Supplemental Appendices for Distributional tests in multivariate dynamic models with Normal and Student *t* innovations*

Javier Mencía
Bank of Spain
<javier.mencia@bde.es>

Enrique Sentana
CEMFI
<sentana@cemfi.es>

June 2010

*Address for correspondence: Casado del Alisal 5, E-28014 Madrid, Spain, tel: +34 91 429 05 51, fax: +34 91 429 1056.

1 The density function

Consider an N-dimensional random vector $\mathbf{u}$, which can be expressed in terms of the following Location-Scale Mixture of Normals (*LSMN*):

$$\mathbf{u} = \boldsymbol{\alpha} + \xi^{-1}\boldsymbol{\Upsilon}\boldsymbol{\beta} + \xi^{-1/2}\boldsymbol{\Upsilon}^{1/2}\mathbf{r}, \tag{1}$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are N-dimensional vectors, $\boldsymbol{\Upsilon}$ is a positive definite matrix of order N, $\mathbf{r} \sim N(\mathbf{0}, \mathbf{I}_N)$, and ξ is an independent positive mixing variable. If the mixing variable follows a Generalised Inverse Gaussian distribution (*GIG*), then the distribution of $\mathbf{u}$ will be the Generalised Hyperbolic distribution (*GH*) introduced by Barndorff-Nielsen (1977). More explicitly, if $\xi \sim GIG(-\nu, \gamma, \delta)$ then the density of the $N \times 1$ *GH* random vector $\mathbf{u}$ will be given by

$$\begin{aligned} f_{GH}(\mathbf{u}) &= \frac{\left(\frac{\gamma}{\delta}\right)^{\nu}}{(2\pi)^{\frac{N}{2}}\left[\boldsymbol{\beta}'\boldsymbol{\Upsilon}\boldsymbol{\beta}+\gamma^2\right]^{\nu-\frac{N}{2}}|\boldsymbol{\Upsilon}|^{\frac{1}{2}}K_{\nu}(\delta\gamma)}\left\{\sqrt{\boldsymbol{\beta}'\boldsymbol{\Upsilon}\boldsymbol{\beta}+\gamma^2}\,\delta q\left[\delta^{-1}(\mathbf{u}-\boldsymbol{\alpha})\right]\right\}^{\nu-\frac{N}{2}} \\ &\quad \times K_{\nu-\frac{N}{2}}\left\{\sqrt{\boldsymbol{\beta}'\boldsymbol{\Upsilon}\boldsymbol{\beta}+\gamma^2}\,\delta q\left[\delta^{-1}(\mathbf{u}-\boldsymbol{\alpha})\right]\right\}\exp\left[\boldsymbol{\beta}'(\mathbf{u}-\boldsymbol{\alpha})\right], \end{aligned} \tag{2}$$

where $-\infty < \nu < \infty$, $\gamma > 0$, $q\left[\delta^{-1}(\mathbf{u}-\boldsymbol{\alpha})\right] = \sqrt{1+\delta^{-2}(\mathbf{u}-\boldsymbol{\alpha})'\boldsymbol{\Upsilon}^{-1}(\mathbf{u}-\boldsymbol{\alpha})}$ and $K_{\nu}(\cdot)$ is the modified Bessel function of the third kind (see Abramowitz and Stegun, 1965, p. 374, as well as Section 4).

Given that δ and $\boldsymbol{\Upsilon}$ are not separately identified, Barndorff-Nielsen and Shephard (2001) set the determinant of $\boldsymbol{\Upsilon}$ equal to 1. However, it is more convenient to set $\delta = 1$ instead in order to reparametrise the *GH* distribution so that it has mean vector $\mathbf{0}$ and covariance matrix $\mathbf{I}_N$. It is then straightforward to use Proposition 1 in Mencía and Sentana (2009) to obtain a standardised *GH* distribution. Specifically, we set $\delta = 1$, $\boldsymbol{\alpha} = -c(\boldsymbol{\beta}, \nu, \gamma)\boldsymbol{\beta}$ and

$$\boldsymbol{\Upsilon} = \frac{\gamma}{R_{\nu}(\gamma)}\left[\mathbf{I}_N + \frac{c(\boldsymbol{\beta},\nu,\gamma)-1}{\boldsymbol{\beta}'\boldsymbol{\beta}}\boldsymbol{\beta}\boldsymbol{\beta}'\right], \tag{3}$$

where

$$c(\boldsymbol{\beta},\nu,\gamma) = \frac{-1+\sqrt{1+4[D_{\nu+1}(\gamma)-1]\boldsymbol{\beta}'\boldsymbol{\beta}}}{2[D_{\nu+1}(\gamma)-1]\boldsymbol{\beta}'\boldsymbol{\beta}}. \tag{4}$$

where $R_{\nu}(\gamma) = K_{\nu+1}(\gamma)/K_{\nu}(\gamma)$ and $D_{\nu+1}(\gamma) = K_{\nu+2}(\gamma)K_{\nu}(\gamma)/K^2_{\nu+1}(\gamma)$. Thus, the distribution of $\mathbf{u}$ depends on two shape parameters, ν and γ, and a vector of N skewness parameters, denoted by $\boldsymbol{\beta}$.

One of the most attractive properties of the GH distribution is that it contains as particular cases several of the most important multivariate distributions already used in the literature. The best known examples are:

• **Normal**, which can be achieved in three different ways: (i) when $\nu \to -\infty$ or (ii) $\nu \to +\infty$, regardless of the values of γ and $\boldsymbol{\beta}$; and (iii) when $\gamma \to \infty$ irrespective of the values of ν and $\boldsymbol{\beta}$.

• **Symmetric Student** t, obtained when $-\infty < \nu < -2$, $\gamma = 0$ and $\boldsymbol{\beta} = \mathbf{0}$.

• **Asymmetric Student** t, which is like its symmetric counterpart except that the vector $\boldsymbol{\beta}$ of skewness parameters is no longer zero.

• **Asymmetric Normal-Gamma**, which is obtained when $\gamma = 0$ and $0 < \nu < \infty$ (see Madan and Milne, 1991).

• **Normal Inverse Gaussian**, for $\nu = -.5$ (see Aas, Dimakos, and Haff, 2005).

• **Hyperbolic**, for $\nu = 1$ (see Chen, Härdle, and Jeong, 2008)

• **Asymmetric Laplace**, for $\nu = 1$ and $\gamma = 0$ (see Cajigas and Urga, 2007).

2 Skewness and kurtosis of GH distributions

We can tediously show that

$$
\begin{gathered}
E\left[vec\left(\boldsymbol{\varepsilon}^*\boldsymbol{\varepsilon}^{*\prime}\right)\boldsymbol{\varepsilon}^{*\prime}\right] = E\left[\left(\boldsymbol{\varepsilon}^*\otimes\boldsymbol{\varepsilon}^*\right)\boldsymbol{\varepsilon}^{*\prime}\right] \\
= c^3(\boldsymbol{\beta},\nu,\gamma)\left[\frac{K_{\nu+3}\left(\gamma\right)K_{\nu}^2\left(\gamma\right)}{K_{\nu+1}^3\left(\gamma\right)} - 3D_{\nu+1}\left(\gamma\right) + 2\right] vec\left(\boldsymbol{\beta}\boldsymbol{\beta}'\right)\boldsymbol{\beta}' \\
+c(\boldsymbol{\beta},\nu,\gamma)\left[D_{\nu+1}\left(\gamma\right)\text{-}1\right]\left(\mathbf{K}_{NN}+\mathbf{I}_{N^2}\right)\left(\boldsymbol{\beta}\otimes\mathbf{A}\right) + c(\boldsymbol{\beta},\nu,\gamma)\left[D_{\nu+1}\left(\gamma\right)\text{-}1\right] vec\left(\mathbf{A}\right)\boldsymbol{\beta}', \qquad (5)
\end{gathered}
$$

and

$$
\begin{gathered}
E\left[vec\left(\boldsymbol{\varepsilon}^*\boldsymbol{\varepsilon}^{*\prime}\right)vec'\left(\boldsymbol{\varepsilon}^*\boldsymbol{\varepsilon}^{*\prime}\right)\right] = E\left[\boldsymbol{\varepsilon}^*\boldsymbol{\varepsilon}^{*\prime}\otimes\boldsymbol{\varepsilon}^*\boldsymbol{\varepsilon}^{*\prime}\right] \\
= c^4(\boldsymbol{\beta},\nu,\gamma)\left[\frac{K_{\nu+4}\left(\gamma\right)K_{\nu}^3\left(\gamma\right)}{K_{\nu+1}^4\left(\gamma\right)} - 4\frac{K_{\nu+3}\left(\gamma\right)K_{\nu}^2\left(\gamma\right)}{K_{\nu+1}^3\left(\gamma\right)} + 6D_{\nu+1}\left(\gamma\right) - 3\right] vec\left(\boldsymbol{\beta}\boldsymbol{\beta}'\right)vec'\left(\boldsymbol{\beta}\boldsymbol{\beta}'\right) \\
+c^2(\boldsymbol{\beta},\nu,\gamma)\left[\frac{K_{\nu+3}\left(\gamma\right)K_{\nu}^2\left(\gamma\right)}{K_{\nu+1}^3\left(\gamma\right)} - 2D_{\nu+1}\left(\gamma\right) + 1\right] \\
\times\left\{vec\left(\boldsymbol{\beta}\boldsymbol{\beta}'\right)vec'\left(\mathbf{A}\right) + vec\left(\mathbf{A}\right)vec'\left(\boldsymbol{\beta}\boldsymbol{\beta}'\right) + \left(\mathbf{K}_{NN}+\mathbf{I}_{N^2}\right)\left[\boldsymbol{\beta}\boldsymbol{\beta}'\otimes\mathbf{A}\right]\left(\mathbf{K}_{NN}+\mathbf{I}_{N^2}\right)\right\} \\
+D_{\nu+1}\left(\gamma\right)\left\{\left[\mathbf{A}\otimes\mathbf{A}\right]\left(\mathbf{K}_{NN}+\mathbf{I}_{N^2}\right) + vec\left(\mathbf{A}\right)vec'\left(\mathbf{A}\right)\right\}, \qquad (6)
\end{gathered}
$$

