The Continuous-Time Limit of Score-Driven Volatility Models

Giuseppe Buccheri[*] Fulvio Corsi[†] Franco Flandoli[‡] Giulia Livieri[§]

ONLINE APPENDIX

[*]University of Rome Tor Vergata, Rome. E-mail: giuseppe.buccheri@uniroma2.it
[†]University of Pisa and City University of London. E-mail: fulvio.corsi@unipi.it
[‡]Scuola Normale Superiore, Pisa. E-mail: franco.flandoli@sns.it
[§]Scuola Normale Superiore, Pisa. E-mail: giulia.livieri@sns.it

A Weak convergence of stochastic systems

In this section, we gather a set of conditions for the weak convergence of a system of discrete-time stochastic difference equations to a system of stochastic differential equations (SDEs) (see the works of Stroock and Varadhan, 1979; Kushner, 1984; Ethier and Kurtz, 1986; Nelson, 1990).

Let $h > 0$ and $k \in \mathbb{N}$. Let $N \in \mathbb{N}$ be the dimension of a discrete-time Markov chain indexed by h, which will be denoted by $x_{kh}^{(h)}$. Let $\mathcal{F}_{kh}$ be the σ-algebra generated by the following set of variables $\left(kh, x_0^{(h)}, x_h^{(h)}, x_{2h}^{(h)}, \ldots, x_{kh}^{(h)}\right)$. Let ν_h be a probability measure on $\left(\mathbb{R}^N, \mathcal{B}\left(\mathbb{R}^N\right)\right)$, where $\mathcal{B}\left(\mathbb{R}^N\right)$ are the Borel sets on $\mathbb{R}^N$. Let $\Pi_{kh}^{(h)}\left(x, \cdot\right)$ be a transition function on $\mathbb{R}^N$, and $\mathbb{P}^{(h)}$ a fixed probability measure on $\mathcal{D}\left([0,\infty), \mathbb{R}^N\right)$. Now, we specify the distribution of the starting point $x_0^{(h)}$, the transition probabilities of the N-dimensional discrete-time Markov process $x_{kh}^{(h)}$, and we construct a continuous-time process $x_t^{(h)}$ from $x_{kh}^{(h)}$ through $\mathbb{P}^{(h)}$ in the following way, respectively:

$$\mathbb{P}^{(h)}\left[x_0^{(h)} \in \Gamma\right] = \nu_h\left(\Gamma\right) \quad \text{for any} \quad \Gamma \in \mathcal{B}\left(\mathbb{R}^N\right) \tag{A.1}$$

$$\mathbb{P}^{(h)}\left[x_t^{(h)} = x_{kh}^{(h)},\, kh < t < (k+1)h\right] = 1 \tag{A.2}$$

$$\mathbb{P}^{(h)}\left[x_{kh}^{(h)} \in \Gamma | \mathcal{F}_{kh}^{(h)}\right] = \Pi_{kh}^{(h)}\left(x_{(k-1)h}^{(h)}, \Gamma\right) \tag{A.3}$$

a.s. under $\mathbb{P}^{(h)}$, $\forall\, k \geq 0$ and $\Gamma \in \mathcal{B}\left(\mathbb{R}^N\right)$. We denote now with x_t the continuous-time process obtained from $x_t^{(h)}$ by shrinking h towards zero which represents, under suitable assumptions (see below), the diffusion limit process to which $x_t^{(h)}$ weakly converges as $h \downarrow 0$.

In what follows, we give assumptions for this convergence result (Nelson, 1990).

Assumption 1. *There exists a continuous mapping $a\left(x,t\right)$ from $\mathbb{R}^N \times [0,\infty)$ into the space of $\mathbb{N} \times \mathbb{N}$ non-negative definite symmetric matrices and a continuous, measurable mapping $b\left(x,t\right)$ from $\mathbb{R}^N \times [0,\infty)$ into $\mathbb{R}^N$ such that for all $R > 0$ and $(k-1)h < t < kh$ we have:*

(**A1.1**) $\lim_{h\to 0} \sup_{\|x\|\leq R} \left\| h^{-1}\mathbb{E}\left[x_{(k+1)h}^{(h)} - x_{kh}^{(h)} | x_{kh}^{(h)} = x\right] - b(x,t)\right\| = 0.$

(**A1.2**) $\lim_{h\to 0} \sup_{\|x\|\leq R} \left\| h^{-1}\mathbb{E}\left[\left(x_{(k+1)h}^{(h)} - x_{kh}^{(h)}\right)\left(x_{(k+1)h}^{(h)} - x_{kh}^{(h)}\right)^{\prime} \middle| x_{kh}^{(h)} = x\right] - a\left(x,t\right)\right\| = 0$

(**A1.3**) $\exists \delta > 0 : \lim_{h\to 0} \sup_{\|x\|\leq R} \left\| h^{-1}\mathbb{E}\left[\left|\left(x_{(k+1)h}^{(h)} - x_{kh}^{(h)}\right)_i\right|^{2+\delta} \middle| x_{kh}^{(h)} = x\right]\right\| = 0$

where $(\cdot)_i$ is the i^{th} element of the vector $\left(x_{(k+1)h}^{(h)} - x_{kh}^{(h)}\right)$.

Assumption 2. *There exists a continuous mapping $\sigma\left(x,t\right)$ from $\mathbb{R}^N \times [0,\infty)$ into the space of $\mathbb{N} \times \mathbb{N}$ matrices, such that for all $x \in \mathbb{R}^N$ and all $t \geq 0$, $a\left(x,t\right) = \sigma\left(x,t\right)\sigma\left(x,t\right)^{\prime}$.*

Assumption 3. *$x_0^{(h)}$ converges in distribution, as $h \to 0$, to a random variable x_0 with probability measure ν_0 on $\left(\mathbb{R}^N, \mathcal{B}\left(\mathbb{R}^N\right)\right)$.*

Assumption 4. *ν_0, $b\left(x,t\right)$, $a\left(x,t\right)$ uniquely specify the distribution of a diffusion process x_t with initial distribution ν_0, drift vector $b\left(x,t\right)$ and diffusion matrix $a(x,t)$.*

We remind the following Theorem (see Nelson, 1990, Theorem 2.1):

Theorem A.1. *Under Assumptions 1-4 the sequence of $x_t^{(h)}$ converges weakly (i.e., in distribution) as $h \to 0$ to the x_t process defined by the stochastic integral equation:*

$$x_t = x_0 + \int_0^t b\left(x_s, s\right)\, ds + \int_0^t \sigma\left(x_s, s\right) dW_{N,s} \tag{A.4}$$

where $W_{N,s}$ is a N-dimensional standard Brownian motion, independent of x_0, and where for any $\Gamma \in \mathcal{B}\left(\mathbb{R}^N\right)$, $\mathbb{P}\left[x_0 \in \Gamma\right] = \nu_0\left(\Gamma\right)$. Such an x_t process exists and is distributional unique. This distribution does not depend on the choice of $\sigma\left(\cdot,\cdot\right)$ made in Assumption 2. Finally, x_t remains finite in finite time intervals almost surely.

B Example in Remark 2

In Theorem 2.1, we show that the class of scale family models driven by the score of the conditional likelihood converges in distribution to a non-degenerate diffusion. As score-driven models are characterized by a single source of uncertainty, it may seem unnatural that they converge to a SDE characterized by two independent noises. Here, we describe an elementary example of weak convergence of random variables to a set of Brownian motions in which the number of independent noises is not invariant in the limit. Specifically, we show that two independent Brownian motions may arise from manipulation of a single sequence of independent random variables in a very natural way.

