Supplementary Appendix of

Score-driven time-varying parameter models with spline-based densities

Janneke van Brummelen, Paolo Gorgi and Siem Jan Koopman

C. Methodology overview

Algorithm 1: Computation of the log-likelihood for a given parameter vector $\boldsymbol{\theta}$

Input: Knot vector $\boldsymbol{t}$, static parameter vector $\boldsymbol{\theta}$, link function $g(f_t, \varepsilon_t)$, scaling function $S(f_t)$, starting value $\hat{f}_1$, and data $\{x_t\}_{t=1}^T$.

Step 1: Compute the spline coefficients a_i, b_i, c_i, d_i from the knot vector $\boldsymbol{t}$ and ordinates $\boldsymbol{y} = \tau(\boldsymbol{\psi})$, and construct the spline function h accordingly.

Step 2: Construct the antiderivative H of h, using formula in Section 2.4, imposing the normalization $H(0) = 0$.

Step 3: Given H, approximate the normalization constant C and the mean μ from H using the procedure described in Section F.

Step 4: Use $\{x_t\}_{t=1}^T$, C and μ to recursively compute the filtered sequence $\{\hat{f}_t(\boldsymbol{\theta})\}_{t=1}^T$, starting from $\hat{f}_1$, using the update equation (13) with $s(f,x) = S(f) \cdot \nabla(f,x)$, and $\nabla(f,x)$ as defined in (11).

Step 5: Use $\{\hat{f}_t(\boldsymbol{\theta})\}_{t=1}^T$, $\{x_t\}_{t=1}^T$, C and μ to calculate the average log-likelihood $\frac{1}{T}\sum_{t=1}^T \hat{\ell}_t(\boldsymbol{\theta})$ based on the expression in (14).

D. Proofs of lemmas

Proof of Lemma B.1. See Section E.1 for the expressions of the derivatives.

That for any $n > 0$, $\mathbb{E}\sup_{\boldsymbol{\theta}\in\Theta} |s(f_t(\boldsymbol{\theta}), x_t)|^n < \infty$, $\mathbb{E}\sup_{\boldsymbol{\theta}\in\Theta} \|\partial s(f_t(\boldsymbol{\theta}), x_t)/\partial\boldsymbol{\psi}\|^n < \infty$, and $\mathbb{E}\sup_{\boldsymbol{\theta}\in\Theta} |\partial s(f_t(\boldsymbol{\theta}), x_t)/\partial\sigma^2|^n < \infty$, follows from the assumptions on Θ and the fact that $s(f_t(\boldsymbol{\theta}), x_t)$, $\partial s(f_t(\boldsymbol{\theta}), x_t)/\partial\boldsymbol{\psi}$ and $\partial s(f_t(\boldsymbol{\theta}), x_t)/\partial\sigma^2$ are all at most third order piecewise polynomials evaluated in $x_t - f_t(\boldsymbol{\theta})$, and x_t and $f_t(\boldsymbol{\theta})$ are known to have bounded moments of any order by Propositions 1 and 2(i). In particular, due to the compactness of Θ, the conditions on $\boldsymbol{\psi}$ that ensure the spline-based density is well-defined and $\sigma > 0$ for any $\boldsymbol{\theta} \in \Theta$, the expectation of the supremum over Θ of these quantities will be finite.

For the remaining expectations, so those with a supremum over $f \in \mathbb{R}$, it can be argued that the derivatives are uniformly bounded over $\boldsymbol{\theta} \in \Theta$, $f \in \mathbb{R}$ and $x \in \mathbb{R}$. In particular, it is clear that $\sup_{z\in\mathbb{R},\boldsymbol{\psi}\in\boldsymbol{\Psi}} |h'(z;\boldsymbol{\psi})| < D$ for some constant $D < \infty$. Furthermore, the second order derivative of h is not only such that $\sup_{z\in\mathbb{R},\boldsymbol{\psi}\in\boldsymbol{\Psi}} |h''(z;\boldsymbol{\psi})| < D$ for some $D < \infty$, but also has value zero outside the outer knots, due to h being linear outside the outer knots. It follows that also $\sup_{z\in\mathbb{R},\boldsymbol{\psi}\in\boldsymbol{\Psi}} |zh''(z;\boldsymbol{\psi})| < D$ for some $D < \infty$. By again taking into account the assumptions on Θ, it is clear that the derivatives are indeed uniformly bounded over f and $\boldsymbol{\theta}$, and therefore have bounded moments of any order. □

Proof of Lemma B.2. See Section E.1 for the expressions of the derivatives.

For the expectations without a supremum over $f \in \mathbb{R}$ and the expectations with a supremum over $f \in \mathbb{R}$, exactly the same two respective argumentations can be used as in the proof of Lemma B.1. Here we also need to use that $\sup_{z\in\mathbb{R},\boldsymbol{\psi}\in\boldsymbol{\Psi}} |z^2h''(z;\boldsymbol{\psi})| < D$ for some constant $D < \infty$, which holds because $h''(z;\boldsymbol{\psi}) = 0$ whenever $z < t_1$ or $z > t_k$. □

Proof of Lemma B.3. The derivative

$$\frac{\partial^2 s(f, x_t)}{\partial f \partial f} = \frac{1}{\sigma^3} h''\left(\frac{x-f}{\sigma};\boldsymbol{\psi}\right),$$

is a continuous piecewise linear function of $x - f$. Due to the compactness of Θ and the other conditions in place, it is clear that the function h'' is Lipschitz continuous uniformly over $\boldsymbol{\psi} \in \boldsymbol{\Psi}$, i.e. there exists a constant K such that $\sup_{\boldsymbol{\psi}\in\boldsymbol{\Psi}} |h''(z_1;\boldsymbol{\psi}) - h''(z_2;\boldsymbol{\psi})| < K|z_1 - z_2|$ for any $z_1, z_2 \in \mathbb{R}$. Therefore,

$$\begin{aligned}
&\sup_{\boldsymbol{\theta}\in\Theta}\left| \left.\frac{\partial^2 s(f, x_t)}{\partial f \partial f}\right|_{f=\hat{f}_t(\boldsymbol{\theta})} - \left.\frac{\partial^2 s(f, x_t)}{\partial f \partial f}\right|_{f=f_t(\boldsymbol{\theta})}\right| \\
&\quad = \sup_{\boldsymbol{\theta}\in\Theta}\left| \frac{1}{\sigma^3} h''\left(\frac{x_t-\hat{f}_t(\boldsymbol{\theta})}{\sigma};\boldsymbol{\psi}\right) - \frac{1}{\sigma^3} h''\left(\frac{x_t-f_t(\boldsymbol{\theta})}{\sigma};\boldsymbol{\psi}\right)\right| \\
&\quad \le \sup_{\boldsymbol{\theta}\in\Theta}\frac{1}{\sigma^3} K \sup_{\boldsymbol{\theta}\in\Theta}\frac{1}{\sigma}\sup_{\boldsymbol{\theta}\in\Theta}\left|f_t(\boldsymbol{\theta}) - \hat{f}_t(\boldsymbol{\theta})\right| \overset{e.a.s}{\to} 0,
\end{aligned}$$

as $t \to \infty$, because $\sigma > 0$ for any $\boldsymbol{\theta} \in \Theta$, Θ is compact and $\hat{f}_t(\boldsymbol{\theta})$ converges e.a.s. to $f_t(\boldsymbol{\theta})$ uniformly over Θ by Proposition 2(i).

For the second derivative under consideration,

$$\frac{\partial^2 s(f, x;\boldsymbol{\gamma})}{\partial \sigma^2 \partial f} = \frac{1}{\sigma^4} h'\left(\frac{x-f}{\sigma};\boldsymbol{\psi}\right) - \frac{x-f}{\sigma^5} h''\left(\frac{x-f}{\sigma};\boldsymbol{\psi}\right),$$

we will show the e.a.s. convergence of the two terms separately, which implies convergence of the total expression by the subadditivity of the sup-norm. The first term is continuously differentiable in f, so an application of the mean value theorem gives:

$$\begin{aligned}\sup_{\boldsymbol{\theta}\in\Theta}&\left|\frac{1}{\sigma^4}h'\left(\frac{x_t-\hat{f}_t(\boldsymbol{\theta})}{\sigma};\boldsymbol{\psi}\right)-\frac{1}{\sigma^4}h'\left(\frac{x_t-\hat{f}_t(\boldsymbol{\theta})}{\sigma};\boldsymbol{\psi}\right)\right|\\ &\le \sup_{f\in\mathbb{R},\boldsymbol{\theta}\in\Theta}\left|\frac{1}{\sigma^5}h''\left(\frac{x_t-f}{\sigma};\boldsymbol{\psi}\right)\right||\hat{f}_t(\boldsymbol{\theta})-f_t(\boldsymbol{\theta})|\\ &\le D|\hat{f}_t(\boldsymbol{\theta})-f_t(\boldsymbol{\theta})| \overset{e.a.s.}{\to} 0\,,\end{aligned}$$

as $t\to\infty$, where $D=\sup_{\boldsymbol{\theta},z}|h''(z;\boldsymbol{\psi})/\sigma^5|$, which is finite because h'' is continuous and is zero for any $z<t_1$ and $z>t_k$ by construction and Θ is compact with $\sigma>0$.

For the remaining term, let $z_t(\boldsymbol{\theta})=(x_t-f_t(\boldsymbol{\theta}))/\sigma$ and $\hat{z}_t(\boldsymbol{\theta})=(x_t-\hat{f}_t(\boldsymbol{\theta}))/\sigma$. Then,

$$\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{\hat{z}_t(\boldsymbol{\theta})}{\sigma^4}h''\left(\hat{z}_t(\boldsymbol{\theta});\boldsymbol{\psi}\right)-\frac{z_t(\boldsymbol{\theta})}{\sigma^4}h''\left(z_t(\boldsymbol{\theta});\boldsymbol{\psi}\right)\right| \overset{e.a.s.}{\to} 0\,,$$

as $t\to\infty$ by Lemma TA.14 of Blasques et al. (2022), because clearly $\sup_{\boldsymbol{\theta}\in\Theta}|\hat{z}_t(\boldsymbol{\theta})-z_t(\boldsymbol{\theta})| \overset{e.a.s.}{\to} 0$ and $\sup_{\boldsymbol{\theta}\in\Theta}|h''\left(\hat{z}_t(\boldsymbol{\theta});\boldsymbol{\psi}\right)-h''\left(z_t(\boldsymbol{\theta});\boldsymbol{\psi}\right)| \overset{e.a.s.}{\to} 0$, by the uniform Lipschitz continuity of h'', and $z_t(\boldsymbol{\theta})$ and $h''\left(z_t(\boldsymbol{\theta});\boldsymbol{\psi}\right)$ have a bounded $\log^+$-moment uniformly over $\boldsymbol{\theta}\in\Theta$.

Finally, given that

$$\frac{\partial^2 s(f,x;\gamma)}{\partial\sigma^2\partial\sigma^2}=\frac{3}{4\sigma^5}h\left(\frac{x-f}{\sigma};\boldsymbol{\psi}\right)+\frac{5(x-f)}{4\sigma^6}h'\left(\frac{x-f}{\sigma};\boldsymbol{\psi}\right)\,,$$

the final convergence result follows from an application of the mean value theorem, the e.a.s. convergence of $\hat{f}_t(\boldsymbol{\theta})$ to $f_t(\boldsymbol{\theta})$ uniformly over Θ, and the observation that the derivative of the expression with respect to f is bounded:

$$\sup_{x\in\mathbb{R},f\in\mathbb{R},\boldsymbol{\theta}\in\Theta}\left|\frac{3}{4\sigma^6}h'\left(\frac{x-f}{\sigma};\boldsymbol{\psi}\right)+\frac{5(x-f)}{4\sigma^7}h''\left(\frac{x-f}{\sigma};\boldsymbol{\psi}\right)\right|=\tilde{D}\,,$$

for some finite constant $\tilde{D}$, due to $h'(z)$ and $zh''(z)$ being bounded in z and Θ being compact. □

