepsilon_{t_1} x_{T,t_2}^3\right] 1_{\{t_2 > t_1\}} \quad (3) \\ &\quad + \mu_4 T^{-2} \sum_{t=1+[T\lambda_1]}^{[T\lambda_2]} E\left[x_{T,t}^4\right] \quad (4) \end{aligned}$$

noting that, since $x_{T,t}\varepsilon_t$ is a martingale difference (as $x_{T,t}$ and ε_t are independent and zero mean), the product moment $E\left[x_{T,t_1}\varepsilon_{t_1} x_{T,t_2}\varepsilon_{t_2} x_{T,t_3}\varepsilon_{t_3} x_{T,t_4}\varepsilon_{t_4}\right]$ is only different from zero when the largest two values in $\{t_1, t_2, t_3, t_4\}$ are equal.

Then, with $t_3 = t_4$, (2) is equal to

$$6\sigma^2 T^{-2} \left(\prod_{i=1}^{3} \sum_{t_i=1+[T\lambda_1]}^{[T\lambda_2]}\right) \left(\prod_{i=1}^{4} \sum_{j_i=1}^{t_i} r_{T,j_i}\right) E\left[\varepsilon_{t_1-j_1}\varepsilon_{t_1}\varepsilon_{t_2-j_2}\varepsilon_{t_2}\varepsilon_{t_3-j_3}\varepsilon_{t_3-j_4}\right] 1_{\{t_3 > \max(t_1,t_2)\}},$$

where we evaluate the expectation, using the formula that relates the expectation of a product of random variables to their cumulants,

$$E\left[X_1 \cdots X_n\right] = \sum_{\pi} \prod_{B \in \pi} \text{cum}(X_i : i \in B) \quad (5)$$

where the summation runs over all partitions π of $\{1, \ldots, n\}$ and $\text{cum}(\cdot)$ denotes joint cumulant.

Given $E[\varepsilon_t] = 0$, then the expression (5) for $E\left[\varepsilon_{t_1-j_1}\varepsilon_{t_1}\varepsilon_{t_2-j_2}\varepsilon_{t_2}\varepsilon_{t_3-j_3}\varepsilon_{t_3-j_4}\right]$ involves only partitions π for which the cardinality $|B|$ of the index sets B involved for each π are $(6), (4,2), (3,3), (2,2,2)$, and, given independence of ε_t, all $i \in B$ have to be equal for each set B for $\kappa(X_i : i \in B) = \kappa_{|B|}$ to be (possibly) different from zero, where κ_i is the i-th order marginal cumulant of ε_t. This fact imposes at least three restrictions among the indexes t_i and j_i to obtain non-zero contributions to the expectation, some of them non-feasible because $j_i > 0$, so, for instance, $\kappa\left(\varepsilon_{t_1-j_1}, \varepsilon_{t_1}\right) = E\left[\varepsilon_{t_1-j_1}\varepsilon_{t_1}\right] = 0$ because $\varepsilon_{t_1-j_1}$ and ε_{t_1} are independent. There are two typical combinations of restrictions involved:

Case a: One restriction $j_i = j_{i'} + t_i - t_{i'}$ for some $i = 1, 2, i' = 1, \ldots, 4, i \neq i'$; and two restrictions $j_k = t_k - t_{k'}$ for $k' = 1, 2$ and some $k \in \{1, \ldots, 4\} / \{i, i', k'\}$.

Case b: Two restrictions $j_i = j_{i'} + t_i - t_{i'}$ for some $i, i' = 1, \ldots, 4$, $i \neq i'$; and the restriction $t_1 = t_2$.

Then, assuming $\varsigma > 0$ w.l.o.g., we can bound the contribution to (2) of the Case a terms of (5) by

$$\begin{aligned} & KT^{-2} \left(\prod_{i=1}^{3} \sum_{t_i=1+[T\lambda_1]}^{[T\lambda_2]}\right) \sum_{j_4=1}^{t_3=t_4} \left|r_{T,j_4+|t_4-t_1|}\right| \left|r_{T,|t_2-t_1|}\right| \left|r_{T,|t_2-t_3|}\right| \left|r_{T,j_4}\right| \\ \leq\ & KT^{-2} \left(\prod_{i=1}^{3} \sum_{t_i=1+[T\lambda_1]}^{[T\lambda_2]}\right) \left|r_{T,|t_2-t_1|}\right| \left|r_{T,|t_2-t_3|}\right| \sum_{j_4=1}^{t_3=t_4} \left|r_{T,j_4}\right|^2 \\ \leq\ & KT^{-2} \sum_{t_1=1+[T\lambda_1]}^{[T\lambda_2]} \sum_{t_2=1+[T\lambda_1]}^{[T\lambda_2]} |t_2 - t_1|_+^{\varsigma-1} \sum_{t_3=1+[T\lambda_1]}^{[T\lambda_2]} |t_2 - t_3|_+^{\varsigma-1} \\ \leq\ & KT^{-2} \left(T|\lambda_2 - \lambda_1|\right)^{1+2\varsigma} \leq KT^{2\varsigma-1} |\lambda_2 - \lambda_1|^{1+2\varsigma} \quad (6) \end{aligned}$$

because $\varsigma < 0.5$, while the Case b terms, excluding the combinations $(1,2)$ and $(3,4)$, can be bounded by

$$\begin{aligned} & KT^{-2}\left(\prod_{i=2}^{3}\sum_{t_i=1+[T\lambda_1]}^{[T\lambda_2]}\right)\sum_{j_2=1}^{t_2}\sum_{j_4=1}^{t_3=t_4}\left|r_{T,j_2}\right|\left|r_{T,j_4}\right|\left|r_{T,j_4+|t_1-t_3|}\right|\left|r_{T,j_2+|t_2-t_3|}\right| \\ \leq\ & KT^{-2}\left(\prod_{i=2}^{3}\sum_{t_i=1+[T\lambda_1]}^{[T\lambda_2]}\right)\sup_{j_4}\left|r_{T,j_4+|t_1-t_3|}\right|\sup_{j_2}\left|r_{T,j_2+|t_2-t_3|}\right|\sum_{j_2=1}^{t_2}\sum_{j_4=1}^{t_3=t_4}\left|r_{T,j_2}\right|\left|r_{T,j_4}\right| \\ \leq\ & KT^{-2}\sum_{t_3=1+[T\lambda_1]}^{[T\lambda_2]}\left|t_1-t_3\right|_+^{\varsigma-1}\sum_{t_2=1+[T\lambda_1]}^{[T\lambda_2]}\left|t_2-t_3\right|_+^{\varsigma-1}T^{2\varsigma} \\ \leq\ & KT^{2\varsigma-2}\left(T\left|\lambda_2-\lambda_1\right|\right)^{2\varsigma}\leq KT^{4\varsigma-2}\left|\lambda_2-\lambda_1\right|^{2\varsigma}, \end{aligned} \tag{7}$$

while the case with $j_1 = j_2 + t_1 - t_2 = j_2$ and $j_3 = j_4 + t_3 - t_4 = j_4$ is bounded using $\varsigma < 0.5$ by

$$\begin{aligned} & KT^{-2}\left(\prod_{i=2}^{3}\sum_{t_i=1+[T\lambda_1]}^{[T\lambda_2]}\right)\sum_{j_2=1}^{t_2}\sum_{j_4=1}^{t_3=t_4}\left|r_{T,j_2}\right|^2\left|r_{T,j_4}\right|^2 \\ \leq\ & KT^{-2}\sum_{t_3=1+[T\lambda_1]}^{[T\lambda_2]}\sum_{t_2=1+[T\lambda_1]}^{[T\lambda_2]}1 \\ \leq\ & K\left|\lambda_2-\lambda_1\right|^2. \end{aligned} \tag{8}$$

Since we can consider only $|\lambda_2 - \lambda_1| \geq T^{-1}$ because otherwise $z_T(\lambda_2) - z_T(\lambda_1) = 0$, we have that (6), (7) and (8) are bounded by $K|\lambda_2 - \lambda_1|^{2-2\varsigma}$, as are the contributions in the presence of additional restrictions and from (3) and (4) following a similar, but simpler argument. Then (1) holds for $\alpha = 2 - 2\varsigma > 1$ because $\varsigma < 0.5$, and the lemma follows. □

Proof of Theorem 1

The consistency and rate of convergence of the restricted estimators $\left(\tilde{d}_{0T}, \tilde{\boldsymbol{\alpha}}_{0T}', \tilde{\mu}_{0T}\right)$ under the null of no break follow from Hualde and Nielsen (2017),

$$\begin{aligned} \tilde{d}_{0T} - d_0 &= O_p(T^{-1/2}), \\ \tilde{\mu}_{0T} - \mu_0 &= O_p(T^{d_0-1/2}), \\ \tilde{\boldsymbol{\alpha}}_{0T} - \boldsymbol{\alpha}_0 &= O_p(T^{-1/2}), \end{aligned}$$

and the same result can be shown under $H_{1,T}(\lambda_0)$ with minor modifications of their arguments in the present situation, which is simpler because the order of the deterministic trend is assumed known, though asymptotic distributions are affected by a local drift.