Example 1. *Let, for $i \in \mathbb{N}$, $Z_i \overset{d}{\sim} \mathcal{N}\left(0,1\right)$ be independent standard Gaussian random variables. It is well known (see, for instance Pierre, 1971) that sample mean $\overline{Z}_n := \frac{Z_1+...+Z_n}{n}$ and sample variance $S_n^2 := \frac{1}{n-1}\sum_{i=1}^n\left(Z_i - \overline{Z}_n\right)^2$ are independent. For every $N \in \mathbb{N}$, let us introduce the times $t_n^N = \frac{n}{N}$, $n = 1, 2, 3, ...$ and the following two discrete time processes:*

$$B_{t_n^N}^N := \frac{Z_1 + ... + Z_n}{\sqrt{N}}, \qquad W_{t_n^N}^N := \frac{\left(Z_1^2 - 1\right) + ... + \left(Z_n^2 - 1\right)}{\sqrt{2}\sqrt{N}}.$$

The previous two processes are centered and normalized in such a way that

$$\mathbb{E}\left[\left|B_{t_n^N}^N\right|^2\right] = \mathbb{E}\left[\left|W_{t_n^N}^N\right|^2\right] = t_n^N.$$

Now, let us call B_t^N and W_t^N the piecewise linear continuous processes, defined for all $t \geq 0$, that coincide with the previous ones at the discrete times t_n^N. In particular, Donsker invariance principle (Donsker, 1951) applies to both of them and gives that the process $\left(B_t^N\right)_{t\geq 0}$ converges in law to a Brownian motion $\left(B_t\right)_{t\geq 0}$ and $\left(W_t^N\right)_{t\geq 0}$ converges in law to a Brownian motion $\left(W_t\right)_{t\geq 0}$. In addition, using a simple algebraic manipulation, thanks to the independence of sample mean and sample variance one gets that $B_{t_n^N}^N$ and $W_{t_n^N}^N - \frac{n\overline{Z}_n^2}{\sqrt{2}\sqrt{N}}$ are independent. Besides, $\frac{n\overline{Z}_n^2}{\sqrt{2}\sqrt{N}}$ converges to zero in probability as $N \to \infty$ since $\mathbb{E}\left[\overline{Z}_n^2\right] = \frac{1}{N}$. Therefore, one can deduce that $B_{t_n^N}^N$ and $W_{t_n^N}^N$ are asymptotically independent and so that that $\left(B_t\right)_{t\geq 0}$ and $\left(W_t\right)_{t\geq 0}$ are independent.

The previous example shows that from the same Gaussian sequence $\left(Z_i\right)_{i\in\mathbb{N}}$ we can construct, by a scaling limit, a pair of independent Brownian motions, using as a main step the well known fact that the sample mean and sample variance are independent. It is important to note that the increments $\frac{\left(Z_i^2-1\right)}{\sqrt{2}\sqrt{N}}$ of the process W_t^N have similar form as those of GARCH models, which are particular instances of score-driven models. Hence the example is not just artificial, but corresponds to the examples of this work, although being more elementary. The limit theorems proved in the present work generalize the simple example above into two directions. First, they are related to non-normal random variables. A known theorem (see e.g. Lukacs, 1963) states that sample mean and sample variance are independent if and only if the underlying variables are normal; but *asymptotic* independence may hold for a much wider class of sequences of random variables (see, for instance Hudson and Tucker, 1979; Polymenis et al., 2017). Second, the update of volatility is determined by the score, which is generally different from the squared returns. Therefore, the results in this work suggest that it might be possible to construct examples similar to the one above involving non-normal random variables and sample scores. Nonetheless, the construction of such examples requires further investigation and is left for further research.

C Proof of results of Section 2

C.1 Proof of Theorem 2.2

The notation in this proof is concise to avoid clutter but is explicit in the main text. First, we compute the score ∇.

$$\begin{aligned}
\nabla &= \frac{\partial \log p\left(y|\Lambda\left(\lambda\right),\Theta\right)}{\partial\lambda} = \frac{\partial}{\partial\lambda}\left[\log\frac{1}{\sqrt{\Lambda(\lambda)h}}\Psi\left(\frac{y}{\sqrt{\Lambda(\lambda)h}},\Theta\right)\right] \\
&= -\frac{1}{2}\frac{\Lambda'(\lambda)}{\Lambda(\lambda)} - \frac{1}{2}\frac{\Lambda'(\lambda)}{\Lambda(\lambda)}\frac{\Psi'\left(\frac{y}{\sqrt{\Lambda(\lambda)h}},\Theta\right)}{\Psi\left(\frac{y}{\sqrt{\Lambda(\lambda)h}},\Theta\right)}\frac{y}{\sqrt{\Lambda(\lambda)h}} \\
&= -\frac{1}{2}\frac{\Lambda'(\lambda)}{\Lambda(\lambda)}\left[1+\frac{\Psi'\left(\frac{y}{\sqrt{\Lambda(\lambda)h}},\Theta\right)}{\Psi\left(\frac{y}{\sqrt{\Lambda(\lambda)h}},\Theta\right)}\frac{y}{\sqrt{\Lambda(\lambda)h}}\right] \qquad \text{(C.1)}
\end{aligned}$$

Therefore, the ℓ-th moment of the score is given by:

$$\begin{aligned}
\mathbb{E}\left[\nabla^{\ell}\right] &= \int_{-\infty}^{\infty}\nabla^{\ell}\,p\left(y|\Lambda(\lambda),\Theta\right)\,dy \\
&= (-1)^{\ell}\left[\frac{1}{2}\frac{\Lambda'(\lambda)}{\Lambda(\lambda)}\right]^{\ell}\int_{-\infty}^{\infty}\left[1+\frac{\Psi'\left(\frac{y}{\sqrt{\Lambda(\lambda)h}},\Theta\right)}{\Psi\left(\frac{y}{\sqrt{\Lambda(\lambda)h}},\Theta\right)}\frac{y}{\sqrt{\Lambda(\lambda)h}}\right]^{\ell}\frac{1}{\sqrt{\Lambda(\lambda)h}}\Psi\left(\frac{y}{\sqrt{\Lambda(\lambda)h}},\Theta\right)\,dy \\
&= (-1)^{\ell}\left[\frac{1}{2}\frac{\Lambda'(\lambda)}{\Lambda(\lambda)}\right]^{\ell}\int_{-\infty}^{\infty}\left(1+\frac{\Psi'(z,\Theta)}{\Psi(z,\Theta)}z\right)^{\ell}\Psi(z,\Theta)\,dz \qquad \text{(C.2)}
\end{aligned}$$

In particular, $\mathbb{E}\left[\nabla^{\ell}\right]$ is independent of h. For further reference, we show that $\mathbb{E}\left[\nabla\,\epsilon\right]=0$, where $\epsilon \overset{d}{\sim} h^{-1/2}\Psi(\cdot)$ and has zero mean. We have:

$$\begin{aligned}
\mathbb{E}\left[\nabla\,\epsilon\right] &= -\int_{-\infty}^{\infty}\frac{1}{2}\frac{\Lambda^{'}(\lambda)}{\Lambda\left(\lambda\right)}\left[1+\frac{\Psi^{'}\left(\frac{y}{\sqrt{h\Lambda(\lambda)}},\Theta\right)}{\Psi\left(\frac{y}{\sqrt{h\Lambda(\lambda)}},\Theta\right)}\frac{y}{\sqrt{h\Lambda\left(\lambda\right)}}\right]\frac{y\sqrt{h}}{\sqrt{h\Lambda(\lambda)}}\frac{1}{\sqrt{h\Lambda(\lambda)}}\Psi\left(\frac{y}{\sqrt{hf(\Lambda)}},\Theta\right)\,dy \\
&= -\frac{1}{2}\frac{\Lambda^{'}(\lambda)}{\Lambda(\lambda)}\int_{-\infty}^{\infty}\left[1+\frac{\Psi^{'}(z,\Theta)}{\Psi(z,\Theta)}z\right]z\sqrt{h}\Psi(z,\Theta)\,dz \\
&= -\frac{1}{2}\frac{\Lambda^{'}(\lambda)}{\Lambda(\lambda)}\sqrt{h}\int_{-\infty}^{\infty}\left[1+\frac{\Psi^{'}(z,\Theta)}{\Psi(z,\Theta)}z\right]z\Psi(z,\Theta)\,dz.
\end{aligned}$$

At this point, to conclude it is sufficient to note that:

$$\int_{-\infty}^{+\infty}\left[1+\frac{\Psi^{'}(z,\Theta)}{\Psi(z,\Theta)}z\right]z\Psi\left(z,\Theta\right)\,dz=\int_{-\infty}^{\infty}z\Psi(z,\Theta)dz+\int_{-\infty}^{+\infty}z^{2}\Psi^{'}\left(z,\Theta\right)dz=0$$

where in the last passage we integrated by parts and used the fact that z has zero-mean by assumption.

C.2 Computation of (A1.1), (A1.2) and (A1.3) for dynamic scale family models.