Proof of Lemma B.4. To prove the result holds, it is sufficient to show that $\sup_{\boldsymbol{\theta}\in\Theta}|\hat{\ell}_t(\boldsymbol{\theta})-\ell_t(\boldsymbol{\theta})| \overset{e.a.s.}{\to} 0$ by Lemma 2.1 of Straumann and Mikosch (2006). The expression of $\ell_t(\boldsymbol{\theta})$ is given in (E.1). Only the first term, $-H((x_t-f_t(\boldsymbol{\theta}))/\sigma;\boldsymbol{\psi})$, contains $f_t(\boldsymbol{\theta})$, so we can disregard the other terms. By construction, $H(\cdot;\boldsymbol{\psi})$ is a continuously differentiable function. Thus, if we let $\hat{z}_t(\boldsymbol{\theta}) := (x_t - \hat{f}_t(\boldsymbol{\theta}))/\sigma$ and $z_t(\boldsymbol{\theta}) := (x_t - f_t(\boldsymbol{\theta}))/\sigma$ we can apply the mean value theorem to obtain:

$$\left|H\left(\hat{z}_t(\boldsymbol{\theta});\boldsymbol{\psi}\right) - H\left(z_t(\boldsymbol{\theta});\boldsymbol{\psi}\right)\right| \leq \left|h\left(z_t^*(\boldsymbol{\theta})\right)\right| \left|\hat{z}_t(\boldsymbol{\theta}) - z_t(\boldsymbol{\theta})\right| \,,$$

for some processes $\{\hat{z}_t(\boldsymbol{\theta})\}$ and $\{z_t(\boldsymbol{\theta})\}$, and where $z_t^*(\boldsymbol{\theta})$ is on the line segment between $\hat{z}_t(\boldsymbol{\theta})$ and $z_t(\boldsymbol{\theta})$. As h is also continuously differentiable, we can apply the mean value theorem to the first factor on the right-hand side, which gives

$$\begin{aligned}\left|h\left(z_t^*(\boldsymbol{\theta});\boldsymbol{\psi}\right)\right| &\leq \left|h\left(z_t(\boldsymbol{\theta});\boldsymbol{\psi}\right)\right| + \left|h'\left(z_t^{**}(\boldsymbol{\theta});\boldsymbol{\psi}\right)\right| \left|z_t^*(\boldsymbol{\theta}) - z_t(\boldsymbol{\theta})\right| \\ &\leq \left|h\left(z_t(\boldsymbol{\theta});\boldsymbol{\psi}\right)\right| + \sup_{z\in\mathbb{R}} \left|h'\left(z;\boldsymbol{\psi}\right)\right| \left|\hat{z}_t(\boldsymbol{\theta}) - z_t(\boldsymbol{\theta})\right| \,,\end{aligned}$$

for some value $z_t^{**}(\boldsymbol{\theta})$ on the line segment between $z_t^*(\boldsymbol{\theta})$ and $z_t(\boldsymbol{\theta})$, and where we use for the second inequality that $z_t^*(\boldsymbol{\theta})$ is on the line segment between $z_t(\boldsymbol{\theta})$ and $\hat{z}_t(\boldsymbol{\theta})$. It therefore follows that:

$$\begin{aligned}&\sup_{\boldsymbol{\theta}\in\Theta} \left|H\left(\hat{z}_t(\boldsymbol{\theta});\boldsymbol{\psi}\right) - H\left(z_t(\boldsymbol{\theta});\boldsymbol{\psi}\right)\right| \\ &\qquad \leq \left(\sup_{\boldsymbol{\theta}\in\Theta} \left|h\left(z_t(\boldsymbol{\theta});\boldsymbol{\psi}\right)\right| + \sup_{\boldsymbol{\theta}\in\Theta}\sup_{z\in\mathbb{R}} \left|h'\left(z;\boldsymbol{\psi}\right)\right| \sup_{\boldsymbol{\theta}\in\Theta} \left|\hat{z}_t(\boldsymbol{\theta}) - z_t(\boldsymbol{\theta})\right|\right) \sup_{\boldsymbol{\theta}\in\Theta} \left|\hat{z}_t(\boldsymbol{\theta}) - z_t(\boldsymbol{\theta})\right| \,,\end{aligned}$$

which converges e.a.s. to zero as $t \to \infty$ as long as $\sup_{\boldsymbol{\theta}\in\Theta} |\hat{z}_t(\boldsymbol{\theta}) - z_t(\boldsymbol{\theta})| \overset{e.a.s.}{\to} 0$ as $t \to \infty$ and $\{z_t\}_{t\in\mathbb{Z}}$ is stationary with a bounded $\log^+$-moment. This follows because Θ is compact and $\sup_{\boldsymbol{\theta}\in\Theta}\sup_{z\in\mathbb{R}} |h'(z;\boldsymbol{\psi})| = D$ for some finite constant D due to the boundedness of h' for every $\boldsymbol{\psi} \in \boldsymbol{\Psi}$. The result then follows from Lemma 2.1 of Straumann and Mikosch (2006), which says that if $\xi_t \overset{e.a.s.}{\to} 0$ as $t \to \infty$ and $\{v_t\}$ is strictly stationary with $\mathbb{E}\log^+|v_t| < \infty$, then $v_t\xi_t \overset{e.a.s.}{\to} 0$ as $t \to \infty$.

Now it remains to be shown that $\{\hat{z}_t\}$ converges to $\{z_t\}$ e.a.s. and $\{z_t\}$ has the required properties. We have that

$$\begin{aligned}\sup_{\boldsymbol{\theta}\in\Theta} \left|\hat{z}_t(\boldsymbol{\theta}) - z_t(\boldsymbol{\theta})\right| &= \sup_{\boldsymbol{\theta}\in\Theta} \left|\frac{x_t - \hat{f}_t(\boldsymbol{\theta})}{\sigma} - \frac{x_t - f_t(\boldsymbol{\theta})}{\sigma}\right| \\ &\leq \sup_{\boldsymbol{\theta}\in\Theta} \frac{1}{\sigma} \cdot \sup_{\boldsymbol{\theta}\in\Theta} \left|\hat{f}_t(\boldsymbol{\theta}) - f_t(\boldsymbol{\theta})\right| \overset{e.a.s.}{\to} 0 \,,\end{aligned}$$

as $t \to \infty$, due to the compactness of Θ and $\sigma^2 > 0$ for any $\boldsymbol{\theta} \in \Theta$ by assumptions, and the convergence result of Proposition 2(i). Finally, $\{z_t(\boldsymbol{\theta})\}_{t\in\mathbb{Z}} = \{(x_t - f_t(\boldsymbol{\theta}))/\sigma\}_{t\in\mathbb{Z}}$ is clearly SE by Krengel (1985, Proposition 4.3) and it has bounded moments of any order due to x_t and $f_t(\boldsymbol{\theta})$ having bounded moments of any order by Propositions 1 and 2. This concludes the proof.

$\square$

Proof of Lemma B.5. See Equation (E.1) for an expression of ℓ_t. The elements $\ell_t(\cdot)$ are continuous functions from Θ to $\mathbb{R}$. Furthermore, $\{\ell_t(\boldsymbol{\theta})\}_{t\in\mathbb{Z}}$ is SE for every $\boldsymbol{\theta} \in \Theta$ by Krengel (1985, Proposition 4.3), as it is a continuous function of the elements of the SE sequences $\{x_t\}_{t\in\mathbb{Z}}$ and $\{f_t(\boldsymbol{\theta})\}_{t\in\mathbb{Z}}$.

Just as in Lemma TA.6 of Blasques et al. (2022), we apply the ergodic theorem for separable Banach spaces of Rao (1962) to $\{\ell_t(\cdot)\}_{t\in\mathbb{Z}}$, which requires us to show $\mathbb{E}\sup_{\boldsymbol{\theta}\in\Theta}|\ell_t(\boldsymbol{\theta})| < \infty$. As the last two terms of $\ell_t(\boldsymbol{\theta})$ are deterministic and time invariant, we only have to consider the first term here. We can see immediately that

$$E\sup_{\boldsymbol{\theta}\in\Theta}\left|H\left(\frac{x_t - f_t(\boldsymbol{\theta})}{\sigma};\boldsymbol{\psi}\right)\right| < \infty\,,$$

because *(i)* Θ is compact, *(ii)* $\sigma > 0$ for any $\boldsymbol{\theta} \in \Theta$, and *(iii)* $\mathbb{E}|x_t|^n < \infty$ and $\mathbb{E}\sup_{\boldsymbol{\theta}\in\Theta}|f_t(\boldsymbol{\theta})|^n < \infty$ for any $n > 0$ by Propositions 1 and 2, and H is a continuous piecewise quartic polynomial in $x_t - f_t(\boldsymbol{\theta})$. Thus, we can apply the ergodic theorem and the result follows.

□

Proof of Lemma B.6. It follows from the proof of Lemma B.5 that $\mathbb{E}[\ell_t(\boldsymbol{\theta})]$ exists and is bounded for any $\boldsymbol{\theta} \in \Theta$. It can be shown straightforwardly, see for instance Blasques et al. (2018, Theorem 4.1), that under correct specification, in case $\ell_t(\boldsymbol{\theta}) = \ell_t(\boldsymbol{\theta}_0)$ holds almost surely if and only if $\boldsymbol{\theta} = \boldsymbol{\theta}_0$, then $\boldsymbol{\theta}_0$ is the unique maximizer of $\mathbb{E}[\ell_t(\boldsymbol{\theta})]$.

For ease of exposition, we introduce the notation $\ell(f_t(\boldsymbol{\theta}),\boldsymbol{\gamma}) := \ell_t(\boldsymbol{\theta})$. We start by showing that $\ell(f_t(\boldsymbol{\theta}_0),\boldsymbol{\gamma}_0) = \ell(f_t(\boldsymbol{\theta}),\boldsymbol{\gamma})$ a.s. if and only if $f_t(\boldsymbol{\theta}_0) = f_t(\boldsymbol{\theta})$ a.s. and $\boldsymbol{\gamma} = \boldsymbol{\gamma}_0$. Notice that $f_t(\boldsymbol{\theta}_0)$ is the true time-varying parameter, as we know from Proposition 2 that $\{f_t(\boldsymbol{\theta}_0)\}_{t\in\mathbb{Z}}$ is the unique limit sequence of $\{\hat{f}_t(\boldsymbol{\theta}_0)\}_{t\in\mathbb{N}}$, which implies that $x_t = f_t(\boldsymbol{\theta}_0) + \sigma_0\varepsilon_t$.

First, we notice that because each vector $\boldsymbol{\psi} \in \boldsymbol{\Psi}$ leads to a unique cubic spline function $h(\cdot;\boldsymbol{\psi})$, it also leads to a unique anti-derivative $H(\cdot;\boldsymbol{\psi})$. Despite to the anti-derivative H for a given $\boldsymbol{\psi}$ only being unique up to an additive constant, it is clear that for $a, a' \in \mathbb{R}$ and $\boldsymbol{\psi}, \boldsymbol{\psi}' \in \boldsymbol{\Psi}$, we have $a + H(z;\boldsymbol{\psi}) = a' + H(z;\boldsymbol{\psi}')$ for all $z \in \mathbb{R}$ if and only if $(a,\boldsymbol{\psi}) = (a',\boldsymbol{\psi}')$.

Building on this, we notice that for any $a, a', c, c' \in \mathbb{R}$, $b, b' > 0$, and $\boldsymbol{\psi}, \boldsymbol{\psi}' \in \boldsymbol{\Psi}$, $a + H((z-c)/b;\boldsymbol{\psi}) = a' + H((z-c')/b';\boldsymbol{\psi}')$ for any $z \in \mathbb{R}$ if and only if $(a,b,c,\boldsymbol{\psi}) = (a',b',c',\boldsymbol{\psi}')$. This follows because the division by b essentially leads to a scaling of the knots $\boldsymbol{t}$, while c effectively yields a shift of the knots, as it can be seen that $h((z -$

$c)/b; \boldsymbol{t}, \boldsymbol{y}) = h(z - c; b\boldsymbol{t}, \boldsymbol{y}) = h(z; b\boldsymbol{t} + c, \boldsymbol{y})$. Due to the rotational symmetry around the origin that is imposed on the spline h, it is clear that a shift of the knots will always lead to an inherently different spline function, which implies that we must have $c = c'$ for the equality above to hold. Furthermore, from the second part of Assumption 4, it follows that $h((z-c)/b; \boldsymbol{\psi}) = h((z-c)/b'; \boldsymbol{\psi}')$ for any $z \in \mathbb{R}$ if and only if $(b, \boldsymbol{\psi}) = (b, \boldsymbol{\psi}')$. This concludes the proof for the statement above.