where

$$
\mathbf{A} = \mathbf{I}_N + \frac{c(\boldsymbol{\beta},\nu,\gamma) - 1}{\boldsymbol{\beta}'\boldsymbol{\beta}}\boldsymbol{\beta}\boldsymbol{\beta}',
$$

and $\mathbf{K}_{NN}$ is the commutation matrix (see Magnus and Neudecker, 1988). In this respect, note that Mardia's (1970) coefficient of multivariate excess kurtosis will be -1 plus the trace of the fourth moment above divided by $N(N+2)$.

Notice that there are $\binom{N+2}{3}$ and $\binom{N+3}{4}$ non-repeated terms in (5) and (6), respectively (see Yañez et al., 1999). For the particular case of $N=2$, we can write these terms as

$$E(\varepsilon_1^{*3}) = c^3(\boldsymbol{\beta},\nu,\gamma)\left[\frac{K_{\nu+3}(\gamma)\,K_{\nu}^2(\gamma)}{K_{\nu+1}^3(\gamma)} - 3D_{\nu+1}(\gamma)+2\right]\beta_1^3 + 3c(\boldsymbol{\beta},\nu,\gamma)\left[D_{\nu+1}(\gamma)\text{-}1\right]a_{11}\beta_1,$$

$$\begin{aligned} E(\varepsilon_1^{*2}\varepsilon_2^*) &= c^3(\boldsymbol{\beta},\nu,\gamma)\left[\frac{K_{\nu+3}(\gamma)\,K_{\nu}^2(\gamma)}{K_{\nu+1}^3(\gamma)} - 3D_{\nu+1}(\gamma)+2\right]\beta_1^2\beta_2 \\ &\quad + c(\boldsymbol{\beta},\nu,\gamma)\left[D_{\nu+1}(\gamma)\text{-}1\right](2a_{12}\beta_1 + a_{11}\beta_2), \end{aligned}$$

$$\begin{aligned} E\left[\varepsilon_1^{*4}\right] &= c^4(\boldsymbol{\beta},\nu,\gamma)\left[\frac{K_{\nu+4}(\gamma)\,K_{\nu}^3(\gamma)}{K_{\nu+1}^4(\gamma)} - 4\frac{K_{\nu+3}(\gamma)\,K_{\nu}^2(\gamma)}{K_{\nu+1}^3(\gamma)} + 6D_{\nu+1}(\gamma)-3\right]\beta_1^4 \\ &\quad + 6c^2(\boldsymbol{\beta},\nu,\gamma)\left[\frac{K_{\nu+3}(\gamma)\,K_{\nu}^2(\gamma)}{K_{\nu+1}^3(\gamma)} - 2D_{\nu+1}(\gamma)+1\right]a_{11}\beta_1^2 + 3D_{\nu+1}(\gamma)\,a_{11}^2, \end{aligned}$$

$$\begin{aligned} E\left[\varepsilon_1^{*3}\varepsilon_2^*\right] &= c^4(\boldsymbol{\beta},\nu,\gamma)\left[\frac{K_{\nu+4}(\gamma)\,K_{\nu}^3(\gamma)}{K_{\nu+1}^4(\gamma)} - 4\frac{K_{\nu+3}(\gamma)\,K_{\nu}^2(\gamma)}{K_{\nu+1}^3(\gamma)} + 6D_{\nu+1}(\gamma)-3\right]\beta_1^3\beta_2 \\ &\quad + 3c^2(\boldsymbol{\beta},\nu,\gamma)\left[\frac{K_{\nu+3}(\gamma)\,K_{\nu}^2(\gamma)}{K_{\nu+1}^3(\gamma)} - 2D_{\nu+1}(\gamma)+1\right](a_{12}\beta_1^2 + a_{11}\beta_1\beta_2) + 3D_{\nu+1}(\gamma)\,a_{11}a_{12}, \end{aligned}$$

and

$$\begin{aligned} E\left[\varepsilon_1^{*2}\varepsilon_2^{*2}\right] &= c^4(\boldsymbol{\beta},\nu,\gamma)\left[\frac{K_{\nu+4}(\gamma)\,K_{\nu}^3(\gamma)}{K_{\nu+1}^4(\gamma)} - 4\frac{K_{\nu+3}(\gamma)\,K_{\nu}^2(\gamma)}{K_{\nu+1}^3(\gamma)} + 6D_{\nu+1}(\gamma)-3\right]\beta_1^2\beta_2^2 \\ &\quad + c^2(\boldsymbol{\beta},\nu,\gamma)\left[\frac{K_{\nu+3}(\gamma)\,K_{\nu}^2(\gamma)}{K_{\nu+1}^3(\gamma)} - 2D_{\nu+1}(\gamma)+1\right](a_{22}\beta_1^2 + 4a_{12}\beta_1\beta_2 + a_{11}\beta_2^2) \\ &\quad + D_{\nu+1}(\gamma)\,(2a_{12}^2 + a_{11}a_{22}), \end{aligned}$$

where we use subindices to denote the elements of $\boldsymbol{\varepsilon}^*$, $\boldsymbol{\beta}$ and $\mathbf{A}$.

Under symmetry, the distribution of the standardised residuals $\boldsymbol{\varepsilon}^*$ is clearly elliptical, as it can be written as $\boldsymbol{\varepsilon}^* = \sqrt{\zeta/\xi}\sqrt{\gamma/R_\nu(\gamma)}\mathbf{u}$, where $\zeta \sim \chi_N^2$ and $\xi^{-1} \sim GIG(\nu,1,\gamma)$. This is confirmed by the fact that the third moment becomes 0, while

$$E\left[\boldsymbol{\varepsilon}^*\boldsymbol{\varepsilon}^{*\prime} \otimes \boldsymbol{\varepsilon}^*\boldsymbol{\varepsilon}^{*\prime}\right] = D_{\nu+1}(\gamma)\left\{\left[\mathbf{I}_N \otimes \mathbf{I}_N\right](K_{NN} + \mathbf{I}_{N^2}) + vec(\mathbf{I}_N)\,vec'(\mathbf{I}_N)\right\}.$$

In the symmetric case, therefore, the coefficient of multivariate excess kurtosis is simply $D_{\nu+1}(\gamma)$-1, which is always non-negative, but monotonically decreasing in γ and $|\nu|$.

3 The score function

Let $\mathbf{y}_t$ be a vector of N observed variables. To accommodate flexible specifications, we assume the following conditionally heteroskedastic dynamic regression model:

$$\left.\begin{array}{c}\mathbf{y}_t=\boldsymbol{\mu}_t(\boldsymbol{\theta})+\boldsymbol{\Sigma}_t^{\frac{1}{2}}(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t^*,\\ \boldsymbol{\mu}_t(\boldsymbol{\theta})=\boldsymbol{\mu}\left(I_{t-1};\boldsymbol{\theta}\right),\\ \boldsymbol{\Sigma}_t(\boldsymbol{\theta})=\boldsymbol{\Sigma}\left(I_{t-1};\boldsymbol{\theta}\right),\end{array}\right\} \tag{7}$$

where $\boldsymbol{\mu}()$ and $vech\left[\boldsymbol{\Sigma}()\right]$ are N and $N(N+1)/2$-dimensional vectors of functions known up to the $p\times 1$ vector of true parameter values, $\boldsymbol{\theta}_0$, I_{t-1} denotes the information set available at $t-1$, which contains past values of $\mathbf{y}_t$ and possibly other variables, $\boldsymbol{\Sigma}_t^{1/2}(\boldsymbol{\theta})$ is some $N\times N$ "square root" matrix such that $\boldsymbol{\Sigma}_t^{1/2}(\boldsymbol{\theta})\boldsymbol{\Sigma}_t^{1/2\prime}(\boldsymbol{\theta})=\boldsymbol{\Sigma}_t(\boldsymbol{\theta})$, and $\boldsymbol{\varepsilon}_t^*$ is a standardised *GH* vector martingale difference sequence satisfying $E(\boldsymbol{\varepsilon}_t^*|I_{t-1};\boldsymbol{\theta}_0)=\mathbf{0}$ and $V(\boldsymbol{\varepsilon}_t^*|I_{t-1};\boldsymbol{\theta}_0)=\mathbf{I}_N$. As a consequence, $E(\mathbf{y}_t|I_{t-1};\boldsymbol{\theta}_0)=\boldsymbol{\mu}_t(\boldsymbol{\theta}_0)$ and $V(\mathbf{y}_t|I_{t-1};\boldsymbol{\theta}_0)=\boldsymbol{\Sigma}_t(\boldsymbol{\theta}_0)$.