We show that, under the assumptions of Theorem 2.1, the conditional moments in (**A1.1**), (**A1.2**) and (**A1.3**) converge, as the interval between observations shrinks to zero, to well defined limits. The first step is to compute the increments of the process $x_{kh}^{(h)}$ and $\lambda_{kh}^{(h)}$, that is:

$$\begin{aligned} x_{kh}^{(h)}-x_{(k-1)h}^{(h)} &= \sqrt{\Lambda(\lambda_{kh}^{(h)}}\epsilon_{kh}^{(h)}, \\ \lambda_{(k+1)h}^{(h)}-\lambda_{kh}^{(h)} &= \omega_h-(1-\beta_h)\lambda_{kh}^{(h)}+\alpha_h s(\lambda_{kh}^{(h)})\nabla_{kh}^{(h)} \end{aligned}$$

*Condition (**A1.1**)*

$$h^{-1}\mathbb{E}\left[\left(x_{kh}^{(h)}-x_{(k-1)h}^{(h)}\right)|\mathcal{F}_{kh}^{(h)}\right]=h^{-1}\mathbb{E}\left[\sqrt{\Lambda(\lambda_{kh}^{(h)})}\epsilon_{kh}^{(h)}|\mathcal{F}_{kh}^{(h)}\right]=h^{-1}\sqrt{\Lambda\left(\lambda_{kh}\right)}\mathbb{E}\left[\epsilon_{kh}^{(h)}|\mathcal{F}_{kh}^{(h)}\right]=0 \tag{C.3}$$

$$\begin{aligned} &h^{-1}\mathbb{E}\left[\left(\lambda_{(k+1)h}^{(h)}-\lambda_{kh}^{(h)}\right)|\mathcal{F}_{kh}^{(h)}\right]=h^{-1}\omega_h+h^{-1}\left(\beta_h-1\right)\lambda_{kh}^{(h)}+\alpha_h s\left(\lambda_{kh}^{(h)}\right)\mathbb{E}\left[\nabla_{kh}^{(h)}|\mathcal{F}_{kh}^{(h)}\right] \\ &=h^{-1}\omega_h-h^{-1}\left(1-\beta_h\right)\lambda_{kh}^{(h)} \end{aligned} \tag{C.4}$$

In the last step we use the well known fact that $\mathbb{E}\left[\nabla_{kh}^{(h)}|\mathcal{F}_{kh}^{(h)}\right]=0$. In particular, if assumptions of Theorem 2.1 are in force the above quantities are finite when $h\downarrow 0$.

*Condition (**A1.2**)*

$$h^{-1}\mathbb{E}\left[\left(x_{kh}^{(h)}-x_{(k-1)h}^{(h)}\right)^2|\mathcal{F}_{kh}^{(h)}\right]=h^{-1}\Lambda(\lambda_{kh}^{(h)})\mathbb{E}\left[(\epsilon_{kh}^{(h)})^2|\mathcal{F}_{kh}^{(h)}\right]=\Lambda(\lambda_{kh}^{(h)})\int_{\mathbb{R}}z^2\Psi\left(z\right)\,dz=\Lambda(\lambda_{kh}^{(h)})\zeta^{(2)} \tag{C.5}$$

$$\begin{aligned} &h^{-1}\mathbb{E}\left[\left(\lambda_{(k+1)h}^{(h)}-\lambda_{kh}^{(h)}\right)^2|\mathcal{F}_{kh}^{(h)}\right]=h^{-1}\omega_h^2+h^{-1}\left(\beta_h-1\right)^2\left(\lambda_{kh}^{(h)}\right)^2+h^{-1}\alpha_h^2 s\left(\lambda_{kh}^{(h)}\right)^2\xi_{kh}^{(2)} \\ &-\quad 2h^{-1}\omega_h\left(1-\beta_h\right)\lambda_{kh}^{(h)} \end{aligned} \tag{C.6}$$

$$\begin{aligned} &h^{-1}\mathbb{E}\left[\left(x_{kh}^{(h)}-x_{(k-1)h}^{(h)}\right)\left(\lambda_{(k+1)h}^{(h)}-\lambda_{kh}^{(h)}\right)|\mathcal{F}_{kh}^{(h)}\right]=h^{-1}\sqrt{\Lambda(\lambda_{kh}^{(h)})}\omega_h\mathbb{E}\left[\epsilon_{kh}^{(h)}|\mathcal{F}_{kh}^{(h)}\right] \\ &-\quad h^{-1}\sqrt{\Lambda\left(\lambda_{kh}^{(h)}\right)}\left(1-\beta_h\right)\lambda_{kh}^{(h)}\mathbb{E}\left[\epsilon_{kh}^{(h)}|\mathcal{F}_{kh}^{(h)}\right]+h^{-1}\sqrt{\Lambda\left(\lambda_{kh}^{(h)}\right)}\alpha_h s\left(\lambda_{kh}^{(h)}\right)\mathbb{E}\left[\epsilon_{kh}^{(h)}\nabla_{kh}^{(h)}|\mathcal{F}_{kh}^{(h)}\right] \\ &=\quad h^{-1}\sqrt{\Lambda\left(\lambda_{kh}^{(h)}\right)}\alpha_h s\left(\lambda_{kh}^{(h)}\right)\mathbb{E}\left[\epsilon_{kh}^{(h)}\nabla_{kh}^{(h)}|\mathcal{F}_{kh}^{(h)}\right]=0, \end{aligned} \tag{C.7}$$

where in the last step we use the fact that $\mathbb{E}\left[\epsilon_{kh}^{(h)}\nabla_{kh}^{(h)}|\mathcal{F}_{kh}^{(h)}\right]=0$. At this point we use Theorem 2.2 and assumptions of Theorem (2.1) to conclude that:

$$h^{-1}\mathbb{E}\left[\left(\lambda_{(k+1)h}^{(h)}-\lambda_{kh}^{(h)}\right)^{2}|\mathcal{F}_{kh}^{(h)}\right]=\alpha^{2}s\left(\lambda_{kh}^{(h)}\right)^{2}\xi_{kh}^{(2)}+o\left(1\right).$$

*Condition (**A1.3**)*

By straightforward computations as in Nelson (1990), and by setting $\omega_h=h\omega$, $\alpha_h=h^{1/2}\alpha$ and $\beta_h=1-h\theta$, condition (**A1.3**) holds for $\delta=2$. Indeed:

$$h^{-1}\mathbb{E}\left[\left(x_{kh}^{(h)}-x_{(k-1)h}^{(h)}\right)^{4}|\mathcal{F}_{kh}^{(h)}\right]=h^{-1}\Lambda^{2}\left(\lambda_{kh}^{(h)}\right)\mathbb{E}\left[\left(\epsilon_{kh}^{(h)}\right)^{4}|\mathcal{F}_{kh}^{(h)}\right]=h\Lambda^{2}\left(\lambda_{kh}^{(h)}\right)\zeta^{(4)}. \tag{C.8}$$

$$h^{-1}\mathbb{E}\left[\left(\lambda_{kh}^{(h)}-\lambda_{(k-1)h}^{(h)}\right)^{4}|\mathcal{F}_{kh}^{(h)}\right]=h\,\alpha^{4}s\left(\lambda_{kh}^{(h)}\right)^{4}\mathbb{E}\left[\left(\nabla_{kh}^{(h)}\right)^{4}|\mathcal{F}_{kh}^{(h)}\right]+O\left(h^{\gamma}\right),\qquad\gamma\geq 3/2 \tag{C.9}$$

Both quantities on the right hand of the previous equations converge to zero as $h\to 0$, where we use, again, Theorem 2.2 and the fact that $\zeta^{(4)}$ is finite by assumption.

C.3 Proof of Theorems 2.3 and 2.4

We report here only the proof of Theorem 2.3. The proof of Theorem 2.4 is omitted since it follows immediately from the former.

We prove existence and uniqueness of the limiting SDE in Theorem 2.3, which ensures that the initial distribution, the drift and the diffusion coefficients uniquely specify the distribution of a diffusion process characterized by these data; see Assumption 4. The proof is based on two steps. In the first, one assumes that the pair $\left(x_t,\sigma_t^2\right)$ is a solution of the system of SDEs and proves that this solution is actually given by the explicit formula. In the second, one assumes that the process given by the explicit formula is a solution of the system of SDEs. In particular, the first step proves uniqueness and the second existence. We here prove explicitly only uniqueness, since the existence is proved in a similar manner.