Therefore, if $\varepsilon_1 \sim p_\varepsilon(\varepsilon_1; \boldsymbol{\psi})$, where $p_\varepsilon(\cdot; \boldsymbol{\psi})$ denotes the spline-based density for parameter $\boldsymbol{\psi} \in \boldsymbol{\Psi}$, then $a' + H((\varepsilon_1 - c')/b'; \boldsymbol{\psi}') = a'' + H((\varepsilon_1 - c'')/b''; \boldsymbol{\psi}'')$ a.s. if and only if $(a', b', c', \boldsymbol{\psi}') = (a'', b'', c'', \boldsymbol{\psi}'')$, as ε_1 is an absolutely continuous random variable with support on the entire real line. Thus, as $\ell(f_t(\boldsymbol{\theta}), \boldsymbol{\gamma}) = \ell(f_t(\boldsymbol{\theta}_0), \boldsymbol{\gamma}_0)$ a.s. if and only if

$$\begin{aligned} -H\left(\frac{f_t(\boldsymbol{\theta}_0) - f_t(\boldsymbol{\theta}) + \sigma_0\varepsilon_t}{\sigma}; \boldsymbol{\psi}\right) - \log C(\boldsymbol{\psi}) - \frac{1}{2}\log\sigma^2 \\ = -H\left(\varepsilon_t; \boldsymbol{\psi}_0\right) - \log C(\boldsymbol{\psi}_0) - \frac{1}{2}\log\sigma_0^2 \,, \end{aligned}$$

holds almost surely, it follows that $\ell(f_t(\boldsymbol{\theta}), \boldsymbol{\gamma}) = \ell(f_t(\boldsymbol{\theta}_0), \boldsymbol{\gamma}_0)$ a.s. if and only if $\boldsymbol{\gamma} = \boldsymbol{\gamma}_0$ and $f_t(\boldsymbol{\theta}_0) = f_t(\boldsymbol{\theta})$ a.s.. To show that $\ell_t(\boldsymbol{\theta}) = \ell_t(\boldsymbol{\theta}_0)$ a.s. if and only if $\boldsymbol{\theta} = \boldsymbol{\theta}_0$, it now only remains to be shown that given $\boldsymbol{\gamma} = \boldsymbol{\gamma}_0$, $f_t(\boldsymbol{\theta}) = f_t(\boldsymbol{\theta}_0)$ a.s. if and only if $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. Because $\{f_t(\boldsymbol{\theta})\}_{t\in\mathbb{Z}}$ is SE for any $\boldsymbol{\theta} \in \Theta$, then if $f_t(\boldsymbol{\theta}) = f_t(\boldsymbol{\theta}_0)$ a.s. for some $t \in \mathbb{Z}$, this equality holds almost surely for any $t \in \mathbb{Z}$. Let us assume that $\boldsymbol{\theta} \in \Theta$ is such that $f_t(\boldsymbol{\theta}) = f_t(\boldsymbol{\theta}_0)$ a.s., then it follows from the form of the updating equation that

$$f_{t+1}(\boldsymbol{\theta}_0) - f_{t+1}(\boldsymbol{\theta}) = \omega_0(1-\beta_0) - \omega(1-\beta) + (\beta_0 - \beta)f_t(\boldsymbol{\theta}_0) + (\alpha_0 - \alpha)\frac{1}{\sigma_0}h\left(\varepsilon_t; \boldsymbol{\psi}_0\right) \,,$$

where $\varepsilon_t \sim p_\varepsilon(\varepsilon_t; \boldsymbol{\psi}_0)$. We start by proving that to have $f_t(\boldsymbol{\theta}_0) = f_t(\boldsymbol{\theta})$ a.s., we must have $\omega(1-\beta) = \omega_0(1-\beta_0)$. We use a proof by contradiction. Say $\omega(1-\beta) \neq \omega_0(1-\beta_0)$ and $f_{t+1}(\boldsymbol{\theta}_0) - f_{t+1}(\boldsymbol{\theta}) = 0$ a.s., then we must have $(\beta_0 - \beta)f_t(\boldsymbol{\theta}_0) + (\alpha_0 - \alpha)\frac{1}{\sigma_0}h\left(\varepsilon_t; \boldsymbol{\psi}_0\right) = \omega(1-\beta) - \omega_0(1-\beta_0) \neq 0$ a.s. As the first and second term of the expression on the left hand side are independent, this equality can only hold a.s. if both terms are constants and at least one is different from zero. However, neither of these terms can be nonzero constants, because $f_t(\boldsymbol{\theta}_0)$ is non-deterministic due to $\alpha_0 \neq 0$ by Assumption 4 and $h(\varepsilon_t; \boldsymbol{\psi}_0)$ is also clearly non-deterministic under the conditions imposed on $\boldsymbol{\Psi}$. Therefore, we must have $\omega(1-\beta) = \omega_0(1-\beta_0)$. Secondly, we show that $\beta = \beta_0$ is needed to have $f_t(\boldsymbol{\theta}_0) = f_t(\boldsymbol{\theta})$ a.s.. If by contradiction we say that $\beta \neq \beta_0$, then to have $f_{t+1}(\boldsymbol{\theta}_0) - f_{t+1}(\boldsymbol{\theta}) = 0$ a.s., we must have $(\alpha_0 - \alpha)\frac{1}{\sigma_0}h(\varepsilon; \boldsymbol{\psi}_0) = (\beta - \beta_0)f_t(\boldsymbol{\theta}_0) \neq 0$ a.s., which is ruled out by $h(\varepsilon; \boldsymbol{\psi}_0)$ and $f_t(\boldsymbol{\theta}_0)$ being independent and non-deterministic. Therefore, we must also have $\beta = \beta_0$, which together with $\omega(1-\beta) = \omega_0(1-\beta_0)$ and the assumption that $|\beta_0| < 1$, implies that we must have $\omega = \omega_0$. Finally, we must clearly also have $\alpha = \alpha_0$, because this is the only way to have $(\alpha_0 - \alpha)\frac{1}{\sigma_0} h\left(\varepsilon_t; \boldsymbol{\psi}_0\right) = 0$ a.s.. This finishes the proof.

□

Proof of Lemma B.7. The expression of $\ell_t''(\boldsymbol{\theta}) = \partial^2 \ell_t(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top$ is given in Section E.2. By the sub-additivity of the norm $\sup_{\boldsymbol{\theta}\in\Theta} \|\cdot\|$, it suffices to show that the expectation of this norm for each term in this expression is bounded. That all these expectations are bounded follows readily from the following reasons: *(i)* Θ is a compact set such that for any $\boldsymbol{\theta} \in \Theta$, $\sigma > 0$ and $\boldsymbol{\psi}$ is such that $\exp(-H(x; \boldsymbol{\psi}))$ is integrable, *(ii)* by Proposition 1 we have $\mathbb{E}|x_t|^n < \infty$ and by Proposition we have 2 $\mathbb{E}\sup_{\boldsymbol{\theta}\in\Theta} |f_t(\boldsymbol{\theta})| < \infty$, $\mathbb{E}\sup_{\boldsymbol{\theta}\in\Theta} |\partial f_t(\boldsymbol{\theta})/\partial\boldsymbol{\theta}|^n < \infty$ and $\mathbb{E}\sup_{\boldsymbol{\theta}\in\Theta} |\partial^2 f_t(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top|^n < \infty$ for any $n > 0$, and *(iii)* $h(\cdot; \boldsymbol{\psi})$ being a piecewise cubic function and $h'(\cdot; \boldsymbol{\psi})$ being a piecewise quadratic function for any $\boldsymbol{y} \in \mathbb{R}^k$, and *(iv)* the assumptions on $\boldsymbol{\Psi}$ that ensure that for any $\boldsymbol{\psi} \in \boldsymbol{\Psi}$, $H(x; \boldsymbol{\psi})$ diverges at least linearly as $|x| \to \infty$, which implies that $\inf_{\boldsymbol{\psi}\in\boldsymbol{\Psi}} C(\boldsymbol{\psi}) > 0$ and $\sup_{\boldsymbol{\psi}\in\boldsymbol{\Psi}} \|\partial C(\boldsymbol{\psi})/\partial\boldsymbol{\psi}\| < \infty$ and $\sup_{\boldsymbol{\psi}\in\boldsymbol{\Psi}} \|\partial^2 C(\boldsymbol{\psi})/\partial\boldsymbol{\psi}\partial\boldsymbol{\psi}^\top\| < \infty$. For the terms that contain products of random variables, the finiteness of the expectation of the supremum follows from the sub-multiplactivity of the norm and the Cauchy-Schwarz inequalty.

□

Proof of Lemma B.8. We use an approach similar to the proof of Lemma A.5 of Gorgi and Koopman (2023). Due to the correct specification assumption and the existence of a finite moment of the second order derivative of the log likelihood function shown in Lemma B.7, we can use the Fisher information matrix equality $-\mathbb{E}[\ell_t''(\boldsymbol{\theta}_0)] = \mathbb{E}[\ell_t'(\boldsymbol{\theta}_0)\ell_t'(\boldsymbol{\theta}_0)^\top]$. See Equation E.2 for the expression of $\ell_t(\boldsymbol{\theta})$. Because $\mathbb{E}[\ell_t'(\boldsymbol{\theta}_0)\ell_t'(\boldsymbol{\theta}_0)^\top]$ is positive semi-definite by construction, we only have to show that it is full rank, which can be achieved by proving that

$$\boldsymbol{v}^\top \ell_t'(\boldsymbol{\theta}_0) = \boldsymbol{v}^\top \left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ -\frac{1}{2\sigma_0^2}\left(1 - \varepsilon_t h\left(\varepsilon_t; \boldsymbol{\psi}_0\right)\right) \\ -\frac{\partial H(\varepsilon_t; \boldsymbol{\psi}_0)}{\partial \boldsymbol{\psi}} - \frac{1}{C(\boldsymbol{\psi}_0)}\frac{\partial C(\boldsymbol{\psi}_0)}{\partial \boldsymbol{\psi}} \end{pmatrix} + \begin{pmatrix} \frac{\partial f_t(\boldsymbol{\theta}_0)}{\partial \omega} \\ \frac{\partial f_t(\boldsymbol{\theta}_0)}{\partial \beta} \\ \frac{\partial f_t(\boldsymbol{\theta}_0)}{\partial \alpha} \\ \frac{\partial f_t(\boldsymbol{\theta}_0)}{\partial \sigma^2} \\ \frac{\partial f_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\psi}} \end{pmatrix} \frac{1}{\sigma_0} h\left(\varepsilon_t; \boldsymbol{\psi}_0\right) \right] = 0 \qquad \text{a.s.}, \tag{D.1}$$

if and only if $\boldsymbol{v} = 0$, where $\boldsymbol{v} \in \mathbb{R}^{4+q}$. Furthermore, we use that $f_t(\boldsymbol{\theta}_0)$ is almost surely equal to the true value of f_t in $x_t = f_t + \sigma_0 \varepsilon_t$ due to the correct specification assumption and due to $\{f_t(\boldsymbol{\theta}_0)\}_{t\in\mathbb{Z}}$ being the unique limit sequence of $\{\hat{f}_t(\boldsymbol{\theta}_0)\}_{t\in\mathbb{N}}$ for any $\boldsymbol{\theta} \in \Theta$ by Proposition 2.

Let us split the vector $\boldsymbol{v}$ into $\boldsymbol{v} = (\boldsymbol{v}_1^\top, v_2, \boldsymbol{v}_3^\top)^\top$, where $\boldsymbol{v}_1 \in \mathbb{R}^3$, $v_2 \in \mathbb{R}$ and $\boldsymbol{v}_3 \in \mathbb{R}^q$. Because $h(\varepsilon_t; \boldsymbol{\psi}_0) \neq 0$ a.s., with probability one we have that (D.1) implies

$$\boldsymbol{v}_1^\top \begin{pmatrix} \frac{\partial f_t(\boldsymbol{\theta}_0)}{\partial \omega} \\ \frac{\partial f_t(\boldsymbol{\theta}_0)}{\partial \beta} \\ \frac{\partial f_t(\boldsymbol{\theta}_0)}{\partial \alpha} \end{pmatrix} + \begin{pmatrix} v_2 \\ \boldsymbol{v}_3 \end{pmatrix}^\top \begin{pmatrix} \frac{\partial f_t(\boldsymbol{\theta}_0)}{\partial \sigma^2} \\ \frac{\partial f_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\psi}} \end{pmatrix} = \frac{\sigma_0}{h(\varepsilon_t; \boldsymbol{\psi}_0)} \begin{pmatrix} v_2 \\ \boldsymbol{v}_3 \end{pmatrix}^\top \begin{pmatrix} \frac{1}{2\sigma_0^2}\left(1 - \varepsilon_t h\left(\varepsilon_t; \boldsymbol{\psi}_0\right)\right) \\ \frac{\partial H(\varepsilon_t; \boldsymbol{\psi}_0)}{\partial \boldsymbol{\psi}} + \frac{1}{C(\boldsymbol{\psi}_0)} \frac{\partial C(\boldsymbol{\psi}_0)}{\partial \boldsymbol{\psi}} \end{pmatrix}. \tag{D.2}$$

Notice that if $\boldsymbol{v}^\top \ell_t(\boldsymbol{\theta}_0) = 0$ a.s. for $\boldsymbol{v} \neq 0$, then we must be in one of the following cases: *(i)* $(v_2, \boldsymbol{v}_3^\top)^\top \neq 0$, or *(ii)* $\boldsymbol{v}_1 \neq 0$, $v_2 = 0$ and $\boldsymbol{v}_3 = 0$. We will now prove that neither of these options can lead to $\boldsymbol{v}^\top \ell_t(\boldsymbol{\theta}_0) = 0$ a.s. We start by noting that the terms on the left-hand side of (D.2) are $\mathcal{G}_{t-1}$-measurable, while those on the right-hand side are not. Thus, this equality can only hold almost surely if both sides are equal to some constant almost surely.