Importantly, given that $\boldsymbol{\varepsilon}_t^*$ is not generally observable, the choice of "square root" matrix is not irrelevant except in univariate *GH* models, or in multivariate *GH* models in which either $\boldsymbol{\Sigma}_t(\boldsymbol{\theta})$ is time-invariant or $\boldsymbol{\varepsilon}_t^*$ is spherical (i.e. $\boldsymbol{\beta}=\mathbf{0}$). But, if we parametrise $\boldsymbol{\beta}$ as a function of past information and a new vector of parameters $\mathbf{b}$ in the following way:

$$\boldsymbol{\beta}_t(\boldsymbol{\theta},\mathbf{b})=\boldsymbol{\Sigma}_t^{\frac{1}{2}\prime}(\boldsymbol{\theta})\mathbf{b}, \tag{8}$$

then it is straightforward to see that the resulting distribution of $\mathbf{y}_t$ conditional on I_{t-1} will not depend on the choice of $\boldsymbol{\Sigma}_t^{\frac{1}{2}}(\boldsymbol{\theta})$. Finally, it is analytically convenient to replace ν and γ by η and ψ, where $\eta=-.5\nu^{-1}$ and $\psi=(1+\gamma)^{-1}$, although we continue to use ν and γ in some equations for notational simplicity.

Using (2), we can express the density of $\mathbf{y}_t$ as

$$\begin{aligned}l\left(\mathbf{y}_t|I_{t-1};\boldsymbol{\phi}\right)=&-\frac{N}{2}\log(2\pi)-\frac{1}{2}\log\left|\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\right|-\frac{1+N\eta}{2\eta}\log(\gamma)+\frac{N}{2}\log R_\nu\left(\gamma\right)\\ &-\frac{1}{2}\log(c_t(\boldsymbol{\phi}))-\log K_\nu(\gamma)+\mathbf{b}'\boldsymbol{\varepsilon}_t(\boldsymbol{\theta})+c_t(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\\ &+\frac{1+N\eta}{4\eta}\log\left[\frac{\gamma}{R_\nu\left(\gamma\right)}c_t(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}+\gamma^2\right]\\ &-\frac{1+N\eta}{4\eta}\log(Q_t)+\log K_{\nu-.5N}\left[\sqrt{Q_t\left[\frac{\gamma}{R_\nu\left(\gamma\right)}c_t(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}+\gamma^2\right]}\right],\end{aligned} \tag{9}$$

where

$$Q_t = 1 + \frac{R_\nu(\gamma)}{\gamma}\left[\varsigma_t(\boldsymbol{\theta}) + 2c_t(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\varepsilon}_t(\boldsymbol{\theta}) + c_t^2(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\right]$$
$$-\frac{R_\nu(\gamma)}{\gamma}\frac{c_t(\boldsymbol{\phi})-1}{c_t(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\left[\mathbf{b}'\boldsymbol{\varepsilon}_t(\boldsymbol{\theta}) + c_t(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\right]^2$$

If the mean vector and covariance matrix specifications in (7) were constant, it would be potentially advantageous to use the EM algorithm for estimation purposes. In general dynamic models, though, the EM is not as useful because it typically requires numerical maximisation procedures at each M step. However, the EM principle can still be useful to derive the score function of the *GH* distribution. In this context, the procedure that we follow is divided in two parts. In the first step, we derive $l(\mathbf{y}_t|\xi_t, I_{t-1}; \boldsymbol{\phi})$ and $l(\xi_t|I_{t-1}; \boldsymbol{\phi})$ with respect to $\boldsymbol{\phi}$. Then, in the second step, we take the expected value of these derivatives given $I_T = \{\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_T\}$ and the parameter values.

Conditional on ξ_t, $\mathbf{y}_t$ is the following multivariate normal:

$$\mathbf{y}_t|\xi_t, I_{t-1} \sim N\left[\boldsymbol{\mu}_t(\boldsymbol{\theta}) + \boldsymbol{\Sigma}_t(\boldsymbol{\theta})c_t(\boldsymbol{\phi})\mathbf{b}\left[\frac{\gamma}{R_\nu(\gamma)}\frac{1}{\xi_t} - 1\right], \frac{\gamma}{R_\nu(\gamma)}\frac{1}{\xi_t}\boldsymbol{\Sigma}_t^*(\boldsymbol{\phi})\right],$$

where $c_t(\boldsymbol{\phi}) = c[\boldsymbol{\Sigma}_t^{\frac{1}{2}\prime}(\boldsymbol{\theta})\mathbf{b}, \nu, \gamma]$ and

$$\boldsymbol{\Sigma}_t^*(\boldsymbol{\phi}) = \boldsymbol{\Sigma}_t(\theta) + \frac{c_t(\boldsymbol{\phi})-1}{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta}).$$

If we define $\mathbf{p}_t = \mathbf{y}_t - \boldsymbol{\mu}_t(\boldsymbol{\theta}) + c_t(\boldsymbol{\phi})\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}$, then we have the following log-density

$$l(\mathbf{y}_t|\xi_t, I_{t-1}; \boldsymbol{\phi}) = \frac{N}{2}\log\left[\frac{\xi_t R_\nu(\gamma)}{2\pi\gamma}\right] - \frac{1}{2}\log|\boldsymbol{\Sigma}_t^*(\boldsymbol{\phi})| - \frac{\xi_t}{2}\frac{R_\nu(\gamma)}{\gamma}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})\mathbf{p}_t$$
$$+\mathbf{b}'\mathbf{p}_t - \frac{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}{2\xi_t}\frac{\gamma c_t(\boldsymbol{\phi})}{R_\nu(\gamma)}.$$

Similarly, ξ_t is distributed as a *GIG* with parameters $\xi_t|I_{t-1} \sim GIG(-\nu, \gamma, 1)$, with a log-likelihood given by

$$l(\xi_t|I_{t-1}; \boldsymbol{\phi}) = \nu\log\gamma - \log 2 - \log K_\nu(\gamma) - (\nu+1)\log\xi_t - \frac{1}{2}\left(\xi_t + \gamma^2\frac{1}{\xi_t}\right).$$

In order to determine the distribution of ξ_t given all the observable information I_T, we can exploit the serial independence of ξ_t given $I_{t-1}; \boldsymbol{\phi}$ to show that

$$f(\xi_t|I_T; \boldsymbol{\phi}) = \frac{f(\mathbf{y}_t, \xi_t|I_{t-1}; \boldsymbol{\phi})}{f(\mathbf{y}_t|I_{t-1}; \boldsymbol{\phi})} \propto f(\mathbf{y}_t|\xi_t, I_{t-1}; \boldsymbol{\phi})\, f(\xi_t|I_{t-1}; \boldsymbol{\phi})$$
$$\propto \xi_t^{\frac{N}{2}-\nu-1} \times \exp\left\{\frac{-1}{2}\left[\left(\frac{R_\nu(\gamma)}{\gamma}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})\mathbf{p}_t + 1\right)\xi_t + \left(\frac{\gamma c_t(\boldsymbol{\phi})}{R_\nu(\gamma)}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b} + \gamma^2\right)\frac{1}{\xi_t}\right]\right\},$$

which implies that

$$\xi_t|I_T;\phi \sim GIG\left(\frac{N}{2}-\nu, \sqrt{\frac{\gamma c_t(\boldsymbol{\phi})}{R_\nu(\gamma)}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}+\gamma^2}, \sqrt{\frac{R_\nu(\gamma)}{\gamma}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})\mathbf{p}_t+1}\right).$$

From here, we can use (66) and (67) to obtain the required moments. Specifically,

$$E(\xi_t|I_T;\boldsymbol{\phi}) = \frac{\sqrt{\frac{\gamma c_t(\boldsymbol{\phi})}{R_\nu(\gamma)}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}+\gamma^2}}{\sqrt{\frac{R_\nu(\gamma)}{\gamma}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}\mathbf{p}_t+1}}$$

$$\times R_{\frac{N}{2}-\nu}\left[\sqrt{\frac{\gamma c_t(\boldsymbol{\phi})}{R_\nu(\gamma)}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}+\gamma^2}\sqrt{\frac{R_\nu(\gamma)}{\gamma}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}\mathbf{p}_t+1}\right],$$

$$E\left(\frac{1}{\xi_t}\middle| I_T;\boldsymbol{\phi}\right) = \frac{\sqrt{\frac{R_\nu(\gamma)}{\gamma}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}\mathbf{p}_t+1}}{\sqrt{\frac{\gamma c_t(\boldsymbol{\phi})}{R_\nu(\gamma)}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}+\gamma^2}}$$

$$\times\frac{1}{R_{\frac{N}{2}-\nu-1}\left[\sqrt{\frac{\gamma c_t(\boldsymbol{\phi})}{R_\nu(\gamma)}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}+\gamma^2}\sqrt{\frac{R_\nu(\gamma)}{\gamma}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}\mathbf{p}_t+1}\right]},$$

$$E(\log\xi_t|\, I_T;\boldsymbol{\phi}) = \log\left(\sqrt{\frac{\gamma c_t(\boldsymbol{\phi})}{R_\nu(\gamma)}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}+\gamma^2}\right) - \log\left(\sqrt{\frac{R_\nu(\gamma)}{\gamma}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}\mathbf{p}_t+1}\right)$$

$$+\frac{\partial}{\partial x}\log K_x\left[\sqrt{\frac{\gamma c_t(\boldsymbol{\phi})}{R_\nu(\gamma)}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}+\gamma^2}\sqrt{\frac{R_\nu(\gamma)}{\gamma}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}\mathbf{p}_t+1}\right]\Bigg|_{x=\frac{N}{2}-\nu}.$$