For shortness of notation we set $y_t=\sigma_t^2$ and $C=C\left(\Theta\right)$. The SDE for σ_t^2 reads as:

$$dy_t=\left(\omega-\theta y_t\right)dt+Cy_t dW_t^{(2)},$$

which is a closed equation in the single unknown y_t, with globally Lipschitz continuous coefficients. Therefore, it has a unique solution. In particular, the process:

$$\widetilde{y}_t=\exp\left(-\left(\theta+\frac{1}{2}C^2\right)t+CW_t^{(2)}\right)$$

is the unique solution of the following homogeneous equation:

$$d\widetilde{y}_t=-\theta\widetilde{y}_t dt+C\widetilde{y}_t dW_t^{(2)}\quad \widetilde{y}_0=1.$$

Indeed, by Itô formula one obtains:

$$d\widetilde{y}_t\;=\;\widetilde{y}_t\left(-\left(\theta+\frac{1}{2}C^2\right)dt+CdW_t^{(2)}\right)+\frac{1}{2}\widetilde{y}_tC^2dt=-\theta\widetilde{y}_tdt+CdW_t^{(2)}.$$

Then the process:

$$y_t = \widetilde{y}_t \left(\sigma_0^2 + \omega \int_0^t \frac{1}{\widetilde{y}_s} ds \right) = e^{-\left(\theta + \frac{1}{2}C^2\right)t + CW_t^{(2)}} \left(\sigma_0^2 + \omega \int_0^t e^{\left(\theta + \frac{1}{2}C^2\right)s - CW_t^{(2)}} ds \right)$$

is the unique solution of the previous SDE with initial condition σ_0^2. Indeed, by Itô formula one has:

$$\begin{aligned} dy_t &= \left(\sigma_0^2 + \omega \int_0^t \frac{1}{\widetilde{y}_s} ds \right) d\widetilde{y}_t + \widetilde{y}_t \omega \frac{1}{\widetilde{y}_t} dt \\ &= \left(\sigma_0^2 + \omega \int_0^t \frac{1}{\widetilde{y}_s} ds \right) \left(-\theta \widetilde{y}_t dt + C \widetilde{y}_t dW_t^{(2)} \right) + \omega dt \\ &= -\theta y_t dt + C y_t dW_t^{(2)} + \omega dt. \end{aligned}$$

Therefore, we prove that σ_t^2 is given by the correct explicit formula. Let x_t be a process satisfying $dx_t = \sigma_t dW_t^{(1)}$ with given x_0. Then:

$$x_t = x_0 + \int_0^t \sigma_s W_s^{(2)}.$$

However, we know that

$$\sigma_s = e^{-\frac{1}{2}\left(\theta + \frac{1}{2}C^2\right)s + \frac{1}{2}CW_s^{(2)}} \sqrt{\left(\sigma_0^2 + \omega \int_0^s e^{\left(\theta + \frac{1}{2}C^2\right)r - CW_r^{(2)}} dr \right)},$$

whence the thesis.

D Proofs of results of Section 3

D.1 Proof of Theorem 3.1

The proof follows the steps in Section C.2. Thus, we start by the computation of the increments of the processes $x_{kh}^{(h)}$ and $\sigma_{kh}^{(h),2}$:

$$\begin{aligned} &x_{kh}^{(h)} - x_{(k-1)h}^{(h)} = \sigma_{kh}^{(h)} \epsilon_{kh}^{(h)} \\ &\sigma_{(k+1)h}^{(h),2} - \sigma_{kh}^{(h),2} = \omega_h - (1 - \beta_h) \sigma_{kh}^{(h),2} + h^{-1} \alpha_h \sigma_{kh}^{(h),2} \epsilon_{kh}^{(h),2} \end{aligned}$$

*Condition (**A1.1**)*

$$h^{-1} \mathbb{E} \left[\left(x_{kh}^{(h)} - x_{(k-1)h}^{(h)} \right) | \mathcal{F}_{kh}^{(h)} \right] = h^{-1} \sigma_{kh}^{(h)} \mathbb{E} \left[\epsilon_{kh}^{(h)} | \mathcal{F}_{kh}^{(h)} \right] = 0 \quad \text{(D.1)}$$

$$\begin{aligned} & h^{-1} \mathbb{E} \left[\left(\sigma_{(k+1)h}^{(h),2} - \sigma_{kh}^{(h),2} \right) | \mathcal{F}_{kh}^{(h)} \right] = h^{-1} \omega_h - h^{-1} (1 - \beta_h) \sigma_{kh}^{(h),2} + h^{-2} \alpha_h \sigma_{kh}^{(h),2} \mathbb{E} \left[\epsilon_{kh}^{(h),2} | \mathcal{F}_{kh}^{(h)} \right] \\ = \; & h^{-1} \omega_h - h^{-1} (1 - \beta_h - \alpha_h) \sigma_{kh}^{(h),2}. \quad \text{(D.2)} \end{aligned}$$

In particular, if the assumptions of Theorem 3.1 are satisfied the above quantities are finite when $h \downarrow 0$.

*Condition (**A1.2**)*

$$h^{-1}\mathbb{E}\left[\left(x_{kh}^{(h)} - x_{(k-1)h}^{(h)}\right)^2 |\mathcal{F}_{kh}^{(h)}\right] = h^{-1}\sigma_{kh}^{(h),2}\mathbb{E}\left[\epsilon_{kh}^{(h),2}|\mathcal{F}_{kh}^{(h)}\right] = \sigma_{kh}^{(h),2} \tag{D.3}$$

$$\begin{aligned} h^{-1}\mathbb{E}\left[\left(\sigma_{(k+1)h}^{(h),2} - \sigma_{kh}^{(h),2}\right)^2 |\mathcal{F}_{kh}^{(h)}\right] &= h^{-1}\omega_h^2 + h^{-1}\left(1-\beta_h\right)^2 \sigma_{kh}^{(h),4} + h^{-1}\alpha_h^2\sigma_{kh}^{(h),4}\mathbb{E}\left[\epsilon_{kh}^{(h),4}|\mathcal{F}_{kh}^{(h)}\right] \\ - \quad 2h^{-1}\omega_h\left(1-\beta_h\right)\sigma_{kh}^{(h),2} - 2h^{-2}(1-\beta_h)\alpha_h\sigma_{kh}^{(h),2}&\mathbb{E}\left[\epsilon_{kh}^{(h),2}|\mathcal{F}_{kh}^{(h)}\right] + 2h^{-2}\omega_h\,\alpha_h\,\sigma_{kh}^{(h),2}\mathbb{E}\left[\epsilon_{kh}^{(h),2}|\mathcal{F}_{kh}^{(h)}\right] \end{aligned} \tag{D.4}$$

We use now the fact that $\mathbb{E}\left[\epsilon_{kh}^{(h),4}|\mathcal{F}_{kh}^{(h)}\right] = 3h^2\frac{(\nu-2)}{(\nu-4)}$ and $\mathbb{E}\left[\epsilon_{kh}^{(h),2}|\mathcal{F}_{kh}^{(h)}\right] = h$. By rearranging the terms we obtain:

$$\begin{aligned} &h^{-1}\mathbb{E}\left[\left(\sigma_{(k+1)h}^{(h),2} - \sigma_{kh}^{(h),2}\right)^2 |\mathcal{F}_{kh}^{(h)}\right] \\ &= h^{-1}\omega_h^2 + h^{-1}\sigma_{kh}^{(h),4}\left(\alpha_h+\beta_h-1\right)^2 + 2\omega_h\sigma_{kh}^{(h),2}\left(\alpha_h+\beta_h-1\right) + \sigma_{kh}^{(h),4}\alpha_h^2 2\frac{(\nu-1)}{(\nu-4)} \\ &= \sigma_{kh}^{(h),4}\alpha_h^2 2\frac{(\nu-1)}{(\nu-4)} + o(1) \end{aligned}$$

where in the last step we use the assumptions of Theorem 3.1 on the parameters. Now, by using the fact that $\mathbb{E}\left[\epsilon_{kh}^{(h)}|\mathcal{F}_{kh}^{(h)}\right] = 0$ and $\mathbb{E}\left[\epsilon_{kh}^{(h),3}|\mathcal{F}_{kh}^{(h)}\right] = 0$ we obtain:

$$h^{-1}\mathbb{E}\left[\left(x_{kh}^{(h)} - x_{(k-1)h}^{(h)}\right)\left(\sigma_{(k+1)h}^{(h),2} - \sigma_{kh}^{(h),2}\right)\right] = o(1). \tag{D.5}$$

It is immediate to check that the the fourth moments go to zero. The weak existence, the uniqueness and the finiteness of the solution of the SDE on compact sets are guaranteed by Theorem 2.3.