We first consider option *(i)*; $(v_2, \boldsymbol{v}_3^\top)^\top \neq 0$. Due to the linear dependence of the cubic spline coefficients on the vector of ordinates $\boldsymbol{y}$ and the linearity of $\tau(\boldsymbol{\psi})$, we can write for any $\boldsymbol{\psi} \in \mathbb{R}^q$ and any $z \in \mathbb{R}$:

$$H(z; \boldsymbol{\psi}) = \boldsymbol{\psi}^\top \begin{pmatrix} H(z; \boldsymbol{e}_1) \\ H(z; \boldsymbol{e}_2) \\ \vdots \\ H(z; \boldsymbol{e}_q) \end{pmatrix} \quad \text{such that} \quad \frac{\partial H(z; \boldsymbol{\psi})}{\partial \boldsymbol{\psi}} = \begin{pmatrix} H(z; \boldsymbol{e}_1) \\ H(z; \boldsymbol{e}_2) \\ \vdots \\ H(z; \boldsymbol{e}_q) \end{pmatrix}, \tag{D.3}$$

where $\boldsymbol{e}_i$ for $i = 1, \dots, q$ denotes a q-dimensional vector with a 1 on entry i and zeroes elsewhere. It follows that $\boldsymbol{v}_3^\top \partial H(\varepsilon_t; \boldsymbol{\psi}_0)/\partial \boldsymbol{\psi} = H(\varepsilon_t; \boldsymbol{v}_3)$. We simplify the right-hand side of (D.2) as follows

$$\frac{\sigma_0}{h(\varepsilon_t; \boldsymbol{\psi}_0)} \left(\frac{v_2}{2\sigma_0^2} + H(\varepsilon_t; \boldsymbol{v}_3) + \frac{1}{C(\boldsymbol{\psi}_0)} \boldsymbol{v}_3^\top \frac{\partial C(\boldsymbol{\psi}_0)}{\partial \boldsymbol{\psi}} \right) - v_2 \frac{1}{2\sigma_0} \varepsilon_t \, .$$

In order to argue that there exists no nonzero vector $(v_2, \boldsymbol{v}_3^\top)^\top$ for which this expression is constant a.s., it suffices to show that for any constants $B, D \in \mathbb{R}$ the following equality only holds a.s. for for $v_2 = 0$ and $\boldsymbol{v}_3 = 0$:

$$B + H(\varepsilon_t; \boldsymbol{v}_3) = \left(D + v_2 \frac{1}{2\sigma_0^2} \varepsilon_t \right) h(\varepsilon_t; \boldsymbol{\psi}_0).$$

We notice that both functions $(D + dz)h(z; \boldsymbol{\psi}_0)$ for $d = v_2/2\sigma_0^2$, and $B + H(z; \boldsymbol{v}_3)$ are piecewise quartic polynomials in $z \in \mathbb{R}$, but they are inherently different from each other, as the former is only twice continuously differentiable under the assumptions on $\boldsymbol{\Psi}$, while the latter is three times continuously differentiable. In particular, their derivatives are different, as $h(z; \boldsymbol{v}_3)$ is a cubic spline function by construction while, $dh(z; \boldsymbol{\psi}_0) + (D + dz)h'(z; \boldsymbol{\psi}_0)$ is not, as $\boldsymbol{\Psi}$ is such that h'' is non-differentiable at some of the knots $\boldsymbol{t}$ under Assumption 4. This holds even if $d = 0$. Therefore, given that the density function p_ε is nonzero on $\mathbb{R}$, there exist no $(v_2, \boldsymbol{v}_3^\top)^\top \neq 0$ and $B, D \in \mathbb{R}$ such that the equality above holds almost surely. Thus, to have that the right-hand side of (D.2) is equal to a constant a.s., we must have $v_2 = 0$ and $\boldsymbol{v}_3 = 0$, in which case it is equal to zero a.s.

Finally, for case *(ii)*, we would need

$$\frac{1}{\sigma_0} h(\varepsilon_t; \boldsymbol{\psi}_0) \boldsymbol{v}_1^\top \begin{pmatrix} \frac{\partial f_t(\boldsymbol{\theta}_0)}{\partial \omega} \\ \frac{\partial f_t(\boldsymbol{\theta}_0)}{\partial \beta} \\ \frac{\partial f_t(\boldsymbol{\theta}_0)}{\partial \alpha} \end{pmatrix} = 0 \quad \text{a.s.}\,,$$

for $\boldsymbol{v}_1 \neq 0$. First of all, we know that $h(\varepsilon_t; \boldsymbol{\psi}_0)/\sigma_0$ is non-degenerate and independent of the rest of the expression. Thus, for the equality to hold almost surely, we must have that the inner product next to it, is equal to zero almost surely. Looking at the expressions of the equations that describe these derivative processes:

$$\begin{aligned}
\frac{\partial f_{t+1}(\boldsymbol{\theta}_0)}{\partial \omega} &= B_t \frac{\partial f_t(\boldsymbol{\theta}_0)}{\partial \omega} + 1 - \beta_0\,, \\
\frac{\partial f_{t+1}(\boldsymbol{\theta}_0)}{\partial \beta} &= B_t \frac{\partial f_t(\boldsymbol{\theta}_0)}{\partial \beta} - \omega_0 + \beta_0 f_t(\boldsymbol{\theta}_0)\,, \\
\frac{\partial f_{t+1}(\boldsymbol{\theta}_0)}{\partial \alpha} &= B_t \frac{\partial f_t(\boldsymbol{\theta}_0)}{\partial \alpha} + s(f_t(\boldsymbol{\theta}_0), x_t)\,, \\
\text{where} \quad B_t &= \beta_0 + \alpha_0 \left.\frac{\partial s(f, x_t)}{\partial f}\right|_{f = f_t(\boldsymbol{\theta}_0)}\,,
\end{aligned}$$

we see that $\partial f_t(\boldsymbol{\theta}_0)/\partial\omega$, $\partial f_t(\boldsymbol{\theta}_0)/\partial\beta$, and $\partial f_t(\boldsymbol{\theta}_0)/\partial\alpha$, are elements of autoregressive processes that are identical, except for their 'innovations' being equal to $1-\beta_0$, $-\omega_0+f_{t-1}(\boldsymbol{\theta}_0)$ and $h(\varepsilon_{t-1}; \boldsymbol{\psi}_0)/\sigma_0$, respectively. Clearly, none of the three processes is degenerate and their 'innovations' take different values with probability one, as 1 is a constant and the remaining two 'innovations' are continuous random variables that are independent of each other. It follows that case (iv) is also ruled out, which concludes the proof.

□

Proof of Lemma B.9. We apply the central limit theorem of Billingsley (1999) for SE martingale difference sequences to $\{\ell_t'(\boldsymbol{\theta}_0)\}_{t\in\mathbb{Z}}$, which requires us to show that this sequence is an SE martingale difference sequence with $\mathbb{E}|\ell_t'(\boldsymbol{\theta}_0)|^2 < \infty$. An application of Krengel (1985, Proposition 4.3) tells us that $\{\ell_t'(\boldsymbol{\theta}_0)\}_{t\in\mathbb{Z}}$ is SE, as $\ell_t'(\boldsymbol{\theta}_0)$ is a continuous function of $f_t(\boldsymbol{\theta}_0)$, $\partial f_t(\boldsymbol{\theta}_0)/\partial\boldsymbol{\theta}$ and x_t, which are elements of SE sequences by Propositions 1 and 2. That $\mathbb{E}|\ell_t'(\boldsymbol{\theta}_0)|^2 < \infty$ can be seen straightforwardly from the expression of $\ell_t'(\boldsymbol{\theta}_0)$ that can be found in (D.1), as it is known that ε_t and $\partial f_t(\boldsymbol{\theta}_0)/\partial\boldsymbol{\theta}$ have bounded moments of any order, and h and H are piecewise polynomials. Finally, $\{\ell_t'(\boldsymbol{\theta}_0)\}_{t\in\mathbb{Z}}$ is a martingale difference sequence, because using the notation from the proof of Lemma B.8, we can write

$$\mathbb{E}\left[\ell_t'(\boldsymbol{\theta}_0)|\mathcal{G}_{t-1}\right] = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -\frac{1}{2\sigma_0^2}\left(1-\mathbb{E}[\varepsilon_t h\left(\varepsilon_t;\boldsymbol{\psi}_0\right)]\right) \\ -\mathbb{E}\left[\frac{\partial H(\varepsilon_t;\boldsymbol{\psi}_0)}{\partial\boldsymbol{\psi}}\right] - \frac{1}{C(\boldsymbol{\psi}_0)}\frac{\partial C(\boldsymbol{\psi}_0)}{\partial\boldsymbol{\psi}} \end{pmatrix} + \begin{pmatrix} \frac{\partial f_t(\boldsymbol{\theta}_0)}{\partial\omega} \\ \frac{\partial f_t(\boldsymbol{\theta}_0)}{\partial\beta} \\ \frac{\partial f_t(\boldsymbol{\theta}_0)}{\partial\alpha} \\ \frac{\partial f_t(\boldsymbol{\theta}_0)}{\partial\sigma^2} \\ \frac{\partial f_t(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\psi}} \end{pmatrix} \frac{1}{\sigma_0}\mathbb{E}[h(\varepsilon_t;\boldsymbol{\psi}_0)] = 0\,,$$

where we use that the derivative $\partial f_t(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$ is $\mathcal{G}_{t-1}$ measurable and that $\mathbb{E}[h(\varepsilon_t;\boldsymbol{\psi}_0)] = 0$, because h is rotationally symmetric around zero by construction, which implies that p_ε is a symmetric distribution. Furthermore, $\mathbb{E}[\varepsilon_t h(\varepsilon_t;\boldsymbol{\psi}_0)] = 1$ follows from:

$$\begin{aligned}\mathbb{E}[\varepsilon_t h(\varepsilon_t;\boldsymbol{\psi}_0)] &= \int_{-\infty}^{\infty} x h(x;\boldsymbol{\psi}_0) p_\varepsilon(x)\,\mathrm{d}x = -\int_{-\infty}^{\infty} x\frac{\partial \log p_\varepsilon(x)}{\partial x} p_\varepsilon(x)\,\mathrm{d}x \\ &= -\int_{-\infty}^{\infty} x p_\varepsilon'(x)\,\mathrm{d}x = -\left[x p_\varepsilon(x) - \mathbb{P}_{X\sim p_\varepsilon}(X\le x)\right]_{-\infty}^{\infty} = 1\,.\end{aligned}$$

Finally, for $i = 1,\dots,q$, it can be seen using the expressions in Section E.2 that

$$\mathbb{E}\left[\frac{\partial H(\varepsilon_t;\boldsymbol{\psi}_0)}{\partial\psi_i}\right] = \int_{-\infty}^{\infty} H(x;\boldsymbol{e}_i)\frac{\exp(-H(x;\boldsymbol{\psi}_0))}{C(\boldsymbol{\psi}_0)}\,\mathrm{d}x = -\frac{1}{C(\boldsymbol{\psi}_0)}\frac{\partial C(\boldsymbol{\psi}_0)}{\partial\psi_i}\,.$$

Thus, $\{\ell_t'(\boldsymbol{\theta}_0)\}_{t\in\mathbb{Z}}$ is indeed a martingale difference sequence and the central limit theorem can be applied, which together with the Fisher information matrix equality, which was argued to hold in the proof of Lemma B.8, concludes the proof.

□

Proof of Lemma B.10. By Lemma 2.1 of Straumann and Mikosch (2006) and the subadditivity of the sup-norm, it suffices to show that

$$\sup_{\boldsymbol{\theta}\in\Theta}\|\hat{\ell}_t'(\boldsymbol{\theta}) - \ell_t'(\boldsymbol{\theta})\| \overset{e.a.s.}{\to} 0 \quad \text{as } t\to\infty\,.$$