If we put all the pieces together, we will finally have that

$$\begin{aligned}
\frac{\partial l(\mathbf{y}_t|\, I_{t-1};\boldsymbol{\phi})}{\partial\boldsymbol{\theta}'} &= -\frac{1}{2}vec'[\boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\theta})]\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial\boldsymbol{\theta}'} - f(I_T,\boldsymbol{\phi})\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})\frac{\partial\mathbf{p}_t}{\partial\boldsymbol{\theta}'}\\
&-\frac{1}{2}\frac{c_t(\boldsymbol{\phi})-1}{c_t(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\sqrt{1+4(D_{\nu+1}(\gamma)-1)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}vec'(\mathbf{b}\mathbf{b}')\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial\boldsymbol{\theta}'} + \mathbf{b}'\frac{\partial\mathbf{p}_t}{\partial\boldsymbol{\theta}'}\\
&+\frac{1}{2}f(I_T,\boldsymbol{\phi})[\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})\otimes\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})]\frac{\partial vec[\boldsymbol{\Sigma}_t^*(\boldsymbol{\phi})]}{\partial\boldsymbol{\theta}'}\\
&-\frac{1}{2}\frac{g(I_T,\boldsymbol{\phi})}{\sqrt{1+4(D_{\nu+1}(\gamma)-1)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}vec'(\mathbf{b}\mathbf{b}')\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial\boldsymbol{\theta}'},
\end{aligned}\tag{10}$$

$$\begin{aligned}
\frac{\partial l(\mathbf{y}_t|\, I_{t-1};\boldsymbol{\phi})}{\partial\mathbf{b}'} &= -\frac{c_t(\boldsymbol{\phi})-1}{c_t(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\sqrt{1+4(D_{\nu+1}(\gamma)-1)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\\
&-f(I_T,\boldsymbol{\phi})c_t(\boldsymbol{\phi})\mathbf{p}_t' + \boldsymbol{\varepsilon}_t' + f(I_T,\boldsymbol{\phi})\frac{c_t(\boldsymbol{\phi})-1}{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}(\mathbf{b}'\mathbf{p}_t)\\
&\times\left\{\frac{[c_t(\boldsymbol{\phi})-1](\mathbf{b}'\mathbf{p}_t)}{c_t^2(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\sqrt{1+4(D_{\nu+1}(\gamma)-1)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\right.\\
&\left.+\frac{\mathbf{p}_t'}{c_t(\boldsymbol{\phi})} - \frac{1}{\sqrt{1+4(D_{\nu+1}(\gamma)-1)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\right\}\\
&+\frac{[2-g(I_T,\boldsymbol{\phi})]}{\sqrt{1+4(D_{\nu+1}(\gamma)-1)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta}),
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial l\left(\mathbf{y}_t \mid I_{t-1};\boldsymbol{\phi}\right)}{\partial \eta} &= \frac{N}{2}\frac{\partial \log R_\nu\left(\gamma\right)}{\partial \eta} + \left(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b} - \frac{1}{2c_t(\boldsymbol{\phi})}\right)\frac{\partial c_t(\boldsymbol{\phi})}{\partial \eta} + \frac{\log\left(\gamma\right)}{2\eta^2} \\
&- \frac{\partial \log K_\nu\left(\gamma\right)}{\partial \eta} - \frac{1}{2\eta^2}E\left[\log \xi_t | Y_T;\boldsymbol{\phi}\right] - \frac{f\left(I_T,\boldsymbol{\phi}\right)}{2}\left\{\frac{\partial \log R_\nu\left(\gamma\right)}{\partial \eta}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})\mathbf{p}_t\right. \\
&\left. + \frac{\partial c_t(\boldsymbol{\phi})}{\partial \eta}\left[\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b} - \frac{\left(\mathbf{b}'\boldsymbol{\varepsilon}_t\right)^2}{c_t^2(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\right]\right\} \\
&- \frac{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}{2}g\left(I_T,\boldsymbol{\phi}\right)\left\{\frac{\partial c_t(\boldsymbol{\phi})}{\partial \eta} - c_t(\boldsymbol{\phi})\frac{\partial \log R_\nu\left(\gamma\right)}{\partial \eta}\right\},
\end{aligned} \tag{11}
$$

and

$$
\begin{aligned}
\frac{\partial l\left(\mathbf{y}_t \mid I_{t-1};\boldsymbol{\phi}\right)}{\partial \psi} &= \frac{N}{2}\frac{\partial \log R_\nu\left(\gamma\right)}{\partial \psi} + \frac{N}{2\psi\left(1-\psi\right)} + \left(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b} - \frac{1}{2c_t(\boldsymbol{\phi})}\right)\frac{\partial c_t(\boldsymbol{\phi})}{\partial \psi} \\
&+ \frac{1}{2\eta\psi\left(1-\psi\right)} - \frac{\partial \log K_\nu\left(\gamma\right)}{\partial \psi} - \frac{f\left(I_T,\boldsymbol{\phi}\right)}{2}\left\{\left[\frac{\partial \log R_\nu\left(\gamma\right)}{\partial \psi} + \frac{1}{\psi\left(1-\psi\right)}\right]\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})\mathbf{p}_t\right. \\
&\left. + \frac{\partial c_t(\boldsymbol{\phi})}{\partial \psi}\left[\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b} - \frac{\left(\mathbf{b}'\boldsymbol{\varepsilon}_t\right)^2}{c_t^2(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\right]\right\} \\
&- \frac{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}{2}g\left(I_T,\boldsymbol{\phi}\right)\left\{-\frac{c_t(\boldsymbol{\phi})}{\psi\left(1-\psi\right)} + \frac{\partial c_t(\boldsymbol{\phi})}{\partial \psi} - c_t(\boldsymbol{\phi})\frac{\partial \log R_\nu\left(\gamma\right)}{\partial \psi}\right\} + g\left(I_T,\boldsymbol{\phi}\right)\frac{R_\nu\left(\gamma\right)}{\psi^2},
\end{aligned} \tag{12}
$$

where

$$
\begin{aligned}
f\left(I_T,\boldsymbol{\phi}\right) &= \gamma^{-1}R_\nu\left(\gamma\right)E\left(\xi_t | I_T;\boldsymbol{\phi}\right), \\
g\left(I_T,\boldsymbol{\phi}\right) &= \gamma R_\nu^{-1}\left(\gamma\right)E\left(\xi_t^{-1} | I_T;\boldsymbol{\phi}\right),
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial vec[\boldsymbol{\Sigma}_t^*(\boldsymbol{\phi})]}{\partial \boldsymbol{\theta}'} &= \frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}'} + \frac{c_t(\boldsymbol{\phi})-1}{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\left\{\left[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\mathbf{b}' \otimes I_N\right] + \left[I_N \otimes \boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\mathbf{b}'\right]\right\}\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}'} \\
&+ \frac{c_t(\boldsymbol{\phi})-1}{\left[\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\right]^2}\left\{\frac{1}{\sqrt{1+4\left(D_{\nu+1}\left(\gamma\right)-1\right)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}} - 1\right\} \\
&\times vec\left[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\right]vec'\left(\mathbf{b}\mathbf{b}'\right)\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}'},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \mathbf{p}_t}{\partial \boldsymbol{\theta}'} &= -\frac{\partial \boldsymbol{\mu}_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + c_t(\boldsymbol{\phi})\left[\mathbf{b}' \otimes I_N\right]\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}'} \\
&+ \frac{c_t(\boldsymbol{\phi})-1}{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\frac{1}{\sqrt{1+4\left(D_{\nu+1}\left(\gamma\right)-1\right)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}vec'\left(\mathbf{b}\mathbf{b}'\right)\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}'},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial c_t(\boldsymbol{\phi})}{\partial\left(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\right)} &= \frac{c_t(\boldsymbol{\phi})-1}{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\frac{1}{\sqrt{1+4\left(D_{\nu+1}\left(\gamma\right)-1\right)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}, \\
\frac{\partial c_t(\boldsymbol{\phi})}{\partial \eta} &= \frac{c_t(\boldsymbol{\phi})-1}{\left[D_{\nu+1}\left(\gamma\right)-1\right]\sqrt{1+4\left(D_{\nu+1}\left(\gamma\right)-1\right)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}\frac{\partial D_{\nu+1}\left(\gamma\right)}{\partial \eta},
\end{aligned}
$$

and

$$
\frac{\partial c_t(\boldsymbol{\phi})}{\partial \psi} = \frac{c_t(\boldsymbol{\phi})-1}{\left[D_{\nu+1}\left(\gamma\right)-1\right]\sqrt{1+4\left(D_{\nu+1}\left(\gamma\right)-1\right)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}\frac{\partial D_{\nu+1}\left(\gamma\right)}{\partial \psi}.
$$