D.2 Proof of Theorem 3.2

The proof follows the lines of the proof of Theorem (3.1). The only difference is that $\epsilon_{kh}^{(h)} \overset{d}{\sim} \text{GED}\left(\nu, h\right)$, where $\text{GED}\left(\nu, h\right)$ denotes the generalized error distribution with shape parameter ν and scale parameter h. In particular, $\mathbb{E}\left[\epsilon_{kh}^2|\mathcal{F}_{kh}\right] = 2^{\frac{2}{\nu}}h\frac{\Gamma\left(\frac{3}{\nu}\right)}{\Gamma\left(\frac{1}{\nu}\right)}$ and $\mathbb{E}\left[\epsilon_{kh}^4\right] = 2^{\frac{4}{\nu}}h^2\frac{\Gamma\left(\frac{5}{\nu}\right)}{\Gamma\left(\frac{1}{\nu}\right)}$. We report details for the following quantity only:

$$\begin{aligned} h^{-1}\mathbb{E}\left[\left(\sigma_{(k+1)h}^{(h),2} - \sigma_{kh}^{(h),2}\right)^2 |\mathcal{F}_{kh}^{(h)}\right] &= h^{-1}\omega_h^2 - 2h^{-1}\omega_h\sigma_{kh}^{(h),2}\left(1-\beta_h-\alpha_h\right) \\ +h^{-1}\sigma_{kh}^{(h),4}\left(\beta_h+\alpha_h-1\right)^2 + 2h^{-1}\omega_h\sigma_{kh}^{(h),2}\alpha_h&\left(2^{\frac{2}{\nu}}\frac{\Gamma\left(\frac{3}{2}\right)}{\Gamma\left(\frac{1}{2}\right)} - 1\right) + h^{-1}\sigma_{kh}^{(h),4}\left(2^{\frac{4}{\nu}}h^2\frac{\Gamma\left(\frac{5}{\nu}\right)}{\Gamma\left(\frac{1}{\nu}\right)} - 1\right)\alpha_h^2 \\ +2h^{-1}\sigma_{kh}^{(h),4}\beta_h\alpha_h\left(\frac{\Gamma\left(\frac{3}{\nu}\right)}{\Gamma\left(\frac{1}{\nu}\right)}2^{\frac{2}{\nu}} - 1\right) &+ h^{-1}\sigma_{kh}^{(h),4}\left(1 - \frac{\Gamma\left(\frac{3}{\nu}\right)}{\Gamma\left(\frac{1}{\nu}\right)}2^{\frac{2}{\nu}}\right) \end{aligned} \tag{D.6}$$

Under the assumptions of Theorem 3.2, we have:

$$h^{-1}\mathbb{E}\left[\left(\sigma_{(k+1)h}^{(h),2} - \sigma_{kh}^{(h),2}\right)^2 |\mathcal{F}_{kh}^{(h)}\right] = \alpha^2\frac{\Gamma\left(\frac{1}{\nu}\right) + 4^{\frac{1}{\nu}}\left(4^{\frac{1}{\nu}}\Gamma\left(\frac{5}{\nu}\right) - 2\,\Gamma\left(\frac{3}{\nu}\right)\right)}{\Gamma\left(\frac{1}{\nu}\right)} + o(1) \tag{D.7}$$

Weak existence, uniqueness and the finiteness of the solution of the SDE on compact sets are guaranteed by Theorem 2.3.

D.3 Proof of Theorem 3.3

The proof follows the lines of the proof of Theorems (3.1) and (3.2). As in Appendix (D.2), $\epsilon_{kh}^{(h)} \overset{d}{\sim} \text{GED}(\nu, h)$. We only report the computation of the second moment per unit of time of $\lambda_{(k+1)h}^{(h)} - \lambda_{(k)h}^{(h)}$. Weak existence, uniqueness and the finiteness of the solution of the SDE on compact sets are guaranteed by Theorem 2.4. We have:

$$
\begin{aligned}
&h^{-1}\mathbb{E}\left[\left(\lambda_{(k+1)h}^{(h)} - \lambda_{kh}^{(h)}\right)^2 |\mathcal{F}_{kh}^{(h)}\right] = h^{-1}\omega_h^2 + h^{-1}(\beta_h - 1)^2\lambda_{kh}^{(h),2} + \alpha_*^2 \frac{4^{\frac{1}{\nu}}\Gamma(\frac{3}{\nu})}{\Gamma(\frac{1}{\nu})} \\
&+\alpha^2 \frac{4^{\frac{1}{\nu}}\Gamma(\frac{3}{\nu})}{\Gamma(\frac{1}{\nu})} + \alpha^2 \frac{2^{\frac{6}{\nu}}\Gamma(\frac{1}{2}+\frac{1}{\nu})^2}{4\pi} + 2h^{-1}\omega_h(\beta_h - 1)\lambda_{kh}^{(h)} \\
&-2\alpha^2 \frac{2^{\frac{6}{\nu}}\Gamma(\frac{1}{2}+\frac{1}{\nu})^2}{4\pi}
\end{aligned} \tag{D.8}
$$

Under the assumptions of Theorem 3.3, we have:

$$
h^{-1}\mathbb{E}\left[\left(\lambda_{(k+1)h}^{(h)} - \lambda_{kh}^{(h)}\right)^2 |\mathcal{F}_{kh}^{(h)}\right] = \alpha_*^2 \frac{4^{\frac{1}{\nu}}\Gamma(\frac{3}{\nu})}{\Gamma(\frac{1}{\nu})} + \alpha^2 \left(\frac{4^{\frac{1}{\nu}}\Gamma(\frac{3}{\nu})}{\Gamma(\frac{1}{\nu})} - \frac{2^{\frac{6}{\nu}}\Gamma(\frac{1}{2}+\frac{1}{\nu})^2}{4\pi} \right) + o(1) \tag{D.9}
$$

E Proofs of the results of Section 5.1

E.1 Proof of Theorem 5.2

Again, the notation in this proof is concise to avoid clutter. First, we compute ∇_μ:

$$
\nabla_\mu = \frac{\partial \log p\left(y|\mu, \Lambda\left(\lambda\right), \Theta\right)}{\partial \mu} = \frac{\partial}{\partial \mu}\left[\log \frac{1}{\sqrt{\Lambda\left(\lambda\right)h}} \Psi\left(\frac{y}{\sqrt{\Lambda\left(\lambda\right)h}}, \Theta\right)\right] = -\sqrt{\frac{h}{\Lambda\left(\lambda\right)}} \frac{\Psi'\left(\frac{y-h\mu}{\sqrt{\Lambda(\lambda)h}}, \Theta\right)}{\Psi\left(\frac{y-h\mu}{\sqrt{\Lambda(\lambda)h}}, \Theta\right)}.
$$

Therefore, the ℓ-*th* moment of the score ∇^μ is given by:

$$
\begin{aligned}
\mathbb{E}\left[\nabla_\mu^\ell\right] &= \int_{-\infty}^{\infty} \nabla_\mu^\ell\, p\left(y|\mu, \Lambda(\lambda), \Theta\right) dy \\
&= \left(-1\right)^\ell \int_{-\infty}^{\infty} \left[\sqrt{\frac{h}{\Lambda\left(\lambda\right)}} \frac{\Psi'\left(\frac{y-\mu h}{\sqrt{\Lambda(\lambda)h}}, \Theta\right)}{\Psi\left(\frac{y-\mu h}{\sqrt{\Lambda(\lambda)h}}, \Theta\right)}\right]^\ell \frac{1}{\sqrt{\Lambda\left(\lambda\right)h}} \Psi\left(\frac{y-\mu h}{\sqrt{\Lambda\left(\lambda\right)h}}, \Theta\right) dy \\
&= \left(-1\right)^\ell \left(\frac{h}{\Lambda\left(\lambda\right)}\right)^{\ell/2} \int_{\mathbb{R}} \left(\frac{\Psi'\left(z, \Theta\right)}{\Psi\left(z, \Theta\right)}\right)^\ell \Psi\left(z\right) dz.
\end{aligned}
$$

E.2 Computation of (A1.1), (A1.2) and (A1.3) for dynamics location-scale family models.