By considering the expression of the first order derivative in (E.2), and using that $\inf_{\boldsymbol{\theta}\in\Theta} 1/\sigma > 0$ under the imposed assumptions, it is clear that the following convergence results are sufficient for this result to hold:

$$\sup_{\boldsymbol{\theta}\in\Theta}\left\| h\left(\frac{x_t-\hat{f}_t(\boldsymbol{\theta})}{\sigma};\boldsymbol{\psi}\right)\frac{\partial \hat{f}_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} - h\left(\frac{x_t-f_t(\boldsymbol{\theta})}{\sigma};\boldsymbol{\psi}\right)\frac{\partial f_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right\| \overset{e.a.s.}{\to} 0\,,$$

$$\sup_{\boldsymbol{\theta}\in\Theta}\left| h\left(\frac{x_t-\hat{f}_t(\boldsymbol{\theta})}{\sigma};\boldsymbol{\psi}\right)(x_t-\hat{f}_t(\boldsymbol{\theta})) - h\left(\frac{x_t-f_t(\boldsymbol{\theta})}{\sigma};\boldsymbol{\psi}\right)(x_t-f_t(\boldsymbol{\theta}))\right| \overset{e.a.s.}{\to} 0\,,$$

$$\sup_{\boldsymbol{\theta}\in\Theta}\left\| \left.\frac{\partial H(z;\boldsymbol{\psi})}{\partial\boldsymbol{\psi}}\right|_{z=(x_t-\hat{f}_t(\boldsymbol{\theta}))/\sigma} - \left.\frac{\partial H(z;\boldsymbol{\psi})}{\partial\boldsymbol{\psi}}\right|_{z=(x_t-f_t(\boldsymbol{\theta}))/\sigma}\right\| \overset{e.a.s.}{\to} 0\,, \tag{D.4}$$

as $t\to\infty$. The first two results follow from applications of Corollary TA.16 of Blasques et al. (2022), which tells us that if $\{\hat{z}_t(\boldsymbol{\theta})\}_{t\in\mathbb{N}}$ and $\{\hat{y}_t(\boldsymbol{\theta})\}_{t\in\mathbb{N}}$ converge e.a.s. to SE limit sequences $\{z_t(\boldsymbol{\theta})\}_{t\in\mathbb{Z}}$ and $\{y_t(\boldsymbol{\theta})\}_{t\in\mathbb{Z}}$ as $t\to\infty$, respectively, uniformly over Θ, then in case $\mathbb{E}\sup_{\boldsymbol{\theta}\in\Theta}\log^+|z_t(\boldsymbol{\theta})|<\infty$ and $\mathbb{E}\sup_{\boldsymbol{\theta}\in\Theta}\log^+|y_t(\boldsymbol{\theta})|<\infty$, $\sup_{\boldsymbol{\theta}\in\Theta}|\hat{z}_t(\boldsymbol{\theta})\hat{y}_t(\boldsymbol{\theta})-z_t(\boldsymbol{\theta})y_t(\boldsymbol{\theta})| \overset{e.a.s.}{\to} 0$ as $t\to\infty$. It is not hard to see that this result can be extended to the case where either $\{\hat{z}_t(\boldsymbol{\theta})\}_{t\in\mathbb{N}}$ or $\{\hat{y}_t(\boldsymbol{\theta})\}_{t\in\mathbb{N}}$ is a vector process and the absolute value is replaced by a norm $\|\cdot\|$. Thus, here it suffices to show that

$$\sup_{\boldsymbol{\theta}\in\Theta}\left|\hat{f}_t(\boldsymbol{\theta})-f_t(\boldsymbol{\theta})\right| \overset{e.a.s.}{\to} 0\,,$$

$$\sup_{\boldsymbol{\theta}\in\Theta}\left\|\frac{\partial \hat{f}_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}-\frac{\partial f_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right\| \overset{e.a.s.}{\to} 0\,,$$

$$\sup_{\boldsymbol{\theta}\in\Theta}\left| h\left(\frac{x_t-\hat{f}_t(\boldsymbol{\theta})}{\sigma};\boldsymbol{\psi}\right) - h\left(\frac{x_t-f_t(\boldsymbol{\theta})}{\sigma};\boldsymbol{\psi}\right)\right| \overset{e.a.s.}{\to} 0\,,$$

as $t\to\infty$ and that the processes $\{f_t(\boldsymbol{\theta})\}_{t\in\mathbb{Z}}$, $\{\partial f_t(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\}_{t\in\mathbb{Z}}$ and $\{h((x_t-f_t(\boldsymbol{\theta})/\sigma;\boldsymbol{\psi}))\}_{t\in\mathbb{Z}}$ are SE and have a finite $\log^+$-moment uniformly over Θ. For the first two processes, all the required results follow readily from Proposition 2. For the last process, the stationarity and ergodicity follows from Krengel (1985, Proposition 4.3), as it is a continuous function of SE sequences. Furthermore, the finite $\log^+$ moment follows from the assumptions on Θ, the fact that x_t and $f_t(\boldsymbol{\theta})$ have been shown to have finite moments of any order uniformly over Θ and h being a piecewise third degree polynomial. Finally, for the convergence result, we can apply the mean value theorem, as $h(\cdot;\boldsymbol{\psi})$ is continuously differentiable and has a uniformly bounded derivative under the current assumptions:

$$\begin{aligned}&\sup_{\boldsymbol{\theta}\in\Theta}\left| h\left(\frac{x_t-\hat{f}_t(\boldsymbol{\theta})}{\sigma};\boldsymbol{\psi}\right) - h\left(\frac{x_t-f_t(\boldsymbol{\theta})}{\sigma};\boldsymbol{\psi}\right)\right| \\ &\qquad\le \underbrace{\sup_{\boldsymbol{\theta}\in\Theta}\sup_{z\in\mathbb{R}}\left|\frac{h'(z;\boldsymbol{\psi})}{\sigma}\right|}_{=D<\infty}\cdot\sup_{\boldsymbol{\theta}\in\Theta}\left|\hat{f}_t(\boldsymbol{\theta})-f_t(\boldsymbol{\theta})\right| \overset{e.a.s.}{\to} 0\,.\end{aligned}$$

Now it only remains to be shown that (D.4) holds, which can be shown using the same approach as is used in the proof of Lemma B.4, because $\partial H(z;\boldsymbol{\psi})/\partial\psi_i = H(z;\boldsymbol{e}_i)$ for $i = 1,\ldots,q$.

$\square$

Proof of Lemma B.11. Using the same argument as in the proof of Lemma B.4, it is sufficient to show that $\sup_{\boldsymbol{\theta}\in\Theta}|\hat{\ell}_t(\boldsymbol{\theta}) - \ell_t(\boldsymbol{\theta})| \overset{e.a.s.}{\to} 0$ as $t\to\infty$. The expression of $\ell_t(\boldsymbol{\theta})$ is given by

$$\ell_t(\boldsymbol{\theta}) = -H\left(\frac{x_t - m}{\exp(f_t(\boldsymbol{\theta}))};\boldsymbol{\psi}\right) - \log C(\boldsymbol{\psi}) - f_t(\boldsymbol{\theta})\,. \tag{D.5}$$

As only the first and last term depend on $f_t(\boldsymbol{\theta})$, the middle term can be disregarded. For the final term we have

$$\sup_{\boldsymbol{\theta}\in\Theta}|\hat{f}_t(\boldsymbol{\theta}) - f_t(\boldsymbol{\theta})| \overset{e.a.s.}{\to} 0 \quad \text{as } t\to\infty\,,$$

which follows directly from Proposition 3. Thus, it only remains to be shown that

$$\sup_{\boldsymbol{\theta}\in\Theta}\left|H\left(\frac{x_t - m}{\exp(\hat{f}_t(\boldsymbol{\theta}))};\boldsymbol{\psi}\right) - H\left(\frac{x_t - m}{\exp(f_t(\boldsymbol{\theta}))};\boldsymbol{\psi}\right)\right| \overset{e.a.s.}{\to} 0 \quad \text{as } t\to\infty\,,$$

By arguments used in the proof of Lemma B.4, it suffices to show that

$$\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{x_t - m}{\exp(\hat{f}_t(\boldsymbol{\theta}))} - \frac{x_t - m}{\exp(f_t(\boldsymbol{\theta}))}\right| \overset{e.a.s.}{\to} 0 \quad \text{as } t\to\infty\,,$$

and that $\{(x_t - m)/\exp(f_t(\boldsymbol{\theta}))\}_{t\in\mathbb{Z}}$ is SE and has a bounded $\log^+$-moment uniformly over Θ. The convergence result holds by an application of the mean value theorem:

$$\begin{aligned}\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{x_t - m}{\exp(\hat{f}_t(\boldsymbol{\theta}))} - \frac{x_t - m}{\exp(f_t(\boldsymbol{\theta}))}\right| &\le \sup_{\boldsymbol{\theta}\in\Theta, f\in\mathcal{F}}\left|\frac{x_t - m}{\exp(f)}\right| \sup_{\boldsymbol{\theta}\in\Theta}|\hat{f}_t(\boldsymbol{\theta}) - f_t(\boldsymbol{\theta})| \\ &\le \frac{|x_t| + \bar{m}}{\exp(\underline{f})}\sup_{\boldsymbol{\theta}\in\Theta}|\hat{f}_t(\boldsymbol{\theta}) - f_t(\boldsymbol{\theta})|\end{aligned}$$

as we know $f_t(\boldsymbol{\theta}) \in \mathcal{F} = [\underline{f},\infty)$ for some $\underline{f} > -\infty$, and where $\bar{m} = \sup_{\boldsymbol{\theta}\in\Theta} m < \infty$. It follows from Proposition 3 that $\{|x_t|\}_{t\in\mathbb{Z}}$ is SE and also:

$$\begin{aligned}\mathbb{E}(\log^+|x_t|) &= \mathbb{E}(\log^+|m_0 + \exp(f_t)\varepsilon_t|) \\ &\le 2\log 2 + \log^+|m_0| + \mathbb{E}(\log^+|\exp(f_t)\varepsilon_t|) \\ &\le 2\log 2 + \log^+|m_0| + \mathbb{E}(\log^+\exp(f_t)) + \mathbb{E}(\log^+|\varepsilon_t|) < \infty\end{aligned}$$

where the inequalities follow from Lemma 2.2 of Straumann and Mikosch (2006). The final expression is finite, because $\mathbb{E}(\log^+ \exp(f_t)) = \mathbb{E}(\mathbb{1}_{f_t>0} f_t) \leq \mathbb{E}|f_t| < \infty$, as f_t has bounded moments of any order by Proposition 3, and $\mathbb{E}(\log^+ |\varepsilon_t|) < \infty$ as ε_t has bounded moments of any order.

Finally, that $\{(x_t - m)/\exp(f_t(\boldsymbol{\theta}))\}_{t\in\mathbb{Z}}$ is SE follows from Krengel (1985, Proposition 4.3) and that the elements of this sequence have a bounded $\log^+$-moment uniformly over Θ holds by another application of Lemma 2.2 of Straumann and Mikosch (2006), and the fact that $f_t(\boldsymbol{\theta})$ is bounded from below by $\underline{f} > -\infty$. □

Proof of Lemma B.12. The log likelihood function $\log p_x(x|f;\boldsymbol{\gamma})$ is uniformly bounded from above over $x \in \mathbb{R}$, $f \in \mathcal{F}$ and $\boldsymbol{\theta} \in \Theta$, as can be seen from its expression in (D.5), as we have that Θ is compact, $f_t(\boldsymbol{\theta}) \geq \underline{f} > -\infty$ a.s., $\boldsymbol{\psi}$ is such that $H(z)$ diverges to $+\infty$ as $|z| \to \infty$, which together with $H(z;\boldsymbol{\psi})$ being finite on $\mathbb{R}$, implies that $H(z;\boldsymbol{\psi})$ is uniformly bounded from below. Also, $C(\boldsymbol{\psi})$ is finite for any $\boldsymbol{\psi} \in \boldsymbol{\Psi}$ under the imposed conditions. Therefore, $\log p_x(x|f;\boldsymbol{\gamma})$ is uniformly bounded from above, implying that $\mathbb{E}\ell_t(\boldsymbol{\theta}) < \infty$ exists for any $\boldsymbol{\theta}$, although it can be equal to $-\infty$.

The second result holds because

$$
\begin{aligned}
\mathbb{E}|\ell_t(\boldsymbol{\theta}_0)| &\leq \mathbb{E}\left|H\left(\frac{x_t - m_0}{\exp(f_t(\boldsymbol{\theta}_0))};\boldsymbol{\psi}_0\right)\right| + |\log C(\boldsymbol{\psi}_0)| + \mathbb{E}|f_t(\boldsymbol{\theta}_0)| \\
&= \mathbb{E}\left|H\left(\varepsilon_t;\boldsymbol{\psi}_0\right)\right| + |\log C(\boldsymbol{\psi}_0)| + \mathbb{E}|f_t(\boldsymbol{\theta}_0)| < \infty\,,
\end{aligned}
$$

where we use that $f_t(\boldsymbol{\theta}_0) = f_t$ a.s., due to $\{f_t(\boldsymbol{\theta}_0)\}_{t\in\mathbb{Z}}$ being a unique limit sequence. Each individual term in the expression on the right-hand side is finite, due to $f_t(\boldsymbol{\theta}_0)$ having bounded moments of any order by Proposition 4 and H being a piecewise quartic polynomial evaluated in ε_t, which has bounded moments of any order due to the exponential decay of the tails of its distribution. Furthermore, clearly $C(\boldsymbol{\psi}_0) < \infty$ under the assumptions that are in place. □

Proof of Lemma B.13. We know from Lemma B.12 that $\ell_t(\boldsymbol{\theta})$ is bounded from above, implying it is integrable, such that $\mathbb{E}\ell_t(\boldsymbol{\theta})$ exists for any $\boldsymbol{\theta} \in \Theta$. That lemma also tells us that $\mathbb{E}|\ell_t(\boldsymbol{\theta}_0)| < \infty$. Clearly, whenever $\boldsymbol{\theta} \in \Theta$ is such that $\mathbb{E}\ell_t(\boldsymbol{\theta}) = -\infty$, then $\mathbb{E}\ell_t(\boldsymbol{\theta}) < \mathbb{E}\ell_t(\boldsymbol{\theta}_0)$, so we just have to consider vectors $\boldsymbol{\theta} \in \Theta$ for which $\mathbb{E}\ell_t(\boldsymbol{\theta})$ is finite.

As for Lemma B.6, we prove the claim by showing that $\ell_t(\boldsymbol{\theta}) = \ell_t(\boldsymbol{\theta}_0)$ a.s. if and only if $\boldsymbol{\theta} = \boldsymbol{\theta}_0$.