3.1 Asymmetric t limit

This case is obtained for $\psi = 1$ and $\eta > 0$. In terms of (1), ξ would be a Gamma variate with mean η^{-1} and variance $2\eta^{-1}$. Using the limiting expression (61), we can show that

$$\begin{aligned}
\lim_{\psi\to 1} \frac{R_\nu(\gamma)}{\gamma} &= \frac{\eta}{1-2\eta} \\
\lim_{\psi\to 1} D_{\nu+1}(\gamma) &= \frac{1-2\eta}{1-4\eta}
\end{aligned}$$

for $\eta < 0.25$. Introducing these results in (9), the asymmetric t log-likelihood can be expressed as

$$\begin{aligned}
l_{at}(\mathbf{y}_t|I_{t-1};\boldsymbol{\phi}) = &-\frac{N}{2}\log(2\pi) - \frac{1}{2}\log|\boldsymbol{\Sigma}_t(\boldsymbol{\theta})| - \frac{N}{2}\log\left(\frac{1-2\eta}{\eta}\right) \\
&-\frac{1}{2}\log(c_{at,t}(\boldsymbol{\phi})) + \left(1-\frac{1}{2\eta}\right)\log 2 - \log\Gamma\left(\frac{1}{2\eta}\right) \\
&+\mathbf{b}'\boldsymbol{\varepsilon}_t(\boldsymbol{\theta}) + c_{at,t}(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b} + \frac{1+N\eta}{4\eta}\log\left[\frac{1-2\eta}{\eta}c_{at,t}(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\right] \\
&-\frac{1+N\eta}{4\eta}\log(Q_{at,t}) + \log K_{\nu-.5N}\left[\sqrt{\frac{1-2\eta}{\eta}c_{at,t}(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}Q_{at,t}}\right]
\end{aligned} \tag{13}$$

where

$$\begin{aligned}
Q_{at,t} &= \lim_{\psi\to 1} Q_t = 1 + \frac{\eta}{1-2\eta}\left[\varsigma_t(\boldsymbol{\theta}) + 2c_{at,t}(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\varepsilon}_t(\boldsymbol{\theta}) + c^2_{at,t}(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\right] \\
&-\frac{\eta}{1-2\eta}\frac{c_{at,t}(\boldsymbol{\phi})-1}{c_{at,t}(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\left[\mathbf{b}'\boldsymbol{\varepsilon}_t(\boldsymbol{\theta}) + c_{at,t}(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\right]^2,
\end{aligned}$$

$$c_{at,t}(\boldsymbol{\phi}) = \lim_{\psi\to 1} c_t(\boldsymbol{\phi}) = \frac{-(1-4\eta) + \sqrt{(1-4\eta)^2 + 8\eta(1-4\eta)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}{4\eta\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}},$$

$$\begin{aligned}
\lim_{\psi\to 1} E(\xi_t|I_T;\boldsymbol{\phi}) &= \frac{1-2\eta}{\eta}\frac{\sqrt{c_{at,t}(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}{\sqrt{\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}\mathbf{p}_t + \frac{1-2\eta}{\eta}}} \\
&\times R_{\frac{N}{2}-\nu}\left[\sqrt{c_{at,t}(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\sqrt{\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}\mathbf{p}_t + \frac{1-2\eta}{\eta}}\right],
\end{aligned}$$

$$\begin{aligned}
\lim_{\psi\to 1} E\left(\frac{1}{\xi_t}\middle| I_T;\boldsymbol{\phi}\right) &= \frac{\eta}{1-2\eta}\frac{\sqrt{\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}\mathbf{p}_t + \frac{1-2\eta}{\eta}}}{\sqrt{c_{at,t}(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}} \\
&\times \frac{1}{R_{\frac{N}{2}-\nu-1}\left[\sqrt{c_{at,t}(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\sqrt{\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}\mathbf{p}_t + \frac{1-2\eta}{\eta}}\right]}
\end{aligned}$$

and

$$\lim_{\psi\to 1} E\left(\log \xi_t | \, I_T; \boldsymbol{\phi}\right) = \log\left(\sqrt{\frac{1-2\eta}{\eta} c_{at,t}(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\right) - \log\left(\sqrt{\frac{\eta}{1-2\eta}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}\mathbf{p}_t + 1}\right)$$
$$+\frac{\partial}{\partial x}\log K_x\left[\sqrt{c_{at,t}(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\sqrt{\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}\mathbf{p}_t + \frac{1-2\eta}{\eta}}\right]\Bigg|_{x=\frac{N}{2}-\nu}.$$

Hence, we can write

$$\lim_{\psi\to 1}\frac{\partial l(\mathbf{y}_t | \, I_{t-1};\boldsymbol{\phi})}{\partial \boldsymbol{\theta}'} = -\frac{1}{2}vec'[\boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\theta})]\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}'} - f_{at}(I_T,\boldsymbol{\phi})\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})\frac{\partial \mathbf{p}_t}{\partial \boldsymbol{\theta}'}$$
$$-\frac{1}{2}\frac{(c_{at,t}(\boldsymbol{\phi})-1)(1-4\eta)}{c_{at,t}(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\sqrt{(1-4\eta)^2+8\eta(1-4\eta)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}vec'(\mathbf{b}\mathbf{b}')\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}'} + \mathbf{b}'\frac{\partial \mathbf{p}_t}{\partial \boldsymbol{\theta}'}$$
$$+\frac{1}{2}f_{at}(I_T,\boldsymbol{\phi})[\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})\otimes \mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})]\frac{\partial vec[\boldsymbol{\Sigma}_t^{*}(\boldsymbol{\phi})]}{\partial \boldsymbol{\theta}'}$$
$$-\frac{1}{2}\frac{g_{at}(I_T,\boldsymbol{\phi})(1-4\eta)}{\sqrt{(1-4\eta)^2+8\eta(1-4\eta)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}vec'(\mathbf{b}\mathbf{b}')\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}'},$$

where

$$\begin{aligned} f_{at}(I_T,\boldsymbol{\phi}) &= \frac{\eta \lim_{\psi\to 1} E\left(\xi_t | I_T;\boldsymbol{\phi}\right)}{1-2\eta}, \\ g_{at}(I_T,\boldsymbol{\phi}) &= \frac{(1-2\eta)\lim_{\psi\to 1} E\left(\xi_t^{-1} | I_T;\boldsymbol{\phi}\right)}{\eta}, \end{aligned}$$

$$\frac{\partial vec[\boldsymbol{\Sigma}_t^{*}(\boldsymbol{\phi})]}{\partial \boldsymbol{\theta}'} = \frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}'} + \frac{c_{at,t}(\boldsymbol{\phi})-1}{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\left\{[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\mathbf{b}'\otimes I_N] + [I_N\otimes \boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\mathbf{b}']\right\}\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}'}$$
$$+\frac{c_{at,t}(\boldsymbol{\phi})-1}{[\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}]^2}\left\{\frac{1-4\eta}{\sqrt{(1-4\eta)^2+8\eta(1-4\eta)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}} - 1\right\}$$
$$\times vec\left[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\right]vec'(\mathbf{b}\mathbf{b}')\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}'},$$

$$\frac{\partial \mathbf{p}_t}{\partial \boldsymbol{\theta}'} = -\frac{\partial \boldsymbol{\mu}_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + c_t(\boldsymbol{\phi})\left[\mathbf{b}'\otimes I_N\right]\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}'}$$
$$+\frac{c_{at,t}(\boldsymbol{\phi})-1}{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\frac{1-4\eta}{\sqrt{(1-4\eta)^2+8\eta(1-4\eta)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}vec'(\mathbf{b}\mathbf{b}')\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}'},$$

$$\frac{\partial c_t(\boldsymbol{\phi})}{\partial\left(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\right)} = \frac{c_{at,t}(\boldsymbol{\phi})-1}{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\frac{1-4\eta}{\sqrt{(1-4\eta)^2+8\eta(1-4\eta)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}$$

and

$$\lim_{\psi\to 1}\frac{\partial l\left(\mathbf{y}_t|\, I_{t-1};\boldsymbol{\phi}\right)}{\partial \mathbf{b}'} = -\frac{\left(c_{at,t}(\boldsymbol{\phi})-1\right)(1-4\eta)}{c_t(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\sqrt{(1-4\eta)^2+8\eta(1-4\eta)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})$$
$$-f\left(I_T,\boldsymbol{\phi}\right)c_{at,t}(\boldsymbol{\phi})\mathbf{p}_t' + \boldsymbol{\varepsilon}_t'(\boldsymbol{\theta}) + f\left(I_T,\boldsymbol{\phi}\right)\frac{c_{at,t}(\boldsymbol{\phi})-1}{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\left(\mathbf{b}'\mathbf{p}_t\right)$$
$$\times\left\{\frac{\left[c_{at,t}(\boldsymbol{\phi})-1\right](1-4\eta)\left(\mathbf{b}'\mathbf{p}_t\right)}{c_{at,t}^2(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\sqrt{(1-4\eta)^2+8\eta(1-4\eta)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\right.$$
$$\left.+\frac{\mathbf{p}_t'}{c_{at,t}(\boldsymbol{\phi})} - \frac{1-4\eta}{\sqrt{(1-4\eta)^2+8\eta(1-4\eta)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\right\}$$
$$+\frac{(1-4\eta)\left[2-g\left(I_T,\boldsymbol{\phi}\right)\right]}{\sqrt{(1-4\eta)^2+8\eta(1-4\eta)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta}). \tag{14}$$

In the case of the derivative with respect to η, we can again use the reflection formula (59) and the limiting expression (61) of $K_\nu(\gamma)$, to show that

$$\begin{aligned}\frac{\partial \log K_\nu(\gamma)}{\partial \eta} &= \frac{\partial \log K_{-.5/\eta}(\gamma)}{\partial \eta} = \frac{\partial \log K_{.5/\eta}(\gamma)}{\partial \eta}\\ &= \frac{-1}{2\eta^2}\psi\left(\frac{1}{2\eta}\right) - \frac{\log 2}{2\eta^2} + \frac{\log\gamma}{2\eta^2}\end{aligned}$$

holds as ψ goes to one, where $\psi(\cdot)$ is the digamma function. Using similar arguments for $R_\nu(\gamma)$, we can show that

$$\lim_{\psi\to 1}\frac{\partial \log R_\nu(\gamma)}{\partial \eta} = \frac{1}{\eta(1-2\eta)}$$

$$\lim_{\psi\to 1}\frac{\partial D_{\nu+1}(\gamma)}{\partial \eta} = \frac{2}{(1-4\eta)^2}$$

and

$$\lim_{\psi\to 1}\frac{\partial c_t(\boldsymbol{\phi})}{\partial \eta} = \frac{c_{at,t}(\boldsymbol{\phi})-1}{\eta\sqrt{(1-4\eta)^2+8\eta(1-4\eta)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}$$

If we introduce these results in (11), we obtain

$$\frac{\partial l_{at}\left(\mathbf{y}_t|\, I_{t-1};\boldsymbol{\phi}\right)}{\partial \eta} = \frac{N}{2\eta(1-2\eta)} + \frac{1}{2\eta^2}\psi\left(\frac{1}{2\eta}\right) + \frac{\log 2}{2\eta^2}$$
$$+\left(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b} - \frac{1}{2c_{at,t}(\boldsymbol{\phi})}\right)\frac{c_{at,t}(\boldsymbol{\phi})-1}{\eta\sqrt{(1-4\eta)^2+8\eta(1-4\eta)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}$$
$$-\frac{\partial \log K_\nu(\gamma)}{\partial \eta} - \frac{1}{2\eta^2}E\left[\log\xi_t|Y_T;\boldsymbol{\phi}\right] - \frac{f\left(I_T,\boldsymbol{\phi}\right)}{2}\left\{\frac{1}{\eta(1-2\eta)}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})\mathbf{p}_t\right.$$
$$\left.+\frac{c_{at,t}(\boldsymbol{\phi})-1}{\eta\sqrt{(1-4\eta)^2+8\eta(1-4\eta)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}\left[\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b} - \frac{\left(\mathbf{b}'\boldsymbol{\varepsilon}_t\right)^2}{c_t^2(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\right]\right\}$$
$$-\frac{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}{2}g\left(I_T,\boldsymbol{\phi}\right)\frac{\left[c_{at,t}(\boldsymbol{\phi})-1\right](1-2\eta) - c_{at,t}(\boldsymbol{\phi})\sqrt{(1-4\eta)^2+8\eta(1-4\eta)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}{\eta(1-2\eta)\sqrt{(1-4\eta)^2+8\eta(1-4\eta)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}.$$

Using the same limiting expressions, we can show for small γ that

$$\frac{\partial \log K_\nu(\gamma)}{\partial \psi} = \frac{1}{2\eta\psi(1-\psi)} \tag{15}$$

$$\frac{\partial \log R_\nu(\gamma)}{\partial \psi} = \frac{-1}{\psi(1-\psi)} \tag{16}$$

$$\frac{\partial D_{\nu+1}(\gamma)}{\partial \psi} = 0$$

and

$$\frac{\partial c_t(\boldsymbol{\phi})}{\partial \psi} = 0.$$

If we introduce these results in (12), we finally obtain

$$\frac{\partial l_{at}(\mathbf{y}_t | I_{t-1}; \boldsymbol{\phi})}{\partial \psi} = g(I_T, \boldsymbol{\phi}) \frac{\eta(1-\psi)}{(1-2\eta)\psi^3},$$

which tends to zero as ψ tends to 1.

3.2 Student t limit

We now take the limit $\mathbf{b} \to \mathbf{0}$ on the asymmetric t log-likelihood. Using L'Hospital rule, it can be shown that, as $\mathbf{b} \to \mathbf{0}$, $c_{at,t}(\boldsymbol{\phi}) \to 1$ and

$$\lim_{\mathbf{b}\to\mathbf{0}} \frac{c_{at,t}(\boldsymbol{\phi}) - 1}{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}} = -2\eta. \tag{17}$$

Hence, $\boldsymbol{\Sigma}_t^*(\boldsymbol{\phi})$ tends to $\boldsymbol{\Sigma}_t(\theta)$ and $\mathbf{p}_t$ tends to $\boldsymbol{\varepsilon}_t(\boldsymbol{\theta})$ and $Q_{at,t}$ becomes

$$Q_{st,t} = \lim_{\mathbf{b}\to\mathbf{0}} Q_{at,t} = 1 + \frac{\eta}{1-2\eta}\varsigma_t(\boldsymbol{\theta}). \tag{18}$$

Since the argument of the Bessel function of the third kind in (13) tends to zero as $\mathbf{b} \to \mathbf{0}$, we can use (61) to show that

$$\begin{aligned}\log K_{\nu-.5N}\left[\sqrt{\frac{1-2\eta}{\eta}c_{at,t}(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}Q_{at,t}}\right] &\simeq -\log 2 + \log\Gamma\left(\frac{1}{2\eta} + \frac{N}{2}\right) \\ &-\frac{1+N\eta}{2\eta}\log\left[\frac{1}{2}\sqrt{\frac{1-2\eta}{\eta}c_{at,t}(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}Q_{at,t}}\right].\end{aligned} \tag{19}$$

If we introduce (18) and (19) in (13) and take the limit $\mathbf{b} \to \mathbf{0}$, we obtain

$$\begin{aligned} l_{st}(\mathbf{y}_t|I_{t-1};\boldsymbol{\phi}) &= -\frac{N}{2}\log(\pi) - \frac{1}{2}\log|\boldsymbol{\Sigma}_t(\boldsymbol{\theta})| - \frac{N}{2}\log\left(\frac{1-2\eta}{\eta}\right) \\ &+\log\Gamma\left(\frac{1+N\eta}{2\eta}\right) - \log\Gamma\left(\frac{1}{2\eta}\right) - \frac{1+N\eta}{4\eta}\log\left[1+\frac{\eta}{1-2\eta}\varsigma_t(\boldsymbol{\theta})\right],\end{aligned}$$

which is the log-likelihood of the symmetric Student t distribution. Hence, the score with respect to $\boldsymbol{\theta}$ and η will coincide with the values given in Fiorentini, Sentana, and Calzolari (2003). Once again, (65) can be used to show that

$$\lim_{\mathbf{b}\to\mathbf{0},\psi\to 1} E\left(\xi_t|I_T;\boldsymbol{\phi}\right) = \frac{(1-2\eta)(1+N\eta)}{\eta(1-2\eta)+\eta^2\varsigma_t(\boldsymbol{\theta})}$$

$$\lim_{\mathbf{b}\to\mathbf{0},\psi\to 1} E\left(\left.\frac{1}{\xi_t}\right| I_T;\boldsymbol{\phi}\right) = \frac{\eta(1-2\eta)+\eta^2\varsigma_t(\boldsymbol{\theta})}{(1-2\eta)(1+(N-2)\eta)}$$

As for the score with respect to $\mathbf{b}$, we can introduce (17) in (14) and obtain

$$\begin{aligned}\frac{\partial l_{st}\left(\mathbf{y}_t|\,I_{t-1};\boldsymbol{\phi}\right)}{\partial\mathbf{b}'} &= [1-f_{st}\left(I_T,\boldsymbol{\phi}\right)]\boldsymbol{\varepsilon}_t'(\boldsymbol{\theta}) \\ &= \frac{\eta[\varsigma_t(\boldsymbol{\theta})-(N+2)]}{1-2\eta+\eta\varsigma_t(\boldsymbol{\theta})}\boldsymbol{\varepsilon}_t'(\boldsymbol{\theta}),\end{aligned}$$

where

$$f_{st}\left(I_T,\boldsymbol{\phi}\right) = \lim_{\mathbf{b}\to\mathbf{0}} f_{at}\left(I_T,\boldsymbol{\phi}\right) = \frac{1+N\eta}{1-2\eta+\eta\varsigma_t(\boldsymbol{\theta})}.$$

3.3 Extremum test of Student t versus symmetric GH innovations

Let us consider the limit under Student t innovations of the second derivative of the log-likelihood with respect to ψ. First, we compute the score with respect to ψ under symmetric GH innovations ($\mathbf{b}=\mathbf{0}$):

$$\begin{aligned}\frac{\partial l_{sGH}\left(\mathbf{y}_t|\,I_{t-1};\boldsymbol{\phi}\right)}{\partial\psi} &= \frac{N}{2}\frac{\partial\log R_\nu\left(\gamma\right)}{\partial\psi} + \frac{1+N\eta}{2\eta\psi\left(1-\psi\right)} - \frac{\partial\log K_\nu\left(\gamma\right)}{\partial\psi} \\ &-\frac{f\left(I_T,\boldsymbol{\phi}\right)}{2}\left[\frac{\partial\log R_\nu\left(\gamma\right)}{\partial\psi}+\frac{1}{\psi\left(1-\psi\right)}\right]\varsigma_t(\boldsymbol{\theta}) + g\left(I_T,\boldsymbol{\phi}\right)\frac{R_\nu\left(\gamma\right)}{\psi^2}. \qquad (20)\end{aligned}$$

If we differentiate (20) with respect to ψ we obtain

$$\begin{aligned}\frac{\partial^2 l_{sGH}\left(\mathbf{y}_t|\,I_{t-1};\boldsymbol{\phi}\right)}{\partial\psi^2} &= \frac{N}{2}\frac{\partial^2\log R_\nu\left(\gamma\right)}{\partial\psi^2} - \frac{(1+N\eta)(1-2\psi)}{2\eta\psi^2\left(1-\psi\right)^2} - \frac{\partial^2\log K_\nu\left(\gamma\right)}{\partial\psi^2} \\ &-\frac{f\left(I_T,\boldsymbol{\phi}\right)}{2}\left[\frac{\partial^2\log R_\nu\left(\gamma\right)}{\partial\psi^2}-\frac{(1-2\psi)}{\psi^2\left(1-\psi\right)^2}\right]\varsigma_t(\boldsymbol{\theta}) \\ &-\frac{1}{2}\frac{\partial f\left(I_T,\boldsymbol{\phi}\right)}{\partial\psi}\left[\frac{\partial\log R_\nu\left(\gamma\right)}{\partial\psi}+\frac{1}{\psi\left(1-\psi\right)}\right]\varsigma_t(\boldsymbol{\theta}) \\ &+g\left(I_T,\boldsymbol{\phi}\right)\left[-2\frac{R_\nu\left(\gamma\right)}{\psi^3}+\frac{1}{\psi^2}\frac{\partial R_\nu\left(\gamma\right)}{\partial\psi}\right]+\frac{\partial g\left(I_T,\boldsymbol{\phi}\right)}{\partial\psi}\frac{R_\nu\left(\gamma\right)}{\psi^2}. \qquad (21)\end{aligned}$$

In order to compute the limit of (21) when ψ tends to one, we need to refine the limiting expressions (15) and (16). Using (64), we can expand $R_\nu\left(\gamma\right)$ around $\gamma=0$ as

$$R_\nu\left(\gamma\right) \simeq \frac{\eta}{1-2\eta}\gamma - \frac{\eta^3}{(1-2\eta)^2(1-4\eta)}\gamma^3 + O(\gamma^4).$$

This expansion is valid for $\nu = -.5/\eta$ when $\eta > 0$. In terms of ψ, we can expand around $\psi = 1$ with

$$R_{\nu}(\gamma) \simeq \frac{\eta(1-\psi)}{1-2\eta}\left[1-(\psi-1)+\frac{7\eta^2-6\eta+1}{(1-2\eta)(1-4\eta)}(\psi-1)^2+O[(\psi-1)^3]\right]. \tag{22}$$

We can also obtain,

$$\frac{\partial \log R_{\nu}(\gamma)}{\partial \psi} \simeq \frac{-1}{1-\psi}-1+O(\psi-1), \tag{23}$$

and

$$\frac{\partial^2 \log R_{\nu}(\gamma)}{\partial \psi^2} \simeq \frac{-1}{(1-\psi)^2}+\frac{6\eta^2-6\eta+1}{(1-2\eta)(1-4\eta)}+O(\psi-1).$$

We can introduce (22) in (60) and obtain

$$\begin{aligned}
\frac{\partial \log K_{\nu}(\gamma)}{\partial \psi} &\simeq \frac{1}{2\eta\psi(1-\psi)}+O(\psi-1),\\
\frac{\partial^2 \log K_{\nu}(\gamma)}{\partial \psi^2} &\simeq -\frac{\eta}{1-2\eta}-\frac{(1-2\psi)}{2\eta\psi^2(1-\psi)^2}+O(\psi-1).
\end{aligned}$$

If we introduce these expressions in (21), we can obtain

$$\lim_{\psi\to 1}\frac{\partial^2 l_{sGH}(\mathbf{y}_t|\, I_{t-1};\boldsymbol{\phi})}{\partial \psi^2} = \frac{\eta}{1-2\eta}-\frac{N\eta^2}{(1-2\eta)(1-4\eta)}$$
$$-\frac{f(I_T,\boldsymbol{\phi})}{2}\left[\frac{-2\eta^2}{(1-2\eta)(1-4\eta)}\right]\varsigma_t(\boldsymbol{\theta})+g(I_T,\boldsymbol{\phi})\left[\frac{-\eta}{1-2\eta}\right].$$

Thus, once we introduce the values of $f(I_T,\boldsymbol{\phi})$ and $g(I_T,\boldsymbol{\phi})$ under the Student t distribution, we finally obtain

$$\lim_{\psi\to 1}\frac{\partial^2 l_{sGH}(\mathbf{y}_t|\, I_{t-1};\boldsymbol{\phi})}{\partial \psi^2} = \frac{\eta^2}{(1-2\eta)(1-4\eta)}\frac{\varsigma_t(\boldsymbol{\theta})-N(1-2\eta)}{1-2\eta+\eta\varsigma_t(\boldsymbol{\theta})}+\frac{\eta^2[N-\varsigma_t(\boldsymbol{\theta})]}{(1-2\eta)(1+(N-2)\eta)}.$$

3.4 Gaussian limits

3.4.1 First case: $\eta \to 0^+$

To obtain this limit, we rely on (63). Since $\nu = -.5/\eta$ is negative in this limit, we need to exploit the reflection property (59) when we apply (63). However, this is not necessary when computing the ratios of Bessel functions in $E(\xi_t|I_T;\boldsymbol{\phi})$, $E\left(\xi_t^{-1}|I_T;\boldsymbol{\phi}\right)$ and $E(\log \xi_t|I_T;\boldsymbol{\phi})$, because $.5N-\nu$ goes to $+\infty$ as $\eta \to 0^+$.

In particular, after tedious but otherwise straightforward algebra, we can write for small $\eta > 0$

$$R_{\nu}(\gamma) = \gamma\eta + 2\gamma\eta^2 + (4\gamma-\gamma^3)\eta^3 + O(\eta^4), \tag{24}$$

$$D_{\nu+1}(\gamma) = 1 + 2\eta + 8\eta^2 + O(\eta^3), \tag{25}$$

and

$$c_t(\boldsymbol{\phi}) = 1 - 2\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\eta + [8(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b})^2 - 8\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}]\eta^2 + O(\eta^3). \tag{26}$$

Also using (63), together with (24) and (26), it is possible to show that the following two Taylor expansions hold:

$$\begin{aligned} f(I_T, \boldsymbol{\phi}) &= 1 + (N - \varsigma_t(\boldsymbol{\theta}) - 2\mathbf{b}'\boldsymbol{\varepsilon}_t + 2)\eta + O(\eta^2), & (27) \\ g(I_T, \boldsymbol{\phi}) &= 1 + (2\mathbf{b}'\boldsymbol{\varepsilon}_t + \varsigma_t(\boldsymbol{\theta}) - N)\eta + O(\eta^2). & (28) \end{aligned}$$

If we introduce these results in (10) and take the limit $\eta \to 0^+$, we can show that the resulting limit is the score of the Gaussian log-likelihood. Similarly it is possible to show that $\lim_{\eta \to 0^+}[\partial l(\mathbf{y}_t | I_{t-1}; \boldsymbol{\phi})/\partial \mathbf{b}] = \mathbf{0}$. In addition, we can also obtain from (63) that

$$\lim_{\eta \to 0^+} \frac{\partial \log R_\nu(\gamma)}{\partial \psi} = \frac{-1}{\psi(1-\psi)},$$

and

$$\lim_{\eta \to 0^+} \frac{\partial \log K_\nu(\gamma)}{\partial \psi} = \frac{1}{2\eta\psi(1-\psi)},$$

which imply that $\lim_{\eta \to 0^+}[\partial l(\mathbf{y}_t | I_{t-1}; \boldsymbol{\phi})/\partial \psi] = 0$.

As for the score with respect to η, once again we can obtain from (63) that

$$\frac{\partial \log R_\nu(\gamma)}{\partial \eta} = \frac{1}{\eta} + 2 + (4 - 2\gamma^2)\eta + (8 - 18\gamma^2)\eta^2 + O(\eta^3) \tag{29}$$

and

$$\frac{\partial \log K_\nu(\gamma)}{\partial \eta} = \frac{\log(\eta\gamma)}{2\eta^2} + \frac{1}{2\eta} - \frac{\gamma^2}{2} + \frac{1}{6} + O(\eta). \tag{30}$$

Then, if we rewrite (11) as

$$\frac{\partial l(\mathbf{y}_t | I_{t-1}; \boldsymbol{\phi})}{\partial \eta} = \varkappa_1 - \frac{1}{2\eta^2} E[\log \xi_t | Y_T; \boldsymbol{\phi}] - \frac{f(I_T, \boldsymbol{\phi})}{2}\varkappa_2 - \frac{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}{2} g(I_T, \boldsymbol{\phi})\varkappa_3, \tag{31}$$

where

$$\begin{aligned} \varkappa_1 &= \frac{N}{2}\frac{\partial \log R_\nu(\gamma)}{\partial \eta} + \left(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b} - \frac{1}{2c_t(\boldsymbol{\phi})}\right)\frac{\partial c_t(\boldsymbol{\phi})}{\partial \eta} \\ &\quad + \frac{\log(\gamma)}{2\eta^2} - \frac{\partial \log K_\nu(\gamma)}{\partial \eta} \end{aligned}$$

$$\varkappa_2 = \frac{\partial \log R_\nu(\gamma)}{\partial \eta}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})\mathbf{p}_t + \frac{\partial c_t(\boldsymbol{\phi})}{\partial \eta}\left[\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b} - \frac{(\mathbf{b}'\boldsymbol{\varepsilon}_t)^2}{c_t^2(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\right]$$

and

$$\varkappa_3 = \frac{\partial c_t(\boldsymbol{\phi})}{\partial \eta} - c_t(\boldsymbol{\phi})\frac{\partial \log R_\nu(\gamma)}{\partial \eta},$$

we can use (24), (26), (29) and (30) to obtain the following expansions:

$$\varkappa_1 = -\frac{\log \eta}{2\eta^2} + \frac{N-1}{2\eta} + N + \frac{\gamma^2}{2} - \frac{1}{6} - 2(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b})^2 + \mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b} + O(\eta), \tag{32}$$

$$\begin{aligned}\varkappa_2 &= \frac{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b} + 2\mathbf{b}'\boldsymbol{\varepsilon}_t + \varsigma_t(\boldsymbol{\theta})}{\eta} + 2\varsigma_t(\boldsymbol{\theta}) + 4\mathbf{b}'\boldsymbol{\varepsilon}_t + 4(\mathbf{b}'\boldsymbol{\varepsilon}_t)^2 \\ &\quad +2\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b} - 4(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b})^2 + O(\eta),\end{aligned} \tag{33}$$

$$\varkappa_3 = -\frac{1}{\eta} - 2 + O(\eta), \tag{34}$$

$$\begin{aligned}&E\left[\log \xi_t | Y_T; \boldsymbol{\phi}\right] = -\log \eta + [N - \varsigma_t(\boldsymbol{\theta}) - 2\mathbf{b}'\boldsymbol{\varepsilon}_t - 1]\eta \\ &+ \left[\begin{array}{c} N - 2\varsigma_t(\boldsymbol{\theta}) - 4\mathbf{b}'\boldsymbol{\varepsilon}_t + 4(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b})(\mathbf{b}'\boldsymbol{\varepsilon}_t) \\ +2(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b})\varsigma_t(\boldsymbol{\theta}) + 2(\mathbf{b}'\boldsymbol{\varepsilon}_t)\varsigma_t(\boldsymbol{\theta}) \\ -2(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b})N + \gamma^2 + \frac{1}{2}\varsigma_t^2(\boldsymbol{\theta}) - \frac{1}{2}N^2 - \frac{1}{3} \end{array}\right]\eta^2 + O(\eta^3).\end{aligned} \tag{35}$$

If we introduce (27), (28), (32), (33), (34) and (35) in (31) we can check that the elements with a pole at $\eta = 0$ cancel out, while the remaining terms yield

$$\begin{aligned}\lim_{\eta \to 0^+} \frac{\partial l\left(\mathbf{y}_t | I_{t-1}; \boldsymbol{\phi}\right)}{\partial \eta} &= \left[\frac{1}{4}\varsigma_t^2(\boldsymbol{\theta}) - \frac{N+2}{2}\varsigma_t(\boldsymbol{\theta}) + \frac{N(N+2)}{4}\right] \\ &\quad +\mathbf{b}'\left\{\boldsymbol{\varepsilon}_t(\boldsymbol{\theta})\left[\varsigma_t(\boldsymbol{\theta}) - (N+2)\right]\right\}.\end{aligned}$$

3.4.2 Second case: $\eta \to 0^-$

To obtain this limit, we also rely on (63). Notice that $\nu = -.5/\eta$ is positive in this case. Hence, we can directly apply (63). However, we need to consider (59) when computing the ratios of Bessel functions in $E(\xi_t | I_T; \boldsymbol{\phi})$, $E\left(\xi_t^{-1} | I_T; \boldsymbol{\phi}\right)$ and $E(\log \xi_t | I_T; \boldsymbol{\phi})$, because now $.5N - \nu$ goes to $-\infty$ as $\eta \to 0^-$.

In particular, we can write for small $\eta < 0$

$$R_\nu(\gamma) = \frac{1 + \gamma^2\eta^2 - 2\gamma^2\eta^3 + O(\eta^4)}{-\gamma\eta}, \tag{36}$$

$$D_{\nu+1}(\gamma) = 1 - 2\eta + 4\gamma^2\eta^3 + O(\eta^4), \tag{37}$$

$$c_t(\boldsymbol{\phi}) = 1 + 2\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\eta + 8(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b})^2\eta^2 + O(\eta^3). \tag{38}$$

Also using (63), together with (36) and (38), it is possible to show that the following two Taylor expansions hold:

$$f(I_T, \boldsymbol{\phi}) = 1 + (-N + \varsigma_t(\boldsymbol{\theta}) + 2\mathbf{b}'\boldsymbol{\varepsilon}_t - 2)\eta + O(\eta^2), \tag{39}$$

$$g(I_T, \boldsymbol{\phi}) = 1 + (-2\mathbf{b}'\boldsymbol{\varepsilon}_t - \varsigma_t(\boldsymbol{\theta}) + N)\eta + O(\eta^2). \tag{40}$$

If we introduce these results in (10) and take the limit $\eta \to 0^-$, we can show that the resulting limit is the score of the Gaussian log-likelihood. Similarly it is possible to show that $\lim_{\eta\to 0^-}[\partial l(\mathbf{y}_t|\, I_{t-1};\boldsymbol{\phi})/\partial \mathbf{b}] = \mathbf{0}$. In addition, we can also obtain from (63) that

$$\lim_{\eta\to 0^-}\frac{\partial \log R_\nu(\gamma)}{\partial \psi} = \frac{1}{\psi(1-\psi)},$$

and

$$\lim_{\eta\to 0^-}\frac{\partial \log K_\nu(\gamma)}{\partial \psi} = \frac{-1}{2\eta\psi(1-\psi)},$$

which imply that $\lim_{\eta\to 0^-}[\partial l(\mathbf{y}_t|\, I_{t-1};\boldsymbol{\phi})/\partial \psi] = 0$.

As for the score with respect to η, once again we can obtain from (63) that

$$\frac{\partial \log R_\nu(\gamma)}{\partial \eta} = \frac{-1}{\eta} + 2\gamma^2\eta - 6\gamma^2\eta^2 + O(\eta^3) \tag{41}$$

and

$$\frac{\partial \log K_\nu(\gamma)}{\partial \eta} = \frac{-\log(-\eta\gamma)}{2\eta^2} + \frac{1}{2\eta} + \frac{\gamma^2}{2} - \frac{1}{6} + O(\eta). \tag{42}$$

Then, if we consider the components of the score as written in (31), we can use (36), (38), (41) and (42) to obtain the following expansions:

$$\begin{aligned}\varkappa_1 &= \frac{\log(-\eta\gamma^2)}{2\eta^2} - \frac{(N+1)}{2\eta} + 2(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b})^2 \\ &\quad -\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b} - \frac{\gamma^2}{2} + \frac{1}{6} + O(\eta),\end{aligned} \tag{43}$$

$$\varkappa_2 = \frac{-\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b} - 2\mathbf{b}'\boldsymbol{\varepsilon}_t - \varsigma_t(\boldsymbol{\theta})}{\eta} + O(\eta), \tag{44}$$

$$\varkappa_3 = \frac{1}{\eta} + 4\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b} + O(\eta), \tag{45}$$

$$\begin{aligned} E\left[\log \xi_t | Y_T;\boldsymbol{\phi}\right] &= -\log(-\eta\gamma^2) + [2\mathbf{b}'\boldsymbol{\varepsilon}_t + \varsigma_t(\boldsymbol{\theta}) - N - 1]\eta \\ + &\left[\begin{array}{c} N - 8\mathbf{b}'\boldsymbol{\varepsilon}_t - 4\varsigma_t(\boldsymbol{\theta}) - 4\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b} + 4(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b})(\mathbf{b}'\boldsymbol{\varepsilon}