We show that under assumptions of Theorem 5.1 the conditional moments in (**A1.1**), (**A1.2**) and (**A1.3**) converge, as the interval between observations shrinks to zero, to well defined limits. We first report the increments of the processes $x_{kh}^{(h)}$, $\mu_{kh}^{(h)}$ and $\lambda_{kh}^{(h)}$:

$$\begin{aligned}
x_{kh}^{(h)} - x_{(k-1)h}^{(h)} &= \mu_{kh}^{(h)} h + \Lambda \left(\lambda_{kh}^{(h)}\right)^{1/2} \epsilon_{kh}^{(h)} \\
\mu_{(k+1)h}^{(h)} - \mu_{kh}^{(h)} &= c_h - (1-b_h)\mu_{kh}^{(h)} + a_h s_\mu(\mu_{kh}^{(h)}, \lambda_{kh}^{(h)}) \nabla_{kh,\mu}^{(h)} \\
\lambda_{(k+1)h}^{(h)} - \lambda_{kh}^{(h)} &= \omega_h - (1-\beta_h)\lambda_{kh}^{(h)} + \alpha_h s_\lambda(\mu_{kh}^{(h)}, \lambda_{kh}^{(h)}) \nabla_{kh,\lambda}^{(h)}
\end{aligned}$$

Note that the expression for the increments of $\lambda_{kh}^{(h)}$ coincides with that of the scale family models. For this reason, some of the computations involving $\lambda_{kh}^{(h)}$ are omitted.

*Condition (**A1.1**)*

$$h^{-1}\mathbb{E}\left[\left(x_{kh}^{(h)} - x_{(k-1)h}^{(h)}\right)|\mathcal{F}_{kh}^{(h)}\right] = \mu_{kh}^{(h)} + h^{-1}\Lambda\left(\lambda_{kh}^{(h)}\right)^{1/2}\mathbb{E}\left[\epsilon_{kh}^{(h)}|\mathcal{F}_{kh}^{(h)}\right] = \mu_{kh}^{(h)} \tag{E.1}$$

$$h^{-1}\mathbb{E}\left[\left(\mu_{(k+1)h}^{(h)} - \mu_{kh}^{(h)}\right)|\mathcal{F}_{kh}^{(h)}\right] = h^{-1}c_h + h^{-1}\left(b_h - 1\right)\mu_{kh}^{(h)} \tag{E.2}$$

$$h^{-1}\mathbb{E}\left[\left(\lambda_{(k+1)}^{(h)} - \lambda_{kh}^{(h)}\right)|\mathcal{F}_{kh}^{(h)}\right] = h^{-1}\omega_h + h^{-1}\left(\beta_h - 1\right)\lambda_{kh}^{(h)}, \tag{E.3}$$

where in the last step of (E.2) we use the fact that $\mathbb{E}\left[\nabla_{kh,\mu}^{(h)}|\mathcal{F}_{kh}^{(h)}\right] = 0$. In particular, under the assumptions of Theorem 5.1 the previous expressions are finite as $h \downarrow 0$.

*Condition (**A1.2**)*

$$\begin{aligned}
&h^{-1}\mathbb{E}\left[\left(x_{kh}^{(h)} - x_{(k-1)h}^{(h)}\right)^2|\mathcal{F}_{kh}^{(h)}\right] \\
&= h\left(\mu_{kh}^{(h)}\right)^2 + 2\mu_{kh}^{(h)}\sqrt{\Lambda\left(\lambda_{kh}^{(h)}\right)}\mathbb{E}\left[\epsilon_{kh}^{(h)}|\mathcal{F}_{kh}^{(h)}\right] + h^{-1}\Lambda\left(\lambda_{kh}^{(h)}\right)\mathbb{E}\left[\left(\epsilon_{kh}^{(h)}\right)^2|\mathcal{F}_{kh}^{(h)}\right] \\
&= h\left(\mu_{kh}^{(h)}\right)^2 + \Lambda\left(\lambda_{kh}^{(h)}\right)\int_{\mathbb{R}} z^2\Psi\left(z\right)\,dz
\end{aligned} \tag{E.4}$$

In particular, when $h \downarrow 0$ we obtain that:

$$h^{-1}\mathbb{E}\left[\left(x_{kh}^{(h)} - x_{(k-1)h}^{(h)}\right)^2|\mathcal{F}_{kh}^{(h)}\right] = \Lambda(\lambda_{kh}^{(h)})\zeta^{(2)} + o\left(1\right). \tag{E.5}$$

Now, let us compute:

$$\begin{aligned}
&h^{-1}\mathbb{E}\left[\left(\mu_{(k+1)h}^{(h)} - \mu_{kh}^{(h)}\right)^2|\mathcal{F}_{kh}^{(h)}\right] \\
&= h^{-1}c_h^2 + h^{-1}\left(b_h - 1\right)^2\left(\mu_{kh}^{(h)}\right)^2 + h^{-1}a_h^2 s_\mu(\mu_{kh}^{(h)}, \lambda_{kh}^{(h)})^2\mathbb{E}\left[\left(\nabla_{kh,\mu}^{(h)}\right)^2|\mathcal{F}_{kh}^{(h)}\right] + 2h^{-1}c_h\,\left(b_h - 1\right)\mu_{kh}^{(h)} \\
&= h^{-1}c_h^2 + h^{-1}\left(b_h - 1\right)^2\left(\mu_{kh}^{(h)}\right)^2 + h^{-1}a_h^2\mathbb{E}\left[\left(\nabla_{kh,\mu}^{(h)}\right)^2|\mathcal{F}_{kh}^{(h)}\right]^{-1} + 2h^{-1}c_h\,\left(b_h - 1\right)\mu_{kh}^{(h)} \\
&= h^{-1}c_h^2 + h^{-1}\left(b_h - 1\right)^2\left(\mu_{kh}^{(h)}\right)^2 + h^{-2}a_h^2\Lambda\left(\lambda_{kh}^{(h)}\right)\left[\int_{-\infty}^{\infty}\left(\frac{\Psi'\left(z\right)}{\Psi\left(z\right)}\right)^2\Psi\left(z\right)\,dz\right]^{-1} + 2h^{-1}c_h\left(b_h - 1\right)\mu_{kh}^{(h)}
\end{aligned} \tag{E.6}$$

where in the last two lines we used the fact that $s_\mu(\mu_{kh}^{(h)}, \lambda_{kh}^{(h)}) = \mathbb{E}\left[\left(\nabla_{kh,\mu}^{(h)}\right)^2|\mathcal{F}_{kh}^{(h)}\right]^{-1}$ and the result of Theorem 5.2. Using the assumptions of Theorem (5.1) and noting that the integral above coincides with the second moment of the score of Ψ with respect to z, $\chi_z = \mathbb{E}\left[\nabla_z^2\right]$, we obtain:

$$h^{-1}\mathbb{E}\left[\left(\mu_{(k+1)h}^{(h)} - \mu_{kh}^{(h)}\right)^2 |\mathcal{F}_{kh}^{(h)}\right] = a^2\Lambda\left(\lambda_{kh}^{(h)}\right)\chi_z^{-1} + o(1) \tag{E.7}$$

Now:

$$h^{-1}\mathbb{E}\left[\left(\lambda_{(k+1)h}^{(h)} - \lambda_{kh}^{(h)}\right)^2 |\mathcal{F}_{kh}^{(h)}\right] = \alpha^2 s_\lambda\left(\mu_{kh}^{(h)}, \lambda_{kh}^{(h)}\right)^2 \xi_{kh,\lambda}^{(2)} + o\left(1\right) \tag{E.8}$$

$$\begin{aligned}
&h^{-1}\mathbb{E}\left[\left(x_{kh}^{(h)} - x_{(k-1)h}^{(h)}\right)\left(\mu_{(k+1)h}^{(h)} - \mu_{kh}^{(h)}\right)|\mathcal{F}_{kh}^{(h)}\right] \\
&= h^{-1}\mathbb{E}\left[\left(h\mu_{kh}^{(h)} + \sqrt{\Lambda(\lambda_{kh}^{(h)})}\epsilon_{kh}^{(h)}\right)\left(c_h + (b_h - 1)\,\mu_{kh}^{(h)} + a_h s_\mu(\mu_{kh}^{(h)}, \lambda_{kh}^{(h)})\nabla_{kh,\mu}^{(h)}\right)|\mathcal{F}_{kh}^{(h)}\right] \\
&= \mu_{kh}^{(h)} c_h - (1 - b_h)\left(\mu_{kh}^{(h)}\right)^2 + h^{-1}\sqrt{\Lambda\left(\lambda_{kh}^{(h)}\right)} a_h s_\mu(\mu_{kh}^{(h)}, \lambda_{kh}^{(h)})\mathbb{E}\left[\epsilon_{kh}^{(h)}\nabla_{kh,\mu}^{(h)}|\mathcal{F}_{kh}^{(h)}\right]
\end{aligned} \tag{E.9}$$