We have argued in the proof of Lemma B.6 that if $\varepsilon_1 \sim p_\varepsilon(\varepsilon_1; \boldsymbol{\psi})$, then $a' + H((\varepsilon_1 - c')/b'; \boldsymbol{\psi}') = a'' + H((\varepsilon_1 - c'')/b''; \boldsymbol{\psi}'')$ a.s. if and only if $(a', b', c', \boldsymbol{\psi}') = (a'', b'', c'', \boldsymbol{\psi}'')$. Thus, because $\ell_t(\boldsymbol{\theta}) = \ell_t(\boldsymbol{\theta}_0)$ a.s. if and only if

$$-H\left(\frac{\exp(f_t(\boldsymbol{\theta}_0))\varepsilon_t + m_0 - m}{\exp(f_t(\boldsymbol{\theta}))}; \boldsymbol{\psi}\right) - \log C(\boldsymbol{\psi}) - f_t(\boldsymbol{\theta})$$
$$= -H(\varepsilon_t; \boldsymbol{\psi}_0) - \log C(\boldsymbol{\psi}_0) - f_t(\boldsymbol{\theta}_0)$$

a.s., it is clear that $\ell_t(\boldsymbol{\theta}) = \ell_t(\boldsymbol{\theta}_0)$ holds almost surely if and only if $\boldsymbol{\psi} = \boldsymbol{\psi}_0$, $m_0 = m$ and $f_t(\boldsymbol{\theta}) = f_t(\boldsymbol{\theta}_0)$ a.s. So it remains to be shown that given $\boldsymbol{\psi} = \boldsymbol{\psi}_0$ and $m = m_0$, $f_t(\boldsymbol{\theta}) = f_t(\boldsymbol{\theta}_0)$ a.s. holds if and only if $(\omega, \beta, \alpha) = (\omega_0, \beta_0, \alpha_0)$. Because $\{f_t(\boldsymbol{\theta})\}_{t\in\mathbb{Z}}$ is SE for any $\boldsymbol{\theta} \in \Theta$, then if we have $f_t(\boldsymbol{\theta}) = f_t(\boldsymbol{\theta}_0)$ almost surely, this must hold for any $t \in \mathbb{Z}$. Say that $\boldsymbol{\theta} \in \Theta$ is such that the equality holds almost surely, then the update equation of $f_t(\boldsymbol{\theta})$ implies that:

$$f_{t+1}(\boldsymbol{\theta}_0) - f_{t+1}(\boldsymbol{\theta})$$
$$= \omega_0(1-\beta_0) - \omega(1-\beta) + (\beta_0 - \beta) f_t(\boldsymbol{\theta}_0) + (\alpha_0 - \alpha)(h(\varepsilon_t; \boldsymbol{\psi}_0)\varepsilon_t - 1) \quad \text{a.s.}$$

To have that the left-hand side is equal to zero almost surely, we must have $\omega_0(1-\beta_0) = \omega(1-\beta)$, because otherwise, the equality implies that $(\beta_0-\beta)f_t(\boldsymbol{\theta}_0)+(\alpha_0-\alpha)(h(\varepsilon_t;\boldsymbol{\psi}_0)\varepsilon_t - 1) = \omega(1-\beta) - \omega_0(1-\beta_0) \neq 0$, which is ruled out, because the left-hand side cannot be equal to a nonzero constant due to $f_t(\boldsymbol{\theta}_0)$ being independent of $h(\varepsilon_t)\varepsilon_t$, and $f_t(\boldsymbol{\theta}_0)$ being non-degenerate due to $\alpha_0 \neq 0$ and $h(\varepsilon_t)\varepsilon_t$ clearly being non-degenerate under the maintained assumptions. Next, we also need $\beta_0 = \beta$ for the equality above to hold a.s., because in case $\beta_0 \neq \beta$, we would need $(\alpha_0 - \alpha)(h(\varepsilon_t; \boldsymbol{\psi}_0)\varepsilon_t - 1) = (\beta - \beta_0) f_t(\boldsymbol{\theta}_0)$, which is not possible due to $h(\varepsilon_t)\varepsilon_t$ and $(\beta - \beta_0)f_t(\boldsymbol{\theta}_0)$ being independent and non-degenerate. It then follows from $\omega_0(1-\beta_0) = \omega(1-\beta)$ and the assumption that $\beta_0 < 1$ that $\omega = \omega_0$. Finally, we must clearly also have $\alpha = \alpha_0$, as there is no other way of having $(\alpha_0 - \alpha)(h(\varepsilon_t; \boldsymbol{\psi}_0)\varepsilon_t - 1) = 0$ a.s.. This concludes the proof.

□

E. Derivatives location model

E.1. Derivatives of score function location model

In this section we provide the first, second and selected third order derivatives of $s(f, x_t; \boldsymbol{\gamma})$ with respect to $\boldsymbol{\theta}$, so effectively $\boldsymbol{\gamma} = (\sigma^2, \boldsymbol{\psi})$, and f, necessary for deriving the properties of the derivatives of $f_t(\boldsymbol{\theta})$. Recall that here $\boldsymbol{\theta} = (\omega, \beta, \alpha, \sigma^2, \boldsymbol{\psi})^\top$, where $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_q)$ for $q = k/2$, and where $\tau(\boldsymbol{\psi})$ is such that $y_1 = -\psi_k$, …, $y_q = -\psi_1$, $y_{q+1} = \psi_1$, …, $y_k = \psi_k$. We have that for any $f \in \mathbb{R}$, $x \in \mathbb{R}$ and $\boldsymbol{\gamma} \in \boldsymbol{\Gamma}$:

$$s(f, x; \boldsymbol{\gamma}) = \frac{1}{\sigma} h\left(\frac{x-f}{\sigma}; \boldsymbol{\psi}\right).$$

For the first order derivatives we will use that the nonzero elements of the derivative $\partial h(z; \boldsymbol{\psi})/\partial \boldsymbol{\theta}$ are given by:

$$\frac{\partial h\left(z; \boldsymbol{\psi}\right)}{\partial \psi_i} = h(z; \boldsymbol{e}_i)\,,$$

for any $z \in \mathbb{R}$ and $i = 1, \ldots, q$, and where $\boldsymbol{e}_j$ for $j = 1, \ldots, q$ denotes a q-dimensional vector with a 1 on entry j and zeroes elsewhere. This follows from the linearity of the function $h(z; \boldsymbol{\psi})$ in $\boldsymbol{\psi}$. It follows that for any $i = 1, \ldots, q$

$$\begin{aligned}
\frac{\partial s(f, x; \boldsymbol{\gamma})}{\partial f} &= -\frac{1}{\sigma^2} h'\left(\frac{x-f}{\sigma}; \boldsymbol{\psi}\right), \\
\frac{\partial s(f, x; \boldsymbol{\gamma})}{\partial \sigma^2} &= -\frac{1}{2\sigma^3}\left[h\left(\frac{x-f}{\sigma}; \boldsymbol{\psi}\right) + h'\left(\frac{x-f}{\sigma}; \boldsymbol{\psi}\right)\frac{x-f}{\sigma}\right], \\
\frac{\partial s(f, x; \boldsymbol{\gamma})}{\partial \psi_i} &= \frac{1}{\sigma} h\left(\frac{x-f}{\sigma}; \boldsymbol{e}_i\right),
\end{aligned}$$

where derivative $h'(z; \boldsymbol{\psi})$ denotes the derivative of the spline function h with respect to z, which is a second degree piecewise polynomial.

For second order derivatives of s, we will use that by the linearity of $h(z; \boldsymbol{\psi})$ in $\boldsymbol{y}$, the second order derivative of h with respect to the static parameters is a matrix of zeroes:

$$\frac{\partial^2 h\left(z; \boldsymbol{\psi}\right)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} = \mathbf{0}_{4+q \times 4+q}\,.$$

Furthermore, the nonzero elements of $\partial h'(z; \boldsymbol{\psi})/\partial \boldsymbol{\theta}$ are given by

$$\frac{\partial h'(z; \boldsymbol{\psi})}{\partial \psi_i} = h'(z; \boldsymbol{e}_i)\,,$$

for any $z \in \mathbb{R}$ and $i = 1, \ldots, q$. Thus, we have for any $i = 1, \ldots, q$ and $j = 1, \ldots, q$

$$\begin{aligned}
\frac{\partial^2 s(f, x; \boldsymbol{\gamma})}{\partial f \partial f} &= \frac{1}{\sigma^3} h''\left(\frac{x-f}{\sigma}; \boldsymbol{\psi}\right) \\
\frac{\partial^2 s(f, x; \boldsymbol{\gamma})}{\partial \sigma^2 \partial f} &= \frac{1}{\sigma^4} h'\left(\frac{x-f}{\sigma}; \boldsymbol{\psi}\right) - \frac{x-f}{\sigma^5} h''\left(\frac{x-f}{\sigma}; \boldsymbol{\psi}\right), \\
\frac{\partial^2 s(f, x; \boldsymbol{\gamma})}{\partial \psi_i \partial f} &= -\frac{1}{\sigma^2} h'\left(\frac{x-f}{\sigma}; \boldsymbol{e}_i\right), \\
\frac{\partial^2 s(f, x; \boldsymbol{\gamma})}{\partial \sigma^2 \partial \sigma^2} &= \frac{3}{4\sigma^5} h\left(\frac{x-f}{\sigma}; \boldsymbol{\psi}\right) + \frac{5(x-f)}{4\sigma^6} h'\left(\frac{x-f}{\sigma}; \boldsymbol{\psi}\right)
\end{aligned}$$

$$
\begin{aligned}
&\quad + \frac{(x-f)^2}{4\sigma^7} h''\left(\frac{x-f}{\sigma};\boldsymbol{\psi}\right)\,, \\
\frac{\partial^2 s(f,x;\boldsymbol{\gamma})}{\partial\sigma^2\partial\psi_i} &= -\frac{1}{2\sigma^3}\left[h\left(\frac{x-f}{\sigma};\boldsymbol{e}_i\right)\right. \\
&\qquad\qquad \left. + h'\left(\frac{x-f}{\sigma};\boldsymbol{e}_i\right)\frac{x-f}{\sigma}\right]\,, \\
\frac{\partial^2 s(f,x;\boldsymbol{\gamma})}{\partial\psi_i\partial\psi_j} &= 0\,.
\end{aligned}
$$

Notice that $h''(z;\boldsymbol{\psi})$ is a continuous piecewise linear function and can be straightforwardly computed given $h'(z;\boldsymbol{\psi})$.

Finally, we compute the following three selected third order derivatives for $i = 1,\dots,q$ and $j = 1,\dots,q$:

$$
\begin{aligned}
\frac{\partial^3 s(f,x;\boldsymbol{\gamma})}{\partial\psi_i\partial\psi_j\partial f} &= 0\,, \\
\frac{\partial^3 s(f,x;\boldsymbol{\gamma})}{\partial\sigma^2\partial\psi_i\partial f} &= \frac{1}{\sigma^4}h'\left(\frac{x-f}{\sigma};\boldsymbol{e}_i\right) - \frac{x-f}{\sigma^5}h''\left(\frac{x-f}{\sigma};\boldsymbol{e}_i\right)\,, \\
\frac{\partial^3 s(f,x;\boldsymbol{\gamma})}{\partial\psi_i\partial f\partial f} &= \frac{1}{\sigma^3}h''\left(\frac{x-f}{\sigma};\boldsymbol{e}_i\right)\,.
\end{aligned}
$$

E.2. Derivatives of log likelihood location model

In this section we give the first and second order derivatives of the log likelihood function of the location model defined in Section 4.1, with respect to $\boldsymbol{\theta}$ where $\boldsymbol{\theta} = (\omega,\beta,\alpha,\sigma^2,\boldsymbol{\psi})^\top$, where $\boldsymbol{\psi} = (\psi_1,\dots,\psi_q)$ for $q = k/2$, and where $\tau(\boldsymbol{\psi})$ is such that $y_1 = -\psi_k,\dots,y_q = -\psi_1$, $y_{q+1} = \psi_1,\dots,y_k = \psi_k$. For any $\boldsymbol{\theta}\in\Theta$, the log likelihood function reads:

$$
\frac{1}{T}\sum_{t=1}^{T}\ell_t(\boldsymbol{\theta})\,,\quad \text{where}\quad \ell_t(\boldsymbol{\theta}) = -H\left(\frac{x_t - f_t(\boldsymbol{\theta})}{\sigma};\boldsymbol{\psi}\right) - \log C(\boldsymbol{\psi}) - \frac{1}{2}\log\sigma^2\,. \tag{E.1}
$$