Denoting $\boldsymbol{\tau} = (d, \boldsymbol{\alpha}')'$, $\boldsymbol{\tau}_0 = (d_0, \boldsymbol{\alpha}_0')'$ and $\tilde{\boldsymbol{\tau}}_{0T} = \left(\tilde{d}_{0T}, \tilde{\boldsymbol{\alpha}}_{0T}'\right)'$ and, after concentrating out μ_0, we can write

$$\tilde{\boldsymbol{\tau}}_{0T} - \boldsymbol{\tau}_0 = -\left.\frac{\partial^2 \mathcal{L}_T^*(\boldsymbol{\tau})}{\partial\boldsymbol{\tau}\partial\boldsymbol{\tau}'}\right|_{\boldsymbol{\tau}=\bar{\boldsymbol{\tau}}_{0T}} \left.\frac{\partial \mathcal{L}_T^*(\boldsymbol{\tau})}{\partial\boldsymbol{\tau}}\right|_{\boldsymbol{\tau}=\boldsymbol{\tau}_0}$$

where $\bar{\boldsymbol{\tau}}_{0T}$ is some intermediate point $\|\bar{\boldsymbol{\tau}}_{0T} - \boldsymbol{\tau}_0\| < \|\tilde{\boldsymbol{\tau}}_{0T} - \boldsymbol{\tau}_0\|$ that may change from row to row of $\frac{\partial^2 \mathcal{L}_T^*(\boldsymbol{\tau})}{\partial\boldsymbol{\tau}\partial\boldsymbol{\tau}'}$, with

$$\mathcal{L}_T^*(\boldsymbol{\tau}) = -\frac{1}{2}\sum_{t=1}^{T}\varepsilon_t^*(\boldsymbol{\tau})^2, \tag{9}$$

and $\varepsilon_t^*(\boldsymbol{\tau}) = \alpha(L)\Delta_t^d(y_t - \tilde{\mu}_{0T}(\boldsymbol{\tau}))$, and

$$\tilde{\mu}_{0T}(\boldsymbol{\tau}) = \left(\sum_{t=1}^{T} m_t(\boldsymbol{\tau})^2\right)^{-1} \sum_{t=1}^{T} m_t(\boldsymbol{\tau})\,\alpha(L)\,\Delta_t^d y_t,$$

with $m_t(\boldsymbol{\tau}) = \alpha(L)\Delta_t^d 1$. Then

$$\begin{aligned} \left.\frac{\partial \mathcal{L}_T^*(\boldsymbol{\tau})}{\partial \boldsymbol{\tau}}\right|_{\boldsymbol{\tau}=\boldsymbol{\tau}_0} &= -\sum_{t=1}^{T} \left.\frac{\partial \varepsilon_t^*(\boldsymbol{\tau})}{\partial \boldsymbol{\tau}}\right|_{\boldsymbol{\tau}=\boldsymbol{\tau}_0} \varepsilon_t^*(\boldsymbol{\tau}_0) = -\sum_{t=1}^{T} \varepsilon_t^0 \sum_{j=1}^{t} \mathbf{s}_j(\boldsymbol{\tau}_0)\,\varepsilon_{t-j}^0 + o_p\left(T^{-1/2}\right) \\ &= -\sum_{t=1}^{T} \left(\mathbf{s}_t(L;\boldsymbol{\tau}_0)\,\varepsilon_t^0\right)\varepsilon_t^0 + o_p\left(T^{-1/2}\right), \end{aligned}$$

where $\varepsilon_t^0 = \varepsilon_t(\psi_0) = \varepsilon_t(0,\mathbf{0},0,d_0,\boldsymbol{\alpha}_0,\mu_0) = \alpha_0(L)\,\Delta_t^{d_0}(y_t - \mu_0)$ and $\mathbf{s}_t(L;\boldsymbol{\tau}_0) = (\partial/\partial\boldsymbol{\tau})\log\Delta^d\alpha(L)\big|_{\boldsymbol{\tau}=\boldsymbol{\tau}_0} = \sum_{j=1}^{t}\mathbf{s}_j(\boldsymbol{\tau}_0)\,L^j = \left(\log\Delta_t, -c(L;\boldsymbol{\tau}_0)(L,\ldots,L^p)'\right)$ where $\mathbf{s}_j(\boldsymbol{\tau}_0) = -\left(j^{-1}, c_{j-1}(\boldsymbol{\tau}_0),\ldots,c_{j-2+p}(\boldsymbol{\tau}_0)\right)'$ so that $c(L;\boldsymbol{\tau}) = 1/\alpha(L) = \sum_{j=0}^{\infty} c_j(\boldsymbol{\tau})\,L^j$.

Then

$$\begin{aligned} \varepsilon_t^0 &= \left(\alpha_0(L) \pm R_t^{(2)}(\lambda_0)\,\beta_0(L)\right)\Delta_t^{d_0 \pm \theta_0 R_t^{(2)}(\lambda_0)}\left(y_t - \mu_0 \pm \nu_0 R_t^{(2)}(\lambda_0)\right) \qquad (10) \\ &= \Delta_t^{-\theta_0 R_t^{(2)}(\lambda_0)}\varepsilon_t - R_t^{(2)}(\lambda_0)\,\beta_0(L)\left[\Delta_t^{d_0}(y_t-\mu_0)\right] + \nu_0\alpha_0(L)\,\Delta_t^{d_0}R_t^{(2)}(\lambda_0) \\ &= \varepsilon_t + T^{-1/2}R_t^{(2)}(\lambda_0)\left[-\delta\log\Delta_t\varepsilon_t - \sum\nolimits_{j=1}^{p}\gamma_j\alpha_0^{-1}(L)\,\varepsilon_{t-j} + \eta T^{d_0}\alpha_0(L)\,\Delta_{t-[\lambda_0 T]}^{d_0}1\right] \\ &\quad + T^{-1}\mathbf{r}_{T,t}(L;\boldsymbol{\tau}_0)\,\varepsilon_t + O\left(T^{d_0-1}\pi_{t-[\lambda_0 T]-1}(d_0-1)\right), \end{aligned}$$

where $\mathbf{r}_{T,t}(L;\boldsymbol{\tau}_0) = \sum_{j=1}^{t}\mathbf{r}_{T,j}(\boldsymbol{\tau}_0)\,L^j$ denotes a filter that might change from case to case, possibly depending on T, with $\|\mathbf{r}_{T,j}(\boldsymbol{\tau}_0)\| \le Kj^{\varsigma-1}$ as $T\to\infty$ for any $\varsigma > 0$.

Therefore, $\mathbf{s}_t(L;\boldsymbol{\tau}_0)\,\varepsilon_t^0$ is equal to

$$\mathbf{s}_t(L;\boldsymbol{\tau}_0)\,\varepsilon_t + T^{-1/2}\mathbf{r}_{T,t}(L;\boldsymbol{\tau}_0)\,\varepsilon_t + \boldsymbol{\zeta}_{T,t},$$

where $\left\|\boldsymbol{\zeta}_{T,t}\right\| \le KT^{d_0-1/2}\log T$ as $T\to\infty$, and it follows that

$$T^{-1/2}\left.\frac{\partial \mathcal{L}_T^*(\boldsymbol{\tau})}{\partial \boldsymbol{\tau}}\right|_{\boldsymbol{\tau}=\boldsymbol{\tau}_0} \to_p \sigma_0^2(1-\lambda_0)\,\boldsymbol{\Xi}\,(\delta,\gamma')' - T^{-1/2}\sum_{t=1}^{T}\left(\mathbf{s}_t(L;\boldsymbol{\tau}_0)\,\varepsilon_t\right)\varepsilon_t. \qquad (11)$$

Since a similar reasoning leads to,

$$\frac{1}{T}\left.\frac{\partial^2 \mathcal{L}_T^*(\boldsymbol{\tau})}{\partial \boldsymbol{\tau}\partial \boldsymbol{\tau}'}\right|_{\boldsymbol{\tau}=\bar{\boldsymbol{\tau}}_{0T}} \to_p \sigma_0^2\boldsymbol{\Xi},$$

we find that

$$T^{1/2}\left(\tilde{\boldsymbol{\tau}}_{0T} - \boldsymbol{\tau}_0\right) \to_p (1-\lambda_0)(\delta,\gamma')' - T^{-1/2}\left(\sigma_0^2\boldsymbol{\Xi}\right)^{-1}\sum_{t=1}^{T}\left(\mathbf{s}_t(L;\boldsymbol{\tau}_0)\,\varepsilon_t\right)\varepsilon_t. \qquad (12)$$

Likewise, using the same arguments on $\tilde{\mu}_{0T} = \tilde{\mu}_{0T}(\tilde{\boldsymbol{\tau}}_{0T})$, under H_{1T},

$$\begin{aligned} T^{1/2-d_0}\left(\tilde{\mu}_{0T} - \mu_0\right) &= T^{1/2-d_0}\left(\sum_{t=1}^{T} m_t(\boldsymbol{\tau}_0)^2\right)^{-1}\sum_{t=1}^{T} m_t(\boldsymbol{\tau}_0)\,\varepsilon_t^0 + o_p(1) \\ &= \eta\left(\sum_{t=1}^{T} m_t(\boldsymbol{\tau}_0)^2\right)^{-1}\sum_{t=[\lambda_0 T]+1}^{T} m_t(\boldsymbol{\tau}_0)\left(\alpha_0(L)\,\Delta_{t-[\lambda_0 T]}^{d_0}1\right) \\ &\quad + T^{1/2-d_0}\left(\sum_{t=1}^{T} m_t(\boldsymbol{\tau}_0)^2\right)^{-1}\sum_{t=1}^{T} m_t(\boldsymbol{\tau}_0)\,\varepsilon_t + o_p(1). \end{aligned}$$