We now compute the expectation $\mathbb{E}\left[\epsilon_{kh}^{(h)}\nabla_{kh,\mu}^{(h)}|\mathcal{F}_{kh}^{(h)}\right]$ integrating by parts and using the expression of $\nabla_{kh,\mu}^{(h)}$ recovered in Theorem E.1 (we omit the dependence on the sampling grid to avoid clutter):

$$\mathbb{E}\left[\epsilon\nabla_\mu\right] = -\frac{h}{\sqrt{\Lambda\left(\lambda\right)}}\int_{\mathbb{R}}\Psi^{'}\left(z\right) z dz = \frac{h}{\sqrt{\Lambda\left(\lambda\right)}}$$

Combining this equation with Eq. (E.9) and by using the assumptions of Theorem (5.1) we have:

$$h^{-1}\mathbb{E}\left[\left(x_{kh}^{(h)} - x_{(k-1)h}^{(h)}\right)\left(\mu_{(k+1)h}^{(h)} - \mu_{kh}^{(h)}\right)|\mathcal{F}_{kh}^{(h)}\right] = a\Lambda(\lambda_{kh}^{(h)})\chi_z^{-1} + o(1) \tag{E.10}$$

Then:

$$\begin{aligned}
&h^{-1}\mathbb{E}\left[\left(x_{kh}^{(h)} - x_{(k-1)h}^{(h)}\right)\left(\lambda_{(k+1)}^{(h)} - \lambda_{kh}^{(h)}\right)|\mathcal{F}_{kh}^{(h)}\right] \\
&= h^{-1}\mathbb{E}\left[\left(h\mu_{kh}^{(h)} + \sqrt{\Lambda\left(\lambda_{kh}^{(h)}\right)}\epsilon_{kh}^{(h)}\right)\left(\omega_h + (\beta_h - 1)\,\lambda_{kh}^{(h)} + \alpha_h s_\lambda\left(\mu_{kh}^{(h)}, \lambda_{kh}^{(h)}\right)\nabla_{kh,\lambda}^{(h)}\right)|\mathcal{F}_{kh}^{(h)}\right] \\
&= \mu_{kh}^{(h)}\omega_h + \mu_{kh}^{(h)}\left(\beta_h - 1\right)\lambda_{kh}^{(h)},
\end{aligned}$$

which is an $o(1)$ under assumptions of Theorem (5.1). Finally we compute:

$$\begin{aligned}
&h^{-1}\mathbb{E}\left[\left(\lambda_{kh}^{(h)} - \lambda_{(k-1)}^{(h)}\right)\left(\mu_{(k+1)h}^{(h)} - \mu_{kh}^{(h)}\right)|\mathcal{F}_{kh}^{(h)}\right] \\
&= h^{-1}\omega_h c_h + h^{-1}\omega_h\left(b_h - 1\right)\mu_{kh}^{(h)} + \left(\beta_h - 1\right)c_h\lambda_{kh}^{(h)} + \left(\beta_h - 1\right)\left(b_h - 1\right)\lambda_{kh}^{(h)}\mu_{kh}^{(h)} \\
+ \quad &\alpha_h a_h s_\lambda\left(\mu_{kh}^{(h)}, \lambda_{kh}^{(h)}\right)s_\mu\left(\mu_{kh}^{(h)}, \lambda_{kh}^{(h)}\right)\mathbb{E}\left[\nabla_{kh,\lambda}^{(h)}\nabla_{kh,\mu}^{(h)}|\mathcal{F}_{kh}^{(h)}\right]
\end{aligned} \tag{E.11}$$

It is straightforward to check that the last conditional expectation is zero. Indeed:

$$\begin{aligned}
\mathbb{E}\left[\nabla_\lambda\nabla_\mu\right] &= \frac{1}{2}\frac{\Lambda^{'}\left(\lambda\right)}{\Lambda\left(\lambda\right)}\sqrt{\frac{h}{\Lambda\left(\lambda\right)}}\int_{\mathbb{R}}\left(1 + \frac{\Psi^{'}\left(z\right)}{\Psi\left(z\right)}z\right)\left(\frac{\Psi^{'}\left(z\right)}{\Psi\left(z\right)}\right)\Psi\left(z\right)\,dz \\
&= \int_{\mathbb{R}}\Psi^{'}\left(z\right)\,dz + \int_{\mathbb{R}}\frac{\Psi^{'}\left(z\right)^2}{\Psi\left(z\right)}z\,dz = 0,
\end{aligned}$$

since the $p.d.f.$ $\Psi(\cdot)$ is symmetric by assumption. So, under assumptions of Theorem (5.1) we have:

$$h^{-1}\mathbb{E}\left[\left(\lambda_{kh}^{(h)}-\lambda_{(k-1)}^{(h)}\right)\left(\mu_{(k+1)h}^{(h)}-\mu_{kh}^{(h)}\right)|\mathcal{F}_{kh}^{(h)}\right]=o\left(1\right).$$

*Condition (**A1.3**)*

By straightforward computation as in Nelson (1990), by setting $\omega_h = h\omega$, $\alpha_h = h^{1/2}\alpha$, $\beta_h = 1-h\theta$, $b_h = 1-h\vartheta$ and $a_h = ha$ and by using Theorems 2.2 - 2.3, condition (**A1.3**) holds for $\delta = 2$. Indeed:

$$h^{-1}\mathbb{E}\left[\left(x_{kh}^{(h)}-x_{(k-1)h}^{(h)}\right)^4|\mathcal{F}_{kh}^{(h)}\right]=h\,\Lambda^2\left(\lambda_{kh}^{(h)}\right)\zeta^{(4)} \tag{E.12}$$

$$h^{-1}\mathbb{E}\left[\left(\mu_{kh}^{(h)}-\mu_{(k-1)h}^{(h)}\right)^4|\mathcal{F}_{kh}^{(h)}\right]=O(h^{\varsigma}),\quad \varsigma\geq 4 \tag{E.13}$$

$$h^{-1}\mathbb{E}\left[\left(\lambda_{kh}^{(h)}-\lambda_{(k-1)h}^{(h)}\right)^4|\mathcal{F}_{kh}^{(h)}\right]=h\,\alpha^4 s_\lambda\left(\mu_{kh}^{(h)},\lambda_{kh}^{(h)}\right)^4\xi_{kh,\lambda}^{(4)}+O\left(h^{\varrho}\right),\quad \varrho\geq 3/2. \tag{E.14}$$

which converges to zero as $h\downarrow 0$.

F Proof of results of Section 5.2

We start by showing that for the class of multivariate conditional distributions the moments of the score are independent of h. We suppress, again, the dependence on t and we denote, for sake of simplicity, $\Sigma = \Sigma(f)$. We remind that $\mathcal{J}$ is the matrix defined implicitly as $\Sigma^{-1} = \mathcal{J}'\mathcal{J}$. In addition, we set $z = (\mathcal{J}y)/\sqrt{h}$. By using the latter notations, the matrix calculus of Abadir and Magnus (2005) we first compute

$$\frac{\partial\left(\frac{y'\Sigma^{-1}y}{h}\right)}{\partial \mathrm{vec}\left(\Sigma\right)'} = -\frac{1}{h}\left(\Sigma^{-1}y\otimes\Sigma^{-1}y\right)' = -\left(\frac{y'\mathcal{J}'\mathcal{J}}{\sqrt{h}}\otimes\frac{y'\mathcal{J}'\mathcal{J}}{\sqrt{h}}\right) = -\left(z'\mathcal{J}\otimes z'\mathcal{J}\right).$$

At this point we have (note that in what follow $\Psi'(\cdot)$ denotes the derivative of $\Psi(\cdot)$ with respect its argument):

$$\begin{aligned}\frac{\partial\log p\left(y|\Sigma;\Theta\right)}{\partial f'} &= \frac{\partial}{\partial f'}\left(-\frac{1}{2}\log|\Sigma|+\log\Psi\left(\frac{y'\Sigma^{-1}y}{h},\Theta\right)\right)\\ &= \left[-\frac{1}{2}\frac{1}{|\Sigma|}\frac{\partial|\Sigma|}{\partial\mathrm{vec}\left(\Sigma\right)'}-\frac{\Psi'\left(z'z,\Theta\right)}{\Psi\left(z'z,\Theta\right)}\left(z'\mathcal{J}\otimes z'\mathcal{J}\right)\right]\frac{\partial\mathrm{vec}\left(\Sigma\right)}{\partial\mathrm{vech}\left(\Sigma\right)'}\frac{\partial\mathrm{vech}\left(\Sigma\right)}{\partial f'}\\ &= \left[-\frac{1}{2}\mathrm{vec}\left(\Sigma^{-1}\right)'-\frac{\Psi'\left(z'z,\Theta\right)}{\Psi\left(z'z,\Theta\right)}z_{\otimes}'\mathcal{J}_{\otimes}\right]\mathcal{D}_n\Upsilon\quad\text{where}\quad\Upsilon=\Upsilon\left(f\right)=\frac{\partial\mathrm{vech}\left(\Sigma\right)}{\partial f'}.\end{aligned}$$

Therefore:

$$\nabla=\frac{\partial\log p\left(y|\Sigma\left(f\right);\Theta'\right)}{\partial f}=-\frac{1}{2}\Upsilon'\mathcal{D}_n'\left[\mathrm{vec}\left(\Sigma^{-1}\right)-\mathcal{J}_{\otimes}'z_{\otimes}'\frac{\Psi'\left(z'z,\Theta\right)}{\Psi\left(z'z,\Theta\right)}\right]. \tag{F.1}$$

Whence the moments of the score do not depend on h.