The first order derivative of $\ell_t(\boldsymbol{\theta})$ is given by:

$$
\begin{aligned}
\frac{\partial\ell_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} &= h\left(\frac{x_t - f_t(\boldsymbol{\theta})}{\sigma};\boldsymbol{\psi}\right)\left[\frac{1}{\sigma}\frac{\partial f_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} - (x_t - f_t(\boldsymbol{\theta}))\frac{\partial\sigma^{-1}}{\partial\boldsymbol{\theta}}\right] - \left.\frac{\partial H(z;\boldsymbol{\psi})}{\partial\boldsymbol{\theta}}\right|_{z=(x_t-f_t(\boldsymbol{\theta}))/\sigma} \\
&\quad - \frac{1}{C(\boldsymbol{\psi})}\frac{\partial C(\boldsymbol{\psi})}{\partial\boldsymbol{\theta}} - \frac{1}{2\sigma^2}\frac{\partial\sigma^2}{\partial\boldsymbol{\theta}}\,,
\end{aligned} \tag{E.2}
$$

where the nonzero elements of $\partial H(z;\boldsymbol{\psi})/\partial\boldsymbol{\theta}$, $\partial C(\boldsymbol{\psi})/\partial\boldsymbol{\theta}$, $\partial\sigma^{-1}/\partial\boldsymbol{\theta}$ and $\partial\sigma^2/\partial\sigma^2$ are given by

$$
\begin{aligned}
\frac{\partial H(z;\boldsymbol{\psi})}{\partial\psi_i} &= H(z;\boldsymbol{e}_i)\,, \\
\frac{\partial C(\boldsymbol{\psi})}{\partial\psi_i} &= \int_{-\infty}^{\infty}\frac{\partial\exp(-H(x;\boldsymbol{\psi}))}{\partial\boldsymbol{\pi}_i}\,\mathrm{d}x
\end{aligned}
$$

$$= \int_{-\infty}^{\infty} - \exp(-H(x; \boldsymbol{\psi}))H(x; \boldsymbol{e}_i)\, \mathrm{d}x\,,$$
$$\frac{\partial \sigma^{-1}}{\partial \sigma^2} = -\frac{1}{2\sigma^3}\,,$$
$$\frac{\partial \sigma^2}{\partial \sigma^2} = 1\,,$$

for $i = 1, \ldots, q$ and any $z \in \mathbb{R}$. The form of $\partial H(z; \boldsymbol{\psi})/\partial \psi_i$ follows from the linearity of the spline coefficients in $\boldsymbol{y}$. We will now argue why the derivative can be taken into the integral in the calculation of $\partial C(\boldsymbol{\psi})/\partial \psi_i$. Because Θ is compact, there must be some real positive value $t^* > 0$ such that for some $\boldsymbol{\psi}^* \in \boldsymbol{\Psi}$ and every $\boldsymbol{\psi} \in \boldsymbol{\Psi}$ and corresponding $\boldsymbol{y}$ and $\boldsymbol{y}^*$, respectively, $|\partial \exp(-H(x; \boldsymbol{\psi}^*))/\partial \psi_i| \geq \partial \exp(-H(x; \boldsymbol{\psi}))/\partial \psi_i$ for every $x > t^*$ or $x < -t^*$. Within $[-t^*, t^*]$, the derivative can be taken into the integral by Leibniz rule and outside of this interval, the derivative can be taken into the integral by an application of the dominated convergence theorem, as $|\partial \exp(-H(x; \boldsymbol{\psi}^*))/\partial \psi_i|$ is clearly integrable under the maintained assumptions.

For the second order derivative, we obtain

$$\begin{aligned}
\frac{\partial^2 \ell_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} = & -h'\left(\frac{x_t - f_t(\boldsymbol{\theta})}{\sigma}; \boldsymbol{\psi}\right) \\
& \qquad \left[\frac{1}{\sigma}\frac{\partial f_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} - (x_t - f_t(\boldsymbol{\theta}))\frac{\partial \sigma^{-1}}{\partial \boldsymbol{\theta}}\right]\left[\frac{1}{\sigma}\frac{\partial f_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^\top} - (x_t - f_t(\boldsymbol{\theta}))\frac{\partial \sigma^{-1}}{\partial \boldsymbol{\theta}^\top}\right] \\
& + h\left(\frac{x_t - f_t(\boldsymbol{\theta})}{\sigma}; \boldsymbol{\psi}\right)\left[\frac{1}{\sigma}\frac{\partial^2 f_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}^\top} + \frac{\partial f_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\frac{\partial \sigma^{-1}}{\partial \boldsymbol{\theta}^\top} + \frac{\partial \sigma^{-1}}{\partial \boldsymbol{\theta}}\frac{\partial f_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^\top}\right. \\
& \qquad \left. -(x_t - f_t(\boldsymbol{\theta}))\frac{\partial \sigma^{-1}}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}^\top}\right] \\
& + \left[\frac{1}{\sigma}\frac{\partial f_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} - (x_t - f_t(\boldsymbol{\theta}))\frac{\partial \sigma^{-1}}{\partial \boldsymbol{\theta}}\right]\left.\frac{\partial h(z; \boldsymbol{\psi})}{\partial \boldsymbol{\theta}^\top}\right|_{z=(x_t - f_t(\boldsymbol{\theta}))/\sigma} \\
& + \left.\frac{\partial h(z; \boldsymbol{\psi})}{\partial \boldsymbol{\theta}}\right|_{z=(x_t - f_t(\boldsymbol{\theta}))/\sigma}\left[\frac{1}{\sigma}\frac{\partial f_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^\top} - (x_t - f_t(\boldsymbol{\theta}))\frac{\partial \sigma^{-1}}{\partial \boldsymbol{\theta}^\top}\right] \\
& - \frac{1}{C(\boldsymbol{\psi})}\frac{\partial^2 C(\boldsymbol{\psi})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}^\top} + \frac{1}{C(\boldsymbol{\psi})^2}\left(\frac{\partial C(\boldsymbol{\psi})}{\partial \boldsymbol{\theta}}\right)^2 - \frac{1}{2}\frac{\partial \sigma^2}{\partial \boldsymbol{\theta}}\frac{\partial \sigma^{-2}}{\partial \boldsymbol{\theta}^\top}\,,
\end{aligned}$$

where we use that

$$\frac{\partial^2 H\left(z; \boldsymbol{\psi}\right)}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}^\top} = \mathbf{0}_{4+q\times 4+q}\,, \quad \text{and} \quad \frac{\partial^2 \sigma^2}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}^\top} = \mathbf{0}_{4+q\times 4+q}\,,$$

because $H(z; \boldsymbol{\psi})$ is linear in $\boldsymbol{y}$ and because $\partial \sigma^2/\partial \sigma^2 = 1$. The expressions of the nonzero elements of $\partial h(z; \boldsymbol{\psi})/\partial \boldsymbol{\theta}$ are given in the previous subsection. Finally, the nonzero elements of $\partial^2 \sigma^{-1}/\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}^\top$, $\partial \sigma^{-2}/\partial \sigma^2$ and $\partial^2 C(\boldsymbol{\psi})/\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}^\top$ are given by

$$\frac{\partial^2 \sigma^{-1}}{\partial \sigma^2 \partial \sigma^2} = \frac{3}{4\sigma^5}\,,$$

$$
\begin{aligned}
\frac{\partial \sigma^{-2}}{\partial \sigma^2} &= -\frac{1}{\sigma^4}\,, \\
\frac{\partial^2 C(\boldsymbol{\psi})}{\partial \psi_i \partial \psi_j} &= \int_{-\infty}^{\infty} \exp(-H(x;\boldsymbol{\psi}))H(x;\boldsymbol{e}_i)H(x;\boldsymbol{t},\boldsymbol{e}_j)\,\mathrm{d}x\,,
\end{aligned}
$$

for any $i, j = 1, \ldots, q$, where we use the same argumentation as for $\partial C(\boldsymbol{\psi})/\partial \boldsymbol{\psi}$ for taking the derivative into the integral.

F. Calculation of moments

Let us derive the moments of the spline probability distribution p_ε. For any non-negative integer n, we have that:

$$
\begin{aligned}
C\mathbb{E}[\varepsilon_t^n] &= \int_{-\infty}^{\infty} \varepsilon^n \ \exp(-H(\varepsilon;\boldsymbol{t},\boldsymbol{y}))\mathrm{d}\varepsilon \\
&= \int_{-\infty}^{t_1} \varepsilon^n \ \exp(-H(\varepsilon;\boldsymbol{t},\boldsymbol{y}))\mathrm{d}\varepsilon + \int_{t_1}^{t_k} \varepsilon^n \ \exp(-H(\varepsilon;\boldsymbol{t},\boldsymbol{y}))\mathrm{d}\varepsilon \\
&\quad + \int_{t_k}^{\infty} \varepsilon^n \ \exp(-H(\varepsilon;\boldsymbol{t},\boldsymbol{y}))\mathrm{d}\varepsilon\,,
\end{aligned}
\tag{F.1}
$$

where C denotes the integrating constant defined in (8). We will discuss how this constant can be calculated in Section F.1. The middle integral can be reliably evaluated numerically, because it is a smooth function that is integrated over a bounded integration range. In particular, we suggest using some quadrature routine between each pair of subsequent knots (t_i, t_{i+1}) for $i = 1, \ldots, k-1$, in order to obtain a good approximation. We use the `quad` routine of the `scipy.integrate` library in Python.

The first and last integrals can be simplified analytically. Note that beyond the outer knots, the natural cubic spline h is linear and therefore its anti-derivative is quadratic. For the tail integrals to be finite, the assumption on the coefficients in Assumption 1 must hold.

To evaluate the tail integrals, we have to distinguish between the case where $b_0 > 0$ ($b_k > 0$) and where $b_0 = 0$ and $a_0 < 0$ ($b_k = 0$ and $a_k > 0$). Let us start with the first integral on the right-hand side of (F.1) for the case $b_0 > 0$:

$$
\begin{aligned}
&\int_{-\infty}^{t_1} \varepsilon^n \exp(-H(\varepsilon;\boldsymbol{t},\boldsymbol{y}))\mathrm{d}\varepsilon = \int_{-\infty}^{t_1} \varepsilon^n \exp(-H_0(\varepsilon;\boldsymbol{t},\boldsymbol{y}))\mathrm{d}\varepsilon \\
&\qquad = \exp(-e_0) \int_{-\infty}^{t_1} \varepsilon^n \exp\left(-a_0(\varepsilon - t_1) - \frac{1}{2}b_0(\varepsilon - t_1)^2\right)\mathrm{d}\varepsilon \\
&\qquad = \exp(-e_0) \int_{-\infty}^{0} (y + t_1)^n \exp\left(-a_0 y - \frac{1}{2}b_0 y^2\right)\mathrm{d}y
\end{aligned}
$$

$$= \sqrt{2\pi b_0^{-1}} \exp\left(-e_0 + \frac{1}{2}a_0^2 b_0^{-1}\right) \int_{-\infty}^{0} (y+t_1)^n \underbrace{\frac{1}{\sqrt{2\pi b_0^{-1}}} \exp\left(-\frac{(y+a_0 b_0^{-1})^2}{2b_0^{-1}}\right)}_{\text{pdf of } \mathcal{N}(-a_0 b_0^{-1}, b_0^{-1})} \mathrm{d}y$$

$$= \tilde{\sigma}_0 \frac{\Phi(\tilde{\beta}_0)}{\varphi(\tilde{\beta}_0)} \exp(-e_0) \int_{-\infty}^{0} (y+t_1)^n \underbrace{\frac{1}{\Phi(\tilde{\beta}_0)\sqrt{2\pi}\tilde{\sigma}_0} \exp\left(-\frac{(y-\tilde{\mu}_0)^2}{2\tilde{\sigma}_0^2}\right)}_{\text{pdf of } tr\mathcal{N}(\tilde{\mu}_0, \tilde{\sigma}_0^2, -\infty, 0)} \mathrm{d}y$$

$$= \tilde{\sigma}_0 \frac{\Phi(\tilde{\beta}_0)}{\varphi(\tilde{\beta}_0)} \exp(-e_0) \mathbb{E}_{Z \sim tr\mathcal{N}(\tilde{\mu}_0, \tilde{\sigma}_0^2, -\infty, 0)} \left[(Z+t_1)^n\right] ,$$

where we use the notation $\tilde{\mu}_0 = -a_0 b_0^{-1}$, $\tilde{\sigma}_0^2 = b_0^{-1}$ and $\tilde{\beta}_0 = -\tilde{\mu}_0/\tilde{\sigma}_0 = a_0\sqrt{b_0^{-1}}$, where we use the change of variable $y = \varepsilon - t_1$ in the third equality and where $tr\mathcal{N}(\mu, \sigma^2, a, b)$ for $\mu \in \mathbb{R}$, $\sigma^2 > 0$ and $-\infty \leq a < b \leq \infty$, denotes the truncated normal distribution with mean μ, variance σ^2 and truncation on the interval (a, b). Furthermore, $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution and $\varphi(\cdot)$ denotes its probability density function. For the last integral in (F.1), under the restriction $b_k > 0$, a similar derivation gives:

$$\int_{t_k}^{\infty} \varepsilon^n \exp(-H(\varepsilon; \boldsymbol{t}, \boldsymbol{y})) \mathrm{d}\varepsilon = \tilde{\sigma}_k \frac{\Phi(-\tilde{\beta}_k)}{\varphi(\tilde{\beta}_k)} \exp(-e_k) \mathbb{E}_{Z \sim tr\mathcal{N}(\tilde{\mu}_k, \tilde{\sigma}_k^2, 0, \infty)} \left[(Z+t_k)^n\right] ,$$

where $\tilde{\mu}_k = -a_k b_k^{-1}$, $\tilde{\sigma}_k^2 = b_k^{-1}$ and $\tilde{\beta}_k = -\tilde{\mu}_k/\tilde{\sigma}_k = a_k\sqrt{b_k^{-1}}$.