Writing $\mathcal{L}^0_{\theta,T}(\lambda) = \mathcal{L}_{\theta,T}\left(0, \mathbf{0}, 0, d_0, \boldsymbol{\alpha}_0, \mu_0, \sigma^2, \lambda\right)$ and $\tilde{\varepsilon}_t = \varepsilon_t(\tilde{\psi}_T)$,

$$\tilde{\varepsilon}_t = \varepsilon^0_t + (\tilde{d}_{0T} - d_0)\log\Delta_t\varepsilon^*_t + (\tilde{\alpha}_{0T}(L) - \alpha_0(L))\frac{\varepsilon^*_t}{\alpha^*_{0T}(L)} - (\tilde{\mu}_{0T} - \mu_0)\,\alpha^*_{0T}(L)\,\Delta^{d^*_{0T}}_t 1, \tag{13}$$

where $\varepsilon^*_t = \alpha^*_{0T}(L)\,\Delta^{d^*_{0T}}_t(y_t - \mu^*_{0T})$ with $\|\psi^*_T - \psi_0\| \leq \left\|\tilde{\psi}_T - \psi_0\right\|$, $\psi^*_{0T} = (d^*_{0T}, \boldsymbol{\alpha}^{*\prime}_{0T}, \mu^*_{0T})'$ yields

$$\begin{aligned}
\tilde{\mathcal{L}}_{\theta,T}(\lambda) &= \mathcal{L}^0_{\theta,T}(\lambda) \\
&\quad -(\tilde{d}_{0T} - d_0)\frac{1}{\tilde{\sigma}^2_T}\sum_{t=[\lambda T]+1}^{T}\left\{\left((\log\Delta_t)^2\varepsilon^*_t\right)\varepsilon^*_t + (\log\Delta_t\varepsilon^*_t)^2\right\} \qquad (14)\\
&\quad +\sum_{j=1}^{p}(\tilde{\alpha}_{j,0T} - \alpha_{j,0T})\frac{1}{\tilde{\sigma}^2_T}\sum_{t=[\lambda T]+1}^{T}\left\{(\log\Delta_t\varepsilon^*_t)\frac{\varepsilon^*_{t-j}}{\alpha^*_{0T}(L)} + \left(\log\Delta_{t-j}\frac{\varepsilon^*_{t-j}}{\alpha^*_{0T}(L)}\right)\varepsilon^*_t\right\} \\
&\quad +(\tilde{\mu}_{0T} - \mu_0)\frac{1}{\tilde{\sigma}^2_T}\sum_{t=[\lambda T]+1}^{T}\left\{\left(\log\Delta_t\alpha^*_{0T}(L)\,\Delta^{d^*}_{t-[\lambda_0 T]}1\right)\varepsilon^*_t + (\log\Delta_t\varepsilon^*_t)\,\alpha^*_{0T}(L)\,\Delta^{d^*}_{t-[\lambda_0 T]}1\right\}.
\end{aligned}$$

Then, uniformly in λ and under H_{1T}, it holds that

$$\begin{aligned}
&T^{-1}\sum_{t=[\lambda T]+1}^{T}\left((\log\Delta_t)^2\varepsilon^*_t\right)\varepsilon^*_t \xrightarrow{p} 0 \\
&T^{-1}\sum_{t=[\lambda T]+1}^{T}(\log\Delta_t\varepsilon^*_t)^2 \xrightarrow{p} (1-\lambda)\,\sigma^2_0\frac{\pi^2}{6} \\
&T^{-1}\sum_{t=[\lambda T]+1}^{T}\left\{(\log\Delta_t\varepsilon^*_t)\frac{\varepsilon^*_{t-j}}{\alpha^*(L)} + \left(\log\Delta_{t-j}\frac{\varepsilon^*_{t-j}}{\alpha^*(L)}\right)\varepsilon^*_t\right\} \xrightarrow{p} (1-\lambda)\,\sigma^2_0\kappa_j \\
&T^{d_0-1}\sum_{t=[\lambda T]+1}^{T}\left(\log\Delta_t\alpha^*(L)\,\Delta^{d_0}_{t-[\lambda_0 T]}1\right)\varepsilon^*_t \xrightarrow{p} 0 \\
&T^{d_0-1}\sum_{t=[\lambda T]+1}^{T}(\log\Delta_t\varepsilon^*_t)\,\alpha^*(L)\,\Delta^{d_0}_{t-[\lambda_0 T]}1 \xrightarrow{p} 0
\end{aligned}$$

using Lemma D.5 of Robinson and Hualde (2003) to Taylor expand the functions of ψ^*_{0T} around ψ_0 until an order high enough and bound uniformly in λ the contribution of each term using Lemmas A.1 and A.2. Then, uniformly in λ,

$$\tilde{\mathcal{L}}_{\theta,T}(\lambda) = \mathcal{L}^0_{\theta,T}(\lambda) - T(1-\lambda)\left(\frac{\pi^2}{6} \quad \boldsymbol{\kappa}'\right)(\tilde{\boldsymbol{\tau}}_{0,T} - \boldsymbol{\tau}_0) + o_p\left(T^{1/2}\right)$$

Similarly, denoting $\mathbf{c}^*_p(L) = c(L;\boldsymbol{\tau}^*_{0T})(L,\ldots,L^p)'$,

$$\begin{aligned}
\tilde{\mathcal{L}}_{\beta,T}(\lambda) &= \mathcal{L}^0_{\beta,T}(\lambda) \qquad (15)\\
&\quad -(\tilde{d}_{0T} - d_0)\frac{1}{\tilde{\sigma}^2_T}\sum_{t=[\lambda T]+1}^{T}\left\{\left(\log\Delta_t\mathbf{c}^*_p(L)\,\varepsilon^*_t\right)\varepsilon^*_t + (\log\Delta_t\varepsilon^*_t)\,\mathbf{c}^*_p(L)\,\varepsilon^*_t\right\} \\
&\quad -\frac{1}{\tilde{\sigma}^2_T}\sum_{t=[\lambda T]+1}^{T}\left(\mathbf{c}^*_p(L)\,\varepsilon^*_t\right)'\left(\mathbf{c}^*_p(L)\,\varepsilon^*_t\right)(\tilde{\alpha}_{0T} - \alpha_{0T}) \\
&\quad +(\tilde{\mu}_{0T} - \mu_0)\frac{1}{\tilde{\sigma}^2_T}\sum_{t=[\lambda T]+1}^{T}\left\{\left(\mathbf{c}^*_p(L)\,\alpha^*(L)\,\Delta^{d_0}_{t-[\lambda_0 T]}1\right)\varepsilon^*_t + \left(\mathbf{c}^*_p(L)\,\varepsilon^*_t\right)\alpha^*(L)\,\Delta^{d_0}_{t-[\lambda_0 T]}1\right\} \\
&= \mathcal{L}^0_{\beta,T}(\lambda) - T(1-\lambda)\left(\boldsymbol{\kappa} \quad \boldsymbol{\Phi}\right)(\tilde{\boldsymbol{\tau}}_{0,T} - \boldsymbol{\tau}_0) + o_p\left(T^{1/2}\right)
\end{aligned}$$

uniformly in λ, and therefore

$$\begin{aligned}\tilde{\mathcal{L}}_{(\theta,\beta),T}(\lambda) &= \mathcal{L}^0_{(\theta,\beta),T}(\lambda) - T(1-\lambda)\,\boldsymbol{\Xi}\left(\tilde{\boldsymbol{\tau}}_{0,T}-\boldsymbol{\tau}_0\right) + o_p\left(T^{1/2}\right) \\ &= \mathcal{L}^0_{(\theta,\beta),T}(\lambda) - T^{1/2}(1-\lambda)(1-\lambda_0)\,\boldsymbol{\Xi}\left(\delta,\gamma'\right)' + (1-\lambda)\frac{1}{\sigma_0^2}\sum_{t=1}^{T}\left(\mathbf{s}_t\left(L;\boldsymbol{\tau}_0\right)\varepsilon_t\right)\varepsilon_t + o_p\left(T^{1/2}\right) \quad (16)\end{aligned}$$

Next, using similar arguments as for (11) and Lemmas A.1 and A.2,

$$\mathcal{L}^0_{(\theta,\beta),T}(\lambda) = T^{1/2}\left(1-\max\{\lambda,\lambda_0\}\right)\boldsymbol{\Xi}\left(\delta,\gamma'\right)' - \frac{1}{\sigma_0^2}\sum_{t=[\lambda T]+1}^{T}\left(\mathbf{s}_t\left(L;\boldsymbol{\tau}_0\right)\varepsilon_t\right)\varepsilon_t + o_p\left(T^{1/2}\right) \quad (17)$$

uniformly in λ under H_{1T}, and therefore, using (16) and (17), $\tilde{\mathcal{L}}_{(\theta,\beta),T}(\lambda)$ is equal to

$$\left(\lambda(1-\lambda_0)-(\lambda-\lambda_0)_+\right)\boldsymbol{\Xi}\left(\delta,\gamma'\right)' + \frac{1}{\sigma_0^2}\sum_{t=1}^{[\lambda T]}\left(\mathbf{s}_t\left(L;\boldsymbol{\tau}_0\right)\varepsilon_t\right)\varepsilon_t - \frac{\lambda}{\sigma_0^2}\sum_{t=1}^{T}\left(\mathbf{s}_t\left(L;\boldsymbol{\tau}_0\right)\varepsilon_t\right)\varepsilon_t + o_p\left(T^{1/2}\right).$$

Hence, under Assumption 1 and using a standard central limit theorem for martingale differences and the tightness from Lemma A.3, we can show that for $\lambda\in\Lambda$ in $\mathcal{D}(\Lambda)$

$$T^{-1/2}\tilde{\mathcal{L}}_{(\theta,\beta),T}(\lambda) \Rightarrow \left(\lambda(1-\lambda_0)-(\lambda-\lambda_0)_+\right)\boldsymbol{\Xi}\left(\delta,\gamma'\right)' + \boldsymbol{\Xi}^{1/2}\left[B_{p+1}(\lambda)-\lambda B_{p+1}(1)\right], \quad (18)$$

where $\Rightarrow$ denotes weak converge with the Skorohod metric.

Next, in the direction of ν, for a given λ,

$$\tilde{\mathcal{L}}_{\nu,T}(\lambda) = \frac{\partial}{\partial\nu}\mathcal{L}\left(\tilde{\psi}_T,\lambda\right)\Big|_{\psi=\tilde{\psi}_T} = \frac{1}{\tilde{\sigma}_T^2}\sum_{t=1}^{T}\left(R_t^{(2)}(\lambda)\,\tilde{\alpha}_{0T}(L)\,\Delta_{t-[\lambda T]}^{\tilde{d}_{0T}}1\right)\tilde{\varepsilon}_t.$$

First, by a Taylor expansion,

$$\begin{aligned}\tilde{\mathcal{L}}_{\nu,T}(\lambda) &= \frac{1}{\tilde{\sigma}_T^2}\sum_{t=[\lambda T]+1}^{T}\left(\alpha_0(L)\,\Delta_{t-[\lambda T]}^{d_0}1\right)\tilde{\varepsilon}_t + \frac{1}{\tilde{\sigma}_T^2}\left(\tilde{d}_{0T}-d_0\right)\sum_{t=[\lambda T]+1}^{T}\left(\tilde{\alpha}^*_{0T}(L)\log\Delta_t\Delta_{t-[\lambda T]}^{d^*_{0T}}1\right)\tilde{\varepsilon}_t \\ &\quad + \frac{1}{\tilde{\sigma}_T^2}\left(\left[\tilde{\alpha}_{0T}(L)-\alpha_0(L)\right]\Delta_{t-[\lambda T]}^{d_0}1\right)\tilde{\varepsilon}_t \\ &= \frac{1}{\tilde{\sigma}_T^2}\sum_{t=[\lambda T]+1}^{T}\left(\alpha_0(L)\,\Delta_{t-[\lambda T]}^{d_0}1\right)\tilde{\varepsilon}_t + o_p(T^{1/2-d_0}),\end{aligned}$$

uniformly in λ, where $\left\|\boldsymbol{\tau}^*_{0T}-\boldsymbol{\tau}_0\right\| \leq \left\|\tilde{\boldsymbol{\tau}}_{0T}-\boldsymbol{\tau}_0\right\|$. Next, using again the Taylor expansion (13) around $(d_0,\boldsymbol{\alpha}_0',\mu_0)$ yields

$$\begin{aligned}\tilde{\mathcal{L}}_{\nu,T}(\lambda) &= \frac{1}{\tilde{\sigma}_T^2}\sum_{t=[\lambda T]+1}^{T}\left(\alpha_0(L)\,\Delta_{t-[\lambda T]}^{d_0}1\right)\varepsilon_t^0 && (19)\\ &\quad + \frac{1}{\tilde{\sigma}_T^2}\left(\tilde{d}_{0T}-d_0\right)\sum_{t=[\lambda T]+1}^{T}\left(\alpha_0(L)\,\Delta_t^{d_0}1\right)\log\Delta_t\varepsilon_t^* && (20)\\ &\quad + \frac{1}{\tilde{\sigma}_T^2}\left(\tilde{\alpha}_{0T}(L)-\alpha_0(L)\right)\sum_{t=[\lambda T]+1}^{T}\left(\alpha_0(L)\,\Delta_t^{d_0}1\right)\frac{\varepsilon_t^*}{\alpha^*_{0T}(L)} && (21)\\ &\quad - \frac{1}{\tilde{\sigma}_T^2}\left(\tilde{\mu}_{0T}-\mu_0\right)\sum_{t=[\lambda T]+1}^{T}\left(\alpha_0(L)\,\Delta_t^{d_0}1\right)\left(\alpha_0(L)\,\Delta_{t-[\lambda T]}^{d_0}1\right) && (22)\\ &\quad + o_p(T^{1/2-d_0}).\end{aligned}$$

The terms (20) and (21) are of smaller order than the remaining ones above by Lemma A.2 and $\tilde{\boldsymbol{\tau}}_{0T}-\boldsymbol{\tau}_0 = O_p\left(T^{-1/2}\right)$.

Next, from the term (22),

$$\left(\tilde{\mu}_{0T}-\mu_0\right)\sum_{t=[\lambda T]+1}^{T}\left(\alpha_0\left(L\right)\Delta_t^{d_0}1\right)\left(\alpha_0\left(L\right)\Delta_{t-[\lambda T]}^{d_0}1\right) = \frac{\sum_{t=[\lambda T]+1}^{T}\left(\Delta_t^{d_0}1\right)\left(\Delta_{t-[\lambda T]}^{d_0}1\right)}{\sum_{t=1}^{T}\left(\Delta_t^{d_0}1\right)^2}\sum_{t=1}^{T}\left(\alpha_0\left(L\right)\Delta_t^{d_0}1\right)\varepsilon_t^0+o_p\left(1\right), \tag{23}$$

since, without loss of generality, for $p=1$,

$$\alpha_0\left(L\right)\Delta_t^{d_0}1 = \Delta_t^{d_0}1-\alpha\Delta_{t-1}^{d_0}1 = \alpha_0\left(1\right)\Delta_t^{d_0}1+\pi_{t-1}\left(d_0\right)$$

where the second term is of order $(t-1)^{-d_0-1}$ and thus negligible and the resulting factor $\alpha_0\left(1\right)$ in (23) cancels. The first factor of (23) converges deterministically in λ to $L\left(d_0;0,\lambda\right)$ where

$$L\left(d_0,a,b\right)\equiv\left(1-2d_0\right)\int_{\max(a,b)}^{1}\left(s-a\right)^{-d_0}\left(s-b\right)^{-d_0}ds$$

with $L\left(d_0,a,a\right)=\left(1-a\right)^{1-2d_0}$. Therefore, combining the term (22) with (19), yields, up to $o_p(T^{1/2-d_0})$ terms, that $\tilde{\mathcal{L}}_{\nu,T}$ is equal to

$$\frac{1}{\tilde{\sigma}_T^2}\sum_{t=[\lambda T]+1}^{T}\left(\alpha_0\left(L\right)\Delta_{t-[\lambda T]}^{d_0}1\right)\varepsilon_t^0-\frac{L\left(d_0;0,\lambda\right)}{\tilde{\sigma}_T^2}\sum_{t=1}^{T}\left(\alpha_0\left(L\right)\Delta_t^{d_0}1\right)\varepsilon_t^0. \tag{24}$$

Using again (10), the first term in (24) can be rewritten as

$$\frac{1}{\tilde{\sigma}_T^2}\sum_{t=[\lambda T]+1}^{T}\left(\alpha_0\left(L\right)\Delta_{t-[\lambda T]}^{d_0}1\right)\left(\varepsilon_t-T^{-1/2}R_t^{(2)}\left(\lambda_0\right)\left(\delta\log\Delta_t\varepsilon_t+\sum_{j=1}^{p}\gamma_j\frac{\varepsilon_{t-j}}{\alpha_0\left(L\right)}\right)+\eta T^{d_0-1/2}R_t^{(2)}\left(\lambda_0\right)\alpha_0\left(L\right)\Delta_{t-[\lambda_0 T]}^{d_0}1\right), \tag{25}$$

plus terms of smaller order of magnitude.

The last term in (25), multiplied by $T^{d_0-1/2}$, is

$$T^{2d_0-1}\frac{\eta}{\sigma_0^2}\sum_{t=[\max(\lambda,\lambda_0)T]+1}^{T}\left(\alpha_0\left(L\right)\Delta_{t-[\lambda T]}^{d_0}1\right)\left(\alpha_0\left(L\right)\Delta_{t-[\lambda_0]T}^{d_0}1\right)\overset{p}{\to}\frac{\eta}{\sigma_0^2}\frac{\alpha_0^2\left(1\right)L\left(d_0;\lambda_0,\lambda\right)}{\left(1-2d_0\right)\Gamma^2\left(1-d_0\right)},$$

uniformly in λ, while the second term of (25) is of smaller order. Multiplied by $T^{d_0-1/2}$, the first term of (25) converges weakly to

$$\frac{\alpha_0\left(1\right)\tilde{W}_{d_0}\left(\lambda,1\right)}{\sigma_0\sqrt{\left(1-2d_0\right)\Gamma^2\left(1-d_0\right)}},$$

where, under Assumption 1, the following convergence holds, similar to Marinucci and Robinson's (2000) results,

$$T^{d_0-1/2}\sum_{t=1}^{[\lambda T]}\left(\alpha_0\left(L\right)\Delta_t^{d_0}1\right)\varepsilon_t\overset{p}{\to}\alpha_0\left(1\right)T^{d_0-1/2}\sum_{t=1}^{[\lambda T]}\pi_{t-1}\left(d_0-1\right)\varepsilon_t\Rightarrow\frac{\sigma_0\alpha_0\left(1\right)\tilde{W}_{d_0}\left(0,\lambda\right)}{\sqrt{\left(1-2d_0\right)\Gamma^2\left(1-d_0\right)}}, \tag{26}$$

$$T^{d_0-1/2}\sum_{t=[\lambda T]+1}^{T}\left(\alpha_0\left(L\right)\Delta_{t-[\lambda T]}^{d_0}1\right)\varepsilon_t\overset{p}{\to}\alpha_0\left(1\right)T^{d_0-1/2}\sum_{t=[\lambda T]+1}^{T}\pi_{t-[\lambda T]-1}\left(d_0-1\right)\varepsilon_t\Rightarrow\frac{\sigma_0\alpha_0\left(1\right)\tilde{W}_{d_0}\left(\lambda,1\right)}{\sqrt{\left(1-2d_0\right)\Gamma^2\left(1-d_0\right)}}$$

with $\tilde{W}_{d_0}(a,b) = (1-2d_0)^{1/2} \int_a^b (s-a)^{-d_0} dB(s)$, (so that $\tilde{W}_{d_0}(0,1)$ has unit variance) and the convergence can be shown to be joint with (18). The fractional Brownian motion $\tilde{W}_{d_0}(0,\lambda)$ has the same marginal distribution as the standard one $W_{d_0}(\lambda) = (1-2d_0)^{1/2} \int_0^\lambda (\lambda - s)^{-d_0} dB(s)$ but has a different covariance,

$$\mathrm{Cov}\left(\tilde{W}_{d_0}(\lambda,1), \tilde{W}_{d_0}(0,1)\right) = L(d_0;0,\lambda). \tag{27}$$

rather than

$$\mathrm{Cov}\left(W_{d_0}(1), W_{d_0}(\lambda)\right) = 1 + \lambda^{1-2d_0} - E\left[W_{d_0}(1) - W_{d_0}(\lambda)\right]^2.$$

The second term of (24) converges to

$$-\frac{\eta}{\sigma_0^2} \frac{\alpha_0^2(1) L(d_0;0,\lambda) L(d_0;0,\lambda_0)}{(1-2d_0)\Gamma^2(1-d_0)} - L(d_0;0,\lambda) \frac{\alpha_0(1)\tilde{W}_{d_0}(0,1)}{\sigma_0\sqrt{(1-2d_0)\Gamma^2(1-d_0)}},$$

so that, combining the limits of the terms in (24), $T^{d_0-1/2}\tilde{\mathcal{L}}_{\nu,T}(\lambda)$ converges weakly to

$$\alpha_0(1) \frac{\tilde{W}_{d_0}(\lambda,1) - L(d_0;0,\lambda)\tilde{W}_{d_0}(0,1)}{\sigma_0\sqrt{(1-2d_0)\Gamma^2(1-d_0)}} + \eta\alpha_0^2(1) \frac{L(d_0;\lambda,\lambda_0) - L(d_0;0,\lambda_0) L(d_0;0,\lambda)}{\sigma_0^2(1-2d_0)\Gamma^2(1-d_0)}. \tag{28}$$

Finally, defining,

$$D_T = diag\left(T^{-1/2}, T^{-1/2}\mathbf{I}_{p\times p}, T^{d_0-1/2}\right),$$

then, by combining the joint convergence of $\tilde{\mathcal{L}}_{(\theta,\beta),T}(\lambda)$ and $\tilde{\mathcal{L}}_{\nu,T}(\lambda)$ as described in (18) and (28), the normalized score for the LM test for a break in all parameters behaves as

$$D_T\tilde{\mathcal{L}}_{(\theta,\boldsymbol{\beta}',\nu),T}(\lambda) = D_T \frac{\partial \mathcal{L}_T(\psi,\lambda)}{\partial(\theta,\boldsymbol{\beta}',\nu)'}|_{\psi=\tilde{\psi}_T}$$

$$\Rightarrow \begin{pmatrix} \boldsymbol{\Xi}^{1/2}\left(B_{p+1}(\lambda) - \lambda B_{p+1}(1)\right) + \left(\lambda(1-\lambda_0) - (\lambda-\lambda_0)_+\right)\boldsymbol{\Xi}\begin{pmatrix}\delta \\ \gamma\end{pmatrix} \\ \alpha_0(1)\frac{\tilde{W}_{d_0}(\lambda,1) - L(0,\lambda)\tilde{W}_{d_0}(0,1)}{\sigma_0\sqrt{(1-2d_0)\Gamma^2(1-d_0)}} + \eta\alpha_0^2(1)\frac{L(d_0;\lambda,\lambda_0) - L(d_0;0,\lambda_0)L(d_0;0,\lambda)}{\sigma_0^2(1-2d_0)\Gamma^2(1-d_0)} \end{pmatrix}, \tag{29}$$

with the two components being uncorrelated.

In addition, using similar methods, the relevant block of the inverse Hessian converges uniformly in λ as

$$D_T^{-1}\left(-\frac{\partial^2\mathcal{L}_T(\psi,\lambda)}{\partial\psi\partial\psi'}\Big|_{\psi=\tilde{\psi}_T}\right)^{-1}_{[1:(p+2),1:(p+2)]} D_T^{-1} \xrightarrow{p} \begin{pmatrix} \lambda(1-\lambda)\boldsymbol{\Xi} & 0 \\ 0 & \alpha_0^2(1)\frac{L(d_0;\lambda,\lambda) - L^2(d_0;0,\lambda)}{\sigma_0^2(1-2d_0)\Gamma^2(1-d_0)} \end{pmatrix}^{-1}, \tag{30}$$

because the limit of the Hessian is block diagonal between σ^2 and the rest of elements of ψ, while

$$-D_T \frac{\partial^2\mathcal{L}_T(\psi,\lambda)}{\partial(d,\boldsymbol{\alpha}',\mu)'\partial(d,\boldsymbol{\alpha}',\mu)}\Big|_{\psi=\tilde{\psi}_T} D_T \xrightarrow{p} \begin{pmatrix} \boldsymbol{\Xi} & 0 \\ 0 & \alpha_0^2(1)\frac{1}{\sigma_0^2(1-2d_0)\Gamma^2(1-d_0)} \end{pmatrix}$$

$$-D_T \frac{\partial^2\mathcal{L}_T(\psi,\lambda)}{\partial(\theta,\boldsymbol{\beta}',\nu)'\partial(\theta,\boldsymbol{\beta}',\nu)}\Big|_{\psi=\tilde{\psi}_T} D_T \xrightarrow{p} \begin{pmatrix} (1-\lambda)\boldsymbol{\Xi} & 0 \\ 0 & \alpha_0^2(1)\frac{L(d_0;\lambda,\lambda)}{\sigma_0^2(1-2d_0)\Gamma^2(1-d_0)} \end{pmatrix}$$

and the cross-derivatives block has the same last limit noticing that $L(d_0;0,\lambda) = L(d_0;\lambda,\lambda)$.

Finally, (29) and (30) establish the result.

Proof of Corollary 1

As a special case of Theorem 1 in DRV, for a known break fraction λ_0, because

$$Var\left[B_i\left(\lambda\right)-\lambda B_i\left(1\right)\right]=\lambda+\lambda^2-2\lambda^2=\lambda\left(1-\lambda\right),i=1,...,p+1,$$

the distribution of the derivative in the direction of memory and autoregressive component is,

$$T^{-1/2}\tilde{\mathcal{L}}_{(\theta,\boldsymbol{\beta}),T}\left(\lambda_0\right)\xrightarrow{d}N\left(\lambda_0\left(1-\lambda_0\right)\boldsymbol{\Xi}\begin{pmatrix}\delta\\ \boldsymbol{\gamma}\end{pmatrix},\lambda_0\left(1-\lambda_0\right)\boldsymbol{\Xi}\right), \tag{31}$$

where the drift and the asymptotic variance are symmetric around $\lambda_0=\frac{1}{2}$. Further, this convergence is joint with that of the derivative in the direction of the level,

$$T^{d_0-1/2}\tilde{\mathcal{L}}_{\nu,T}\left(\lambda_0\right)\xrightarrow{d}N\left(\frac{\eta}{\sigma_0}\alpha_0\left(1\right)\frac{L\left(d_0;\lambda,\lambda\right)-L^2\left(d_0;0,\lambda\right)}{\left(1-2d_0\right)\Gamma^2\left(1-d_0\right)},\alpha_0^2\left(1\right)\frac{L\left(d_0;\lambda,\lambda\right)-L^2\left(d_0;0,\lambda\right)}{\left(1-2d_0\right)\Gamma^2\left(1-d_0\right)}\right),$$

because

$$\begin{aligned}Var\left[\tilde{W}_{d_0}\left(\lambda,1\right)-L\left(d_0;0,\lambda\right)\tilde{W}_{d_0}\left(0,1\right)\right]&=L\left(d_0;\lambda,\lambda\right)+L^2\left(d_0;0,\lambda\right)-2L^2\left(d_0;0,\lambda\right)\\&=L\left(d_0;\lambda,\lambda\right)-L^2\left(d_0;0,\lambda\right),\end{aligned}$$

and with zero correlation with the limit in (31). Combining with the behaviour of the Hessian in (30) establishes the convergence to a $\chi^2_{2+p}(c)$ under $H_{1,T}^{d,\alpha,\mu}\left(\lambda_0\right)$.

Proof of Proposition 1

Consistency of the test in the second regime follows from the fact that the limit of $\left(\tilde{d}_{0T},\tilde{\boldsymbol{\alpha}}'_{0T},\tilde{\mu}_{0T}\right)$, differs from $\left(d_1,\boldsymbol{\alpha}'_1,\mu_1\right)$, which is true since

$$\left(\tilde{d}_{0T},\tilde{\boldsymbol{\alpha}}'_{0T},\tilde{\mu}_{0T}\right)=\arg\min_{d,\boldsymbol{\alpha}',\mu}\sum_{t=1}^{[\lambda_0T]}\left[\alpha\left(L\right)\Delta_t^d\left(y_t-\mu\right)\right]^2+\sum_{t=[\lambda_0T]+1}^{T}\left[\alpha\left(L\right)\Delta_t^d\left(y_t-\mu\right)\right]^2 \tag{32}$$

and noticing that the *argmin* of the sum of two convex terms differs from the *argmins* of the respective summands.

Then, in the MVT theorems in the proof of Theorem 1, equations (14), (15) and (22) respectively, the terms $(\tilde{d}_{0T}-d_0)$, $\left(\tilde{\boldsymbol{\alpha}}'_{0T}-\boldsymbol{\alpha}'_0\right)$ and $(\tilde{\mu}_{0T}-\mu_0)$ converge to nonzero constants. The proof works then by showing that in this case, the rate needed for convergence of the numerator of the test statistics is the square of the rate needed under the null (or local alternative). Thus, the test statistic diverges.

In particular, we analyze how the components $\tilde{\mathcal{L}}_{\theta,T}\left(\lambda\right)$, $\tilde{\mathcal{L}}_{\beta,T}\left(\lambda\right)$ and $\tilde{\mathcal{L}}_{\nu,T}\left(\lambda\right)$ behave under the different alternatives, $H_1^{d,\alpha}\left(\lambda_0\right),H_1^{\mu}\left(\lambda_0\right)$ and $H_1^{d,\alpha,\mu}\left(\lambda_0\right)$. Under both $H_1^{d,\alpha,\mu}\left(\lambda_0\right)$ and $H_1^{d,\alpha}\left(\lambda_0\right)$, since $(\tilde{d}_{0T}-d_1)=O_p\left(1\right)$, the term (14) in $\tilde{\mathcal{L}}_{\theta,T}\left(\lambda\right)$ is of order $O_p\left(T\right)$ rather than $O_p(T^{1/2})$, where the rates are sharp. Under $H_1^{\mu}\left(\lambda_0\right)$, and for $d_0>0$, since $(\tilde{\mu}_{0T}-\mu_0)=O_p\left(1\right)$, the term (22) in $\tilde{\mathcal{L}}_{\nu,T}\left(\lambda\right)$ is of order $O_p\left(T^{1-2d_0}\right)$ rather than $O_p(T^{1/2-d_0})$. For $d_0<0$, the estimator $\tilde{d}_{0T}$ converges, from an argument similar to the one in Proposition 1 in Rachinger (2017), to 0 and thus $\tilde{\mathcal{L}}_{\nu,T}\left(\lambda\right)$ is of order $O_p\left(T\right)$. Combining these results with the behaviour of the denominator (30), the LM test statistics (4) in DRV for a break in both parameters diverge at rate T if there is a break in the dynamics ($H_1^{d,\alpha,\mu}\left(\lambda_0\right)$ or $H_1^{d,\alpha}\left(\lambda_0\right)$) and it diverges under $H_1^{\mu}\left(\lambda_0\right)$ at rate T^{1-2d_0} if $d_0\geq0$ and at rate T if $d_0<0$, the latter because $\tilde{d}_{0T}\xrightarrow{p}0$ in this case.

Proof of Proposition 2

Part a) Under both $H_1^{d,\mu}(\lambda_0)$ and $H_1^d(\lambda_0)$, $\bar{d}_{0T}$ behaves similarly as $\tilde{d}_{0T}$, implying $(\bar{d}_{0T} - d_1) = O_p(1)$. Thus, the corresponding term in $\tilde{\mathcal{L}}_{\theta,T}(\lambda)$ similar to (14) is of order $O_p(T)$ rather than $O_p(T^{1/2})$. Thus, the test diverges. Next, under $H_1^{\mu}(\lambda_0)$, the corresponding term (14) with $\tilde{\mu}_{0T}$ replaced by $\bar{\mu}_{1T}$ behaves as the one in Theorem 1. Thus, similarly as in Corollary 1, we obtain the χ^2 limit.

Part b) Under $H_1^{d,\mu}(\lambda_0)$ and $H_1^{\mu}(\lambda_0)$, $\bar{\mu}_{0T}$ behaves similarly as $\tilde{\mu}_{0T}$, implying $(\bar{\mu}_{0T} - \mu_0) = O_p(1)$. Under $H_1^{\mu}(\lambda_0)$, and for $d_0 \geq 0$, the corresponding term (22) in $\tilde{\mathcal{L}}_{\nu,T}(\lambda)$ is then of order $O_p(T^{1-2d_0})$ rather than $O_p(T^{1/2-d_0})$ and the test diverges at rate T^{1-2d_0}. For $d_0 < 0$, the estimator $\bar{d}_{1T}$ converges, from an argument similar to the one in Proposition 1 in Rachinger (2017), to 0 and thus $\tilde{\mathcal{L}}_{\nu,T}(\lambda)$ is of order $O_p(T)$ and the test diverges at rate T. Under $H_1^{d,\mu}(\lambda_0)$, for $d_1 \geq 0$, $\bar{d}_{1T}$ converges to d_1 and the test diverges at rate T^{1-2d_1}; for $d_1 < 0$, $\bar{d}_{1T}$ converges to 0 and the test diverges at rate T. Finally, under $H_1^d(\lambda_0)$, the corresponding term (22) with $\tilde{d}_{0T}$ replaced by $\bar{d}_{1T}$ behaves as the one in Theorem 1 as well. Therefore, as in Part a) we obtain the χ^2 limit.

Proof of Theorem 2

Part a) First, we consider the unfeasible estimates which use true values of $(d_0, \boldsymbol{\alpha}_0', \mu_0)$ and $(\theta, \boldsymbol{\beta}', \nu)$ for a given λ. Using (10), we have that under $H_{1,T}$

$$\begin{aligned} D_T \sum_{t=1}^{T} X_t^0(\lambda) X_t^0(\lambda)' D_T &= D_T \sum_{t=[\lambda T]+1}^{T} \bar{X}_t^0(\lambda) \bar{X}_t^0(\lambda)' D_T + o_p(1) \\ &= \operatorname{diag}\left\{(1-\lambda)\sigma_0^2 \boldsymbol{\Xi}, \frac{\alpha_0^2(1)(1-\lambda)^{1-2d_0}}{(1-2d_0)\Gamma^2(1-d_0)}\right\} + o_p(1), \end{aligned}$$

uniformly in λ, applying Lemmas A.1 and A.2, where $\bar{X}_t^0(\lambda) = \left(\log \Delta_t \varepsilon_t, \left\{\frac{\varepsilon_{t-j}}{\alpha_0(L)}\right\}_{j=1}^{p}, \ \alpha_0(L) \Delta_t^{d_0} R_t^{(2)}(\lambda)\right)'$ because we can write

$$X_t^0(\lambda) = \left(\left[\frac{1-\Delta_t^{\theta R_t^{(2)}(\lambda)}}{\theta}\right] \varepsilon_t^0, R_t^{(2)}(\lambda)\left\{\frac{\varepsilon_{t-j}}{\alpha_1(L)}\right\}_{j=1}^{p}, \ \alpha_0(L) \Delta_t^{d_0+\theta} R_t^{(2)}(\lambda)\right)'$$

and therefore, for instance,

$$\begin{aligned} \frac{1}{T}\sum_{t=1}^{T}\left(\left[\frac{1-\Delta_t^{\theta R_t^{(2)}(\lambda)}}{\theta}\right]\varepsilon_t^0\right)^2 &= \frac{1}{T}\sum_{t=[\lambda T]+1}^{T}\left(\left[\frac{1-\Delta_t^{\left(\delta T^{-1/2}\right)}}{\left(\delta T^{-1/2}\right)}\right]\varepsilon_t^0\right)^2 \\ &= \frac{1}{T}\sum_{t=[\lambda T]+1}^{T}\left(\left[\log \Delta_t + T^{-1/2} r_{t,T}(L;\delta)\right]\varepsilon_t^0\right)^2 \\ &\xrightarrow{p} (1-\lambda)\sigma_0^2 \sum_{j=1}^{\infty}\frac{1}{j^2} = (1-\lambda)\sigma_0^2\frac{\pi^2}{6}, \end{aligned}$$

where $r_{T,t}(L;\delta) = \sum_{j=1}^{t} r_{T,j}(\delta) L^j$ denotes a filter depending on T, with $|r_{T,j}(\delta_0)| \leq K j^{\varsigma-1}$ as $T \to \infty$ for any $\varsigma > 0$.

Then, using similar arguments, under $H_{1,T}$,

$$\begin{aligned} D_T \sum_{t=1}^{T} X_t^0(\lambda) Y_t^0 &= D_T \sum_{t=[\lambda T]+1}^{T} \bar{X}_t^0(\lambda) \varepsilon_t^0 + o_p(1) \\ &= T^{-1/2} D_T \sum_{t=[\lambda T]+1}^{T} \bar{X}_t^0(\lambda) b_{t,T}(\lambda_0) + D_T \sum_{t=[\lambda T]+1}^{T} \bar{X}_t^0(\lambda) \varepsilon_t + o_p(1) \end{aligned}$$

uniformly in λ, where $b_{t,T}(\lambda_0) = -R_t^{(2)}(\lambda_0)\left(\delta \log \Delta_t \varepsilon_t + \sum_{j=1}^{p} \gamma_j \alpha_0^{-1}(L)\,\varepsilon_{t-j} - \eta T^{d_0} \alpha_0(L)\,\Delta_{t-[\lambda_0 T]}^{d_0} 1\right)$.

Next,

$$T^{-1/2} D_T \sum_{t=[\lambda T]+1}^{T} \bar{X}_t^0(\lambda)\, b_{t,T}(\lambda_0) \to_p \left[-\sigma_0^2 \left(1 - \max(\lambda, \lambda_0)\right)(\delta\ \gamma')\,\mathbf{\Xi},\ \ \eta \frac{\alpha_0^2(1)\, L(d_0; \lambda_0, \lambda)}{(1-2d_0)\,\Gamma^2(1-d_0)}\right]'$$

uniformly in λ, while

$$D_T \sum_{t=[\lambda T]+1}^{T} \bar{X}_t^0(\lambda)\,\varepsilon_t \Rightarrow \left[\sigma_0^2 \left(\mathbf{\Xi}^{1/2}\left(B_{p+1}(1) - B_{p+1}(\lambda)\right)\right)',\ \ \frac{\sigma_0 \alpha_0(1)\, \tilde{W}_{d_0}(\lambda, 1)}{\sqrt{(1-2d_0)\,\Gamma^2(1-d_0)}}\right]' \tag{33}$$

by applying a joint FCLT, where the two terms on the right are independent.

We now deal with the estimation of the parameters. Note that with the case of unknown $\theta = d_1 - d_0$ can be dealt exactly as in LV. Likewise, estimated $\boldsymbol{\beta}' = \boldsymbol{\alpha}_1' - \boldsymbol{\alpha}_0'$ and $\nu = \mu_1 - \mu_0$ do not affect the asymptotic distribution because they are $O_p\left(T^{-1/2}\right)$ and $O_p\left(T^{d_0-1/2}\right)$, respectively, under $H_{1,T}$.

To deal with the effect of estimation of the unknown parameters $(d_0, \boldsymbol{\alpha}_0, \mu_0)$ by CSS under the null using observations for the whole sample, note that, using (13), and for $\tilde{X}_t(\lambda) = X_t\left(\lambda, \hat{\theta}_T(\lambda), \hat{\nu}_T(\lambda), \tilde{d}_{0T}, \tilde{\alpha}_{0T}, \tilde{\mu}_{0T}\right)$, we can write again under $H_{1,T}$,

$$\begin{aligned} D_T \sum_{t=1}^{T} \tilde{X}_t(\lambda)\, \tilde{X}_t(\lambda)'\, D_T &= D_T \sum_{t=[\lambda T]+1}^{T} \bar{X}_t^0(\lambda)\, \bar{X}_t^0(\lambda)'\, D_T + o_p(1) \\ &= \operatorname{diag}\left\{(1-\lambda)\,\sigma_0^2 \mathbf{\Xi},\ \ \frac{\alpha_0^2(1)\,(1-\lambda)^{1-2d_0}}{(1-2d_0)\,\Gamma^2(1-d_0)}\right\} + o_p(1), \end{aligned} \tag{34}$$

uniformly in λ, because $\tilde{\boldsymbol{\tau}}_{0T} - \boldsymbol{\tau}_0 = O_p\left(T^{-1/2}\right)$ and $\tilde{\mu}_{0T} - \mu_0 = O_p\left(T^{d_0-1/2}\right)$, so we can write

$$\tilde{X}_t(\lambda) = \left(\left[\frac{1 - \Delta_t^{\hat{\theta}_T(\lambda) R_t^{(2)}(\lambda)}}{\hat{\theta}_T(\lambda)}\right] \tilde{\varepsilon}_t, R_t^{(2)}(\lambda) \left\{\Delta_t^{\hat{d}_{1T}(\lambda)} \left(\frac{\Delta_t^{-\hat{d}_{1T}(\lambda)} \varepsilon_{t-j}}{\hat{\alpha}_{1T}(L)} + \mu_1 - \hat{\mu}_{1T}(\lambda)\right)\right\}_{j=1}^{p},\ \ \tilde{\alpha}_{0T}(L)\, \Delta_t^{\hat{d}_{1T}(\lambda)} R_t^{(2)}(\lambda)\right)'$$

and, for instance, uniformly in λ,

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^{T} \left(\left[\frac{1 - \Delta_t^{\hat{\theta}_T(\lambda) R_t^{(2)}(\lambda)}}{\hat{\theta}_T(\lambda)}\right] \tilde{\varepsilon}_t\right)^2 &= \frac{1}{T} \sum_{t=[\lambda T]+1}^{T} \left(\left[\log \Delta_t + \hat{\theta}_T(\lambda) r_t(L) + c_T r_{T,t}\left(L; \hat{\theta}_T(\lambda)\right)\right] \varepsilon_t^0\right)^2 + o_p(1) \\ &= \frac{1}{T} \sum_{t=[\lambda T]+1}^{T} \left(\log \Delta_t \varepsilon_t^0\right)^2 + o_p(1) \xrightarrow{p} (1-\lambda)\,\sigma_0^2 \frac{\pi^2}{6}, \end{aligned}$$

where $c_T = O_p\left(T^{-n/2}\right)$ for some $n > 2$ and $r_t(L) = \sum_{j=1}^{t} r_{T,j} L^j$ and $r_{T,t}(L; \theta) = \sum_{j=1}^{t} r_{T,j}(\theta)\, L^j$ denote filters with $|r_{T,j}| \le K j^{\varsigma - 1}$, any $\varsigma > 0$, and $|r_{T,j}(\theta)| \le K j^{\theta-1} \log^n j$ as $T \to \infty$, and then proceeding as in Lemma C.5 in Robinson and Hualde (2003) choosing n large enough.

Then, under $H_{1,T}$, for $\tilde{Y}_t = \tilde{\alpha}_{0T}(L)\, \Delta_t^{\tilde{d}_{0T}}(y_t - \tilde{\mu}_{0T}) = \tilde{\varepsilon}_t$, and using the same methods as in (13) and

Lemma A.1, $D_T \sum_{t=1}^{T} X_t^0(\lambda)\tilde{Y}_t$ equals

$$D_T \sum_{t=[\lambda T]+1}^{T} \bar{X}_t^0(\lambda)\varepsilon_t^0 + D_T \sum_{t=[\lambda T]+1}^{T} \bar{X}_t^0(\lambda) \left\{ \begin{array}{c} (\tilde{d}_{0T} - d_0)\log \Delta_t \varepsilon_t^* \\ +(\tilde{\alpha}_{0T}(L) - \alpha_0(L)) \frac{\varepsilon_t^*}{\alpha_{0T}^*(L)} \\ -(\tilde{\mu}_{0T} - \mu_0)\alpha_0(L)\Delta_t^{d_{0T}^*} 1 \end{array} \right\} + o_p(1)$$

$$= D_T \sum_{t=[\lambda T]+1}^{T} \bar{X}_t^0(\lambda)\varepsilon_t^0 + \text{diag}\left[\sigma_0^2(1-\lambda)\,\boldsymbol{\Xi} \quad - \frac{\alpha_0^2(1)L(d_0;0,\lambda)}{(1-2d_0)\Gamma(1-d_0)^2}\right]\left(\tilde{\boldsymbol{\psi}}_{0T} - \boldsymbol{\psi}_0\right) + o_p(1) \quad (35)$$

$$= D_T \sum_{t=[\lambda T]+1}^{T} \bar{X}_t^0(\lambda)\varepsilon_t^0 + \left[\sigma_0^2(1-\lambda)(1-\lambda_0)(\delta,\gamma')\,\boldsymbol{\Xi} \quad -\eta\frac{\alpha_0^2(1)L(d_0;0,\lambda)L(d_0;0,\lambda_0)}{(1-2d_0)\Gamma^2(1-d_0)}\right]'$$

$$-D_T\left[(1-\lambda)\sum_{t=1}^{T}\varepsilon_t(\mathbf{s}_t(L;\boldsymbol{\tau}_0)\varepsilon_t)' \quad L(d_0;0,\lambda)\sum_{t=1}^{T} m_t(\boldsymbol{\tau}_0)\varepsilon_t\right]' + o_p(1)$$

uniformly in λ, using (12) and (23). Then, combining this result with (33) we obtain

$$D_T \sum_{t=1}^{T} X_t^0(\lambda)\tilde{Y}_t \;\Rightarrow\; \left[-\sigma_0^2(1-\max(\lambda,\lambda_0))(\delta\;\gamma')\,\boldsymbol{\Xi},\;\; \eta\frac{\alpha_0^2(1)L(d_0;\lambda_0,\lambda)}{(1-2d_0)\Gamma^2(1-d_0)}\right]'$$
$$+\left[\sigma_0^2(1-\lambda)(1-\lambda_0)(\delta,\gamma')\,\boldsymbol{\Xi},\;\; -\eta\frac{\alpha_0^2(1)L(d_0;0,\lambda)L(d_0;0,\lambda_0)}{(1-2d_0)\Gamma^2(1-d_0)}\right]'$$
$$+\left[\sigma_0^2\left(\boldsymbol{\Xi}^{1/2}(B_{p+1}(1) - B_{p+1}(\lambda))\right)',\;\; \frac{\sigma_0\alpha_0(1)\tilde{W}_{d_0}(\lambda,1)}{\sqrt{(1-2d_0)\Gamma^2(1-d_0)}}\right]'$$
$$-\left[(1-\lambda)\sigma_0^2\left(\boldsymbol{\Xi}^{1/2}B_{p+1}(1)\right)',\;\; \frac{\sigma_0\alpha_0(1)L(d_0;0,\lambda)\tilde{W}_{d_0}(0,1)}{\sqrt{(1-2d_0)\Gamma^2(1-d_0)}}\right]'$$

and this together with (34) shows that the asymptotic distribution of $\widetilde{LMW}_{2,T}(\lambda)$ is as in Theorem 1.

Part b) The proof follows from Part a) in a similar way as in the proof of Corollary 1.

Proof of Proposition 3

The proof follows as that of Proposition 1 by exploiting the different behaviour of estimates under the null and the alternative. In particular, $\tilde{\psi}_{0T}$ and $\hat{\psi}_{1T}$ are mimizing over different sums, the former as in (32), the latter, for $\lambda < \lambda_0$,

$$\left(\hat{d}_{1T}, \hat{\boldsymbol{\alpha}}_{1T}', \hat{\mu}_{1T}\right) = \arg\min_{d_1,\boldsymbol{\alpha}_1',\mu_1} \sum_{t=[\lambda T]+1}^{[\lambda_0 T]} \left[\alpha_1(L)\Delta_{t-[\lambda T]}^{d_1}(y_t - \mu_1)\right]^2 + \sum_{t=[\lambda_0 T]+1}^{T} \left[\alpha_1(L)\Delta_{t-[\lambda T]}^{d_1}(y_t - \mu_1)\right]^2.$$

As for $\tilde{\psi}_{0T}$, the limit of the first summand is minimized by $(d_0, \boldsymbol{\alpha}_0', \mu_0)$, while the limit of the second summand is minimized by $(d_1, \boldsymbol{\alpha}_1', \mu_1)$. However, the first summand consists for $\hat{\psi}_{1T}$ only of $[(\lambda_0 - \lambda)T]$ terms, with $\lambda > \epsilon$, while for $\tilde{\psi}_{0T}$ it consists of $[\lambda_0 T]$ terms. Thus, similarly in Proposition 1, the limits of limit of $\tilde{\psi}_{0T}$ and $\hat{\psi}_{1T}$ differ. For $\lambda \geq \lambda_0$, the argument simplifies since only a term minimized by $(d_1, \boldsymbol{\alpha}_1', \mu_1)$ remains.

Then, under the fixed alternative of one break in the dynamics in any direction, $\theta \neq 0$ and/or $\boldsymbol{\beta} \neq 0$, under $H_1^{d,\alpha}$ and under $H_1^{d,\alpha,\mu}$ plim$\hat{\theta} \neq 0$ and plim$\hat{\beta} \neq 0$ so that for a similar reason as in Proposition 1, the OLS estimates $\hat{\vartheta}_{1T}(\lambda)$ and $\hat{\vartheta}_{\alpha T}(\lambda)$ converge to nonzero constants and the test statistic diverges.

Under H_1^μ, similarly, $\text{plim}(\hat{\mu}_{1T} - \tilde{\mu}_{0T}) \neq 0$. Further, as before, $\tilde{d}_{0T}$ converges to d_0 (resp. 0), if $d_0 \geq 0$ (resp. < 0). Then, the last term in the second term in (35) is of order $O_p(T^{1/2-d_0})$ if $d_0 \geq 0$ and of order $O_p(\sqrt{T})$, if $d_0 < 0$. Thus, $\widetilde{LMW}_{2,T} = O_p\left(T^{1-2d_0}\right)$, in the former, and $\widetilde{LMW}_{2,T} = O_p\left(T\right)$, in the latter case.

Finite sample properties of LM and LMW-type tests implemented in Regimes 1 and 2

In order to illustrate the gains from the symmetric tests, we repeat the exercise leading to Table 1 in DRV but now for the tests implemented in Regimes 1 and 2 separately. In particular, consider the case of a known break fraction of $\lambda_0 = 0.5$. Also here, the significance level is 0.05 and the sample sizes are $T = 200, 500$ and $1,000$ when considering size, and $T = 200$ as regards power. We take draws from a $N(0,1)$ distribution. For the size, d takes the values $\{-0.4, -0.3, -0.2, -0.1, 0, 0.1, 0.2, 0.3, 0.4\}$ and a non-breaking level of $\mu_0 = 0$. To compute power, we consider $d_0 \in \{-0.2, 0, 0.2\}$, $d_1 \in \{-0.4, -0.2, 0, 0.2, 0.4\}$, $\mu_0 = 0$ and $\mu_1 = \{0, 0.25, 0.5, 1\}$. The number of simulations is $10,000$. Table A1 (panels a through d) displays the size of the LM and LMW-type tests, respectively, when testing in the first and second regime and for breaks in both d and μ. Next, Table A2 displays the power results of the two tests for a break in d and/or μ at $\lambda_0 = 0.5$. Figures in bold characters correspond to size. In comparison to Table 1 in DRV, here size is less stable over different memory parameters and power stronger depends on the direction of the break.

References

[1] Billingsley, P, 1968. *Convergence of Probability Measures*, Wiley, New York.

[2] Hualde, J. and Nielsen, M.O., 2017. "Truncated sum of squares estimation of fractional time series models with deterministic trends". QED Working Paper 1376.

[3] Hualde, J. and Robinson, P.M., 2011. "Gaussian pseudo-maximum likelihood estimation of fractional time series models". *Annals of Statistics* 39, 3152-3181.

[4] Marinucci, D., Robinson, P.M., 2000. "Weak convergence of multivariate fractional processes". *Stochastic Processes and their Applications* 83, 103-120.

[5] Rachinger, H., 2017. "Multiple breaks in long memory time series". University of Vienna, mimeo. https://sites.google.com/site/heikorachinger/Rachinger2017.pdf

[6] Robinson, P.M. and J. Hualde, 2003, "Cointegration in fractional systems with unknown integration orders". *Econometrica* 71 1727-1766.

Table A1: **Simulated size of LM and LMW-type tests for a joint break in memory and level.**

a) LM test in Regime 1

$T \setminus d_0$	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4
200	4.3	4.0	3.8	3.7	3.5	4.2	4.7	5.4	8.6
500	4.6	4.7	4.7	4.6	4.3	4.3	4.9	5.7	7.3
1000	5.3	4.9	4.3	4.4	4.9	4.4	5.2	5.4	6.2

b) LM test in Regime 2

$T \setminus d_0$	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4
200	6.6	6.0	5.4	4.7	4.4	4.5	4.0	4.0	4.3
500	5.6	5.4	5.3	4.8	4.6	4.6	4.3	4.5	4.9
1000	5.7	5.5	5.1	4.8	5.1	4.6	4.8	4.5	4.7

c) LMW-type test in Regime 1

$T \setminus d_0$	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4
200	6.6	6.4	6.2	5.8	7.8	7.4	6.8	6.3	7.1
500	5.4	5.8	6.4	6.6	6.5	6.5	6.7	7.0	7.8
1000	6.1	5.8	5.6	6.1	6.2	6.0	6.3	6.3	6.6

d) LMW-type test in Regime 2

$T \setminus d_0$	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4
200	7.6	7.8	6.8	6.9	5.8	6.4	7.7	6.6	6.8
500	6.7	6.6	6.1	6.4	6.7	6.4	6.3	6.2	6.1
1000	6.4	6.5	6.1	5.9	6.3	5.9	5.9	5.5	5.5

Rejection probabilities of 5% test for joint break in d and μ at $\lambda_0 = 0.5,\ \mu_0 = 0,\ \sigma_0^2 = 1$.

Table A2: **Simulated power of LM and LMW-type tests for a joint break in memory and level.**

a) LM test in Regime 1

d_0	-0.2				0				0.2			
$d_1 \setminus \mu_1$	0	0.25	0.5	1	0	0.25	0.5	1	0	0.25	0.5	1
-0.4	17.6	94.6	100	100	67.4	82.2	97.8	99.9	96.1	94.9	94.3	97.3
-0.2	**3.8**	71.7	99.8	100	18.0	55.6	93.2	99.9	71.1	74.5	85.7	97.8
0	31.8	56.3	96.6	100	**3.5**	18.9	71.6	100	15.5	26.2	51.9	93.4
0.2	83.9	87.9	96.0	99.9	34.7	41.2	60.4	97.8	**4.7**	8.8	23.4	80.0
0.4	99.1	99.4	99.4	100	89.1	89.3	92.0	96.6	42.7	45.6	54.9	79.4

b) LM test in Regime 2

d_0	-0.2				0				0.2			
$d_1 \setminus \mu_1$	0	0.25	0.5	1	0	0.25	0.5	1	0	0.25	0.5	1
-0.4	41.6	96.2	100	100	91.1	90.4	89.7	99.9	99.8	98.4	96.4	97.3
-0.2	**5.4**	74.2	99.7	100	41.6	68.8	95.0	99.9	90.1	92.1	94.4	98.3
0	11.1	49.6	96.5	100	**4.4**	21.5	71.5	99.9	35.1	42.6	61.9	94.2
0.2	58.1	70.7	89.0	99.9	12.9	24.1	50.0	96.8	**4.0**	7.2	21.8	77.9
0.4	95.1	94.4	95.4	100	62.7	66.5	72.2	89.7	18.9	21.7	30.0	64.7

c) LMW-type test in Regime 1

d_0	-0.2				0				0.2			
$d_1 \setminus \mu_1$	0	0.25	0.5	1	0	0.25	0.5	1	0	0.25	0.5	1
-0.4	16.6	92.9	100	100	65.0	80.0	96.5	100	96.5	95.4	94.2	96.8
-0.2	**6.2**	84.9	100	100	17.3	49.7	92.5	100	69.7	74.8	84.1	97.3
0	44.9	83.4	99.5	100	**5.8**	34.6	83.5	100	17.8	27.8	52.3	93.8
0.2	91.3	94.3	98.8	100	45.7	58.5	81.8	99.4	**6.5**	14.4	36.2	85.7
0.4	99.7	99.8	99.8	100	92.5	92.7	94.9	98.3	46.5	50.8	59.4	82.6

d) LMW-type test in Regime 2

d_0	-0.2				0				0.2			
$d_1 \setminus \mu_1$	0	0.25	0.5	1	0	0.25	0.5	1	0	0.25	0.5	1
-0.4	48.3	98.3	100	100	93.1	93.4	99.2	100	99.8	99.3	97.1	99.0
-0.2	**8.8**	84.1	100	100	49.6	79.8	98.8	100	93.0	94.3	97.1	99.9
0	14.0	49.0	95.7	100	**7.8**	34.8	83.5	100	45.8	58.0	80.0	99.1
0.2	63.9	70.1	86.8	99.3	16.0	24.8	51.4	95.0	**7.7**	14.1	33.5	83.2
0.4	96.4	96.8	97.2	99.2	66.7	68.5	75.3	88.6	17.1	21.1	28.6	57.9

Rejection probabilities of 5% test for joint break in d and μ at $\lambda_0 = 0.5$, $\mu_0 = 0$, $\sigma_0^2 = 1$, $T = 200$. Bold numbers correspond to size simulations.