F.1 Proof of Theorem 5.3

We start by reporting the increments of the processes $x_{kh}^{(h)}$ and $f_{kh}^{(h)}$:

$$\begin{aligned} x_{kh}^{(h)} - x_{(k-1)h}^{(h)} &= \eta_{kh}^{(h)} \\ f_{(k+1)h}^{(h)} - f_{kh}^{(h)} &= c_h - (I_k - B_h) + A_h S(f_{kh}^{(h)}) \nabla_{kh}^{(h)}. \end{aligned} \tag{F.2}$$

Now, we show that under assumptions of Theorem 5.3 the conditional moments in (**A1.1**), (**A1.2**) and (**A1.3**) converge, as the interval between observations shrinks to zero, to well defined limits. We start from:

*Condition (**A1.1**)*

$$\begin{aligned} &h^{-1}\mathbb{E}\left[x_{kh}^{(h)} - x_{(k-1)h}^{(h)} | \mathcal{F}_{kh}^{(h)}\right] = h^{-1}\mathbb{E}\left[\eta_{kh}^{(h)} | \mathcal{F}_{kh}^{(h)}\right] = 0_N. \\ &h^{-1}\mathbb{E}\left[f_{(k+1)h}^{(h)} - f_{kh}^{(h)} | \mathcal{F}_{kh}^{(h)}\right] = h^{-1}c_h - h^{-1}\left(I_k - B_h\right) + h^{-1}\mathbb{E}\left[\nabla_{kh}^{(h)} | \mathcal{F}_{kh}^{(h)}\right] = h^{-1}c_h - h^{-1}\left(I_k - B_h\right), \end{aligned}$$

since the conditional moment of the score is 0_K. Under the assumptions in Theorem 5.3 the previous expressions are finite as $h \downarrow 0$.

*Condition (**A1.2**)*

$$\begin{aligned} &h^{-1}\mathbb{E}\left[\left(x_{kh}^{(h)} - x_{(k-1)h}^{(h)}\right)\left(x_{kh}^{(h)} - x_{(k-1)h}^{(h)}\right)' | \mathcal{F}_{kh}^{(h)}\right] = h^{-1}\mathbb{E}\left[\eta_{kh}^{(h)} \eta_{kh}^{(h),'}\right] = \Sigma\left(f_{kh}^{(h)}\right). \\ &h^{-1}\mathbb{E}\left[\left(f_{(k+1)h}^{(h)} - f_{kh}^{(h)}\right)\left(f_{(k+1)h}^{(h)} - f_{kh}^{(h)}\right)'\right] = h^{-1}c_h c_h' + h^{-1}c_h\mathbb{E}\left[\nabla_{kh}^{(h),'} | \mathcal{F}_{kh}^{(h)}\right] S\left(f_{kh}^{(h)}\right)' A_h' \\ &-h^{-1}c_h f_{kh}^{(h),'}\left(I_k - B_h\right) + h^{-1}A_h S\left(f_{kh}^{(h)}\right)\mathbb{E}\left[\nabla_{kh}^{(h)} | \mathcal{F}_{kh}^{(h)}\right] c_h' + h^{-1}A_h S\left(f_{kh}^{(h)}\right)\mathbb{E}\left[\nabla_{kh}^{(h)}\nabla_{kh}^{(h),'} | \mathcal{F}_{kh}^{(h)}\right] \cdot \\ &\cdot S\left(f_{kh}^{(h)}\right)' A_h' - h^{-1}A_h S\left(f_{kh}^{(h)}\right)\mathbb{E}\left[\nabla_{kh}^{(h)} | \mathcal{F}_{kh}^{(h)}\right] f_{kh}^{(h),'}\left(I_k - B_h\right) - h^{-1}\left(I_k - B_h\right) f_{kh}^{(h)} c_h' \\ &-h^{-1}\left(I_k - B_h\right) f_{kh}^{(h)}\mathbb{E}\left[\nabla_{kh}^{(h),'} | \mathcal{F}_{kh}^{(h)}\right] S(f_{kh}^{(h)})' A_h' + h^{-1}\left(I_k - B_h\right) f_{kh}^{(h)} f_{kh}^{(h),'}\left(I_k - B_h\right). \end{aligned} \tag{F.3}$$

At this point, we use the fact that the first moment of the score is equal to zero together with the Assumptions of Theorem 5.3 to conclude that:

$$h^{-1}\mathbb{E}\left[\left(f_{(k+1)h}^{(h)} - f_{kh}^{(h)}\right)\left(f_{(k+1)h}^{(h)} - f_{kh}^{(h)}\right)'\right] = A\,\xi_{kh}^{(2)}\,A' + o(1).$$

Finally, we have:

$$\begin{aligned} &h^{-1}\mathbb{E}\left[\left(x_{kh}^{(h)} - x_{(k-1)h}^{(h)}\right)\left(f_{(k+1)h}^{(h)} - f_{kh}^{(h)}\right)' | \mathcal{F}_{kh}^{(h)}\right] = h^{-1}c_h\mathbb{E}\left[\eta_{kh}^{(h)} | \mathcal{F}_{kh}^{(h)}\right] \\ &+h^{-1}\mathbb{E}\left[\eta_{kh}^{(h)}\nabla_{kh}^{(h),'} | \mathcal{F}_{kh}^{(h)}\right] S\left(f_{kh}^{(h)}\right)' A_h' - h^{-1}\mathbb{E}\left[\eta_{kh}^{(h)} | \mathcal{F}_{kh}^{(h)}\right] f_{kh}^{(h),'}\left(I_k - B_h\right)' = 0_{(N\times K)}, \end{aligned}$$

where we use that $\mathbb{E}\left[\eta_{kh}^{(h)} | \mathcal{F}_{kh}^{(h)}\right] = 0$ and the fact the score is odd.

*Condition (**A1.3**)*

By straightforward computation as in Nelson (1990), under the Assumptions of Theorem 5.3, condition (**A1.3**) holds for $\delta = 2$ and by setting $c_h = ch$, $(I_k - B_h) = h\Lambda$ and $A_h = \sqrt{h}A$.

References

Abadir, K. M., Magnus, J. R., 2005. Matrix algebra. Vol. 1. Cambridge University Press.

Donsker, M. D., 1951. An invariance principle for certain probability limit theorems. Mem. Amer. Math. Soc. 6, 12 pp.

Ethier, S. N., Kurtz, T. G., 1986. Markov processes: characterization and convergence. John Wiley & Sons.

Hudson, W. N., Tucker, H. G., 1979. Asymptotic independence in the multivariate central limit theorem. The Annals of Probability 7 (4), 662–671.

Kushner, H. J., 1984. Approximation and weak convergence methods for random processes, with applications to stochastic systems theory. Vol. 6. MIT press.

Lukacs, E., 1963. Applications of characteristic functions in statistics. Sankhya: The Indian Journal of Statistics, Series A (1961-2002) 25 (2), 175–188.

Nelson, D. B., 1990. Arch models as diffusion approximations. Journal of Econometrics 45 (1-2), 7–38.

Pierre, P. A., 1971. Infinitely divisible distributions, conditions for independence, and central limit theorems. Journal of Mathematical Analysis and Applications 33 (2), 341–354.

Polymenis, A., et al., 2017. Asymptotic relationship between sample mean and sample variance for autoregressive processes of order 1. Journal of Statistical and Econometric Methods 6 (1), 101–108.

Stroock, D. W., Varadhan, S. S., 1979. Multidimensional diffusion processes. Springer.