For the case $b_0 = 0$ and $a_0 < 0$, the first integral on the right-hand side of (F.1) is equal to

$$\begin{aligned}
\int_{-\infty}^{t_1} \varepsilon^n \exp(-H(\varepsilon; \boldsymbol{t}, \boldsymbol{y})) \mathrm{d}\varepsilon &= \int_{-\infty}^{t_1} \varepsilon^n \exp(-H_0(\varepsilon; \boldsymbol{t}, \boldsymbol{y})) \mathrm{d}\varepsilon \\
&= \exp(-e_0) \int_{-\infty}^{t_1} \varepsilon^n \exp\left(-a_0(\varepsilon - t_1)\right) \mathrm{d}\varepsilon \\
&= \exp(-e_0) \int_{0}^{\infty} (t_1 - y)^n \exp\left(-(-a_0)y\right) \mathrm{d}y \\
&= -\frac{\exp(-e_0)}{a_0} \int_{0}^{\infty} (t_1 - y)^n \underbrace{(-a_0)\exp\left(-(-a_0)y\right)}_{\text{pdf of Exp}(\lambda) \text{ with } \lambda = -a_0} \mathrm{d}y \\
&= -\frac{\exp(-e_0)}{a_0} \mathbb{E}_{Z \sim \mathrm{Exp}(-a_0)} \left[(t_1 - Z)^n\right] ,
\end{aligned}$$

where we use the change of variable $y = t_1 - \varepsilon$ in the third equality and where $\mathrm{Exp}(\lambda)$ denotes the exponential distriubtion with parameter $\lambda > 0$. The last integral in (F.1), under the restriction $b_k = 0$ and $a_k > 0$, can be shown to be equal to $\frac{\exp(-e_k)}{a_k} \mathbb{E}_{Z \sim \mathrm{Exp}(a_k)} \left[(Z + t_k)^n\right]$ using a similar derivation.

In summary, we have found that for any integer n,

$$\int_{-\infty}^{t_1} \varepsilon^n \exp(-H(\varepsilon))\mathrm{d}\varepsilon = \begin{cases} \tilde{\sigma}_0 \frac{\Phi(\tilde{\beta}_0)}{\varphi(\tilde{\beta}_0)} \exp(-e_0)\mathbb{E}_{Z\sim tr\mathcal{N}(\tilde{\mu}_0,\tilde{\sigma}_0^2,-\infty,0)}\left[(Z+t_1)^n\right], & \text{if } b_0 > 0\,, \\ -\frac{\exp(-e_0)}{a_0}\mathbb{E}_{Z\sim\mathrm{Exp}(-a_0)}\left[(t_1-Z)^n\right], & \text{if } b_0 = 0 \text{ and } a_0 < 0\,, \end{cases}$$

$$\int_{t_k}^{\infty} \varepsilon^n \exp(-H(\varepsilon))\mathrm{d}\varepsilon = \begin{cases} \tilde{\sigma}_k \frac{\Phi(-\tilde{\beta}_k)}{\varphi(\tilde{\beta}_k)} \exp(-e_k)\mathbb{E}_{Z\sim tr\mathcal{N}(\tilde{\mu}_k,\tilde{\sigma}_k^2,0,\infty)}\left[(Z+t_k)^n\right], & \text{if } b_k > 0\,, \\ \frac{\exp(-e_k)}{a_k}\mathbb{E}_{Z\sim\mathrm{Exp}(a_k)}\left[(Z+t_k)^n\right], & \text{if } b_k = 0 \text{ and } a_k > 0\,, \end{cases} \tag{F.2}$$

where $\tilde{\mu}_0 = -a_0 b_0^{-1}$, $\tilde{\sigma}_0^2 = b_0^{-1}$ and $\tilde{\beta}_0 = -\tilde{\mu}_0/\tilde{\sigma}_0 = a_0\sqrt{b_0^{-1}}$ and where $\tilde{\mu}_k = -a_k b_k^{-1}$, $\tilde{\sigma}_k^2 = b_k^{-1}$ and $\tilde{\beta}_k = -\tilde{\mu}_k/\tilde{\sigma}_k = a_k\sqrt{b_k^{-1}}$.

For the first and second moment we can use that if $Z \sim tr\mathcal{N}(\tilde{\mu}_0, \tilde{\sigma}_0^2, -\infty, 0)$:

$$\mathbb{E}\left[Z\right] = \tilde{\mu}_0 - \frac{\varphi(\tilde{\beta}_0)}{\Phi(\tilde{\beta}_0)}\tilde{\sigma}_0\,, \quad \text{and} \quad \mathbb{V}\mathrm{ar}\left(Z\right) = \tilde{\sigma}_0^2\left[1 - \frac{\tilde{\beta}_0\varphi(\tilde{\beta}_0)}{\Phi(\tilde{\beta}_0)} - \left(\frac{\varphi(\tilde{\beta}_0)}{\Phi(\tilde{\beta}_0)}\right)^2\right]\,,$$

$$\text{such that} \quad \mathbb{E}\left[Z^2\right] = \tilde{\sigma}_0^2\left[1 + \tilde{\beta}_0^2 + \tilde{\beta}_0\frac{\varphi(\tilde{\beta}_0)}{\Phi(\tilde{\beta}_0)}\right]\,,$$

and if $Z \sim tr\mathcal{N}(\tilde{\mu}_k, \tilde{\sigma}_k^2, 0, \infty)$:

$$\mathbb{E}\left[Z\right] = \tilde{\mu}_k + \frac{\varphi(\tilde{\beta}_k)}{\Phi(-\tilde{\beta}_k)}\tilde{\sigma}_k\,, \quad \text{and} \quad \mathbb{V}\mathrm{ar}\left(Z\right) = \tilde{\sigma}_k^2\left[1 + \frac{\tilde{\beta}_k\varphi(\tilde{\beta}_k)}{\Phi(-\tilde{\beta}_k)} - \left(\frac{\varphi(\tilde{\beta}_k)}{\Phi(-\tilde{\beta}_k)}\right)^2\right]\,,$$

$$\text{such that} \quad \mathbb{E}\left[Z^2\right] = \tilde{\sigma}_k^2\left[1 + \tilde{\beta}_k^2 - \tilde{\beta}_k\frac{\varphi(\tilde{\beta}_k)}{\Phi(-\tilde{\beta}_k)}\right]\,.$$

F.1. Normalizing constant

We can compute the normalizing constant C using (F.1) evaluated in $n = 0$. The middle integral on the right-hand side can be approximated numerically, as discussed above. Using (F.2), it is clear that:

$$\int_{-\infty}^{t_1} \exp(-H(\varepsilon))\mathrm{d}\varepsilon = \begin{cases} \tilde{\sigma}_0 \frac{\Phi(\tilde{\beta}_0)}{\varphi(\tilde{\beta}_0)} \exp(-e_0)\,, & \text{if } b_0 > 0\,, \\ -\frac{\exp(-e_0)}{a_0}\,, & \text{if } b_0 = 0 \text{ and } a_0 < 0\,, \end{cases}$$

$$\text{and} \quad \int_{t_k}^{\infty} \exp(-H(\varepsilon))\mathrm{d}\varepsilon = \begin{cases} \tilde{\sigma}_k \frac{\Phi(-\tilde{\beta}_k)}{\varphi(\tilde{\beta}_k)} \exp(-e_k)\,, & \text{if } b_k > 0\,, \\ \frac{\exp(-e_k)}{a_k}\,, & \text{if } b_k = 0 \text{ and } a_k > 0\,. \end{cases}$$

F.2. Expectation

Let us derive the mean μ of ε_t for a given spline density:

$$\mu = \mathbb{E}[\varepsilon_t] = \frac{1}{C}\int_{-\infty}^{\infty} \varepsilon\ \exp(-H(\varepsilon; \boldsymbol{t}, \boldsymbol{y}))\mathrm{d}\varepsilon\,,$$

where C denotes the normalizing constant. Based on (F.2) for $n = 1$ and the form of the expectation of a truncated normal distribution, we have that:

$$\int_{-\infty}^{t_1} \varepsilon \exp(-H(\varepsilon))\mathrm{d}\varepsilon = \begin{cases} \exp(-e_0)\tilde{\sigma}_0 \left[\frac{\Phi(\tilde{\beta}_0)}{\varphi(\tilde{\beta}_0)}(t_1 + \tilde{\mu}_0) - \tilde{\sigma}_0\right], & \text{if } b_0 > 0 \\ -\frac{\exp(-e_0)}{a_0}(t_1 + a_0^{-1}), & \text{if } b_0 = 0 \text{ and } a_0 < 0, \end{cases}$$

$$\text{and} \quad \int_{t_k}^{\infty} \varepsilon \exp(-H(\varepsilon))\mathrm{d}\varepsilon = \begin{cases} \exp(-e_k)\tilde{\sigma}_k \left[\frac{\Phi(-\tilde{\beta}_k)}{\varphi(\tilde{\beta}_k)}(t_k + \tilde{\mu}_k) + \tilde{\sigma}_k\right], & \text{if } b_k > 0, \\ \frac{\exp(-e_k)}{a_k}(t_k + a_k^{-1}), & \text{if } b_k = 0 \text{ and } a_k > 0. \end{cases}$$

F.3. Variance

Now let us turn to the calculation of $\mathbb{E}[\varepsilon_t^2]$, such that we can calculate the variance of ε_t.

$$\mathbb{E}[\varepsilon_t^2] = \frac{1}{C}\int_{-\infty}^{\infty} \varepsilon^2 \ \exp(-H(\varepsilon; \boldsymbol{t}, \boldsymbol{y}))\mathrm{d}\varepsilon.$$

From (F.2) for $n = 2$ it follows that:

$$\int_{-\infty}^{t_1} \varepsilon^2 \exp(-H(\varepsilon))\mathrm{d}\varepsilon = \begin{cases} \exp(-e_0)\tilde{\sigma}_0 \left[\frac{\Phi(\tilde{\beta}_0)}{\varphi(\tilde{\beta}_0)}\left([t_1 + \tilde{\mu}_0]^2 + \tilde{\sigma}_0^2\right) - \tilde{\sigma}_0\left(\tilde{\mu}_0 + 2t_1\right)\right], \\ \hfill \text{if } b_0 > 0, \\ -\frac{\exp(-e_0)}{a_0}\left[t_1^2 + t_1\frac{2}{a_0} + \frac{2}{a_0^2}\right], \quad \text{if } b_0 = 0 \text{ and } a_0 < 0, \end{cases}$$

$$\int_{t_k}^{\infty} \varepsilon^2 \exp(-H(\varepsilon))\mathrm{d}\varepsilon = \begin{cases} \exp(-e_k)\tilde{\sigma}_k \left[\frac{\Phi(-\tilde{\beta}_k)}{\varphi(\tilde{\beta}_k)}\left([t_k + \tilde{\mu}_k]^2 + \tilde{\sigma}_k^2\right) + \tilde{\sigma}_k\left(\tilde{\mu}_k + 2t_k\right)\right], \\ \hfill \text{if } b_k > 0, \\ \frac{\exp(-e_k)}{a_k}\left[t_k^2 + t_k\frac{2}{a_k} + \frac{2}{a_k^2}\right], \quad \text{if } b_k = 0 \text{ and } a_k > 0, \end{cases}$$

Based on the values of $\mu = \mathbb{E}[\varepsilon_t]$ and $\mathbb{E}[\varepsilon_t^2]$, we can calculate $\sigma^2 = \mathbb{V}\mathrm{ar}(\varepsilon_t) = \mathbb{E}[\varepsilon_t^2] - \mu^2$.

References

Billingsley, P. (1999). *Convergence of Probability Measures*. John Wiley & Sons, 2nd edition.

Blasques, F., Gorgi, P., Koopman, S. J., and Wintenberger, O. (2018). Feasible invertibility conditions and maximum likelihood estimation for observation-driven models. *Electronic Journal of Statistics*, 12(1):1019–1052.

Blasques, F., van Brummelen, J., Koopman, S. J., and Lucas, A. (2022). Maximum likelihood estimation for score-driven models. *Journal of Econometrics*, 227(2):325–346.

Gorgi, P. and Koopman, S. J. (2023). Beta observation-driven models with exogenous regressors: A joint analysis of realized correlation and leverage effects. *Journal of Econometrics*, 237(2):105177.

Krengel, U. (1985). *Ergodic theorems*, volume 6 of *De Gruyter Studies in Mathematics*. de Gruyter, Berlin.

Rao, R. R. (1962). Relations between weak and uniform convergence of measures with applications. *The Annals of Mathematical Statistics*, 33(5):659–680.

Straumann, D. and Mikosch, T. (2006). Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: A stochastic recurrence equations approach. *The Annals of Statistics*, 34(5):2449–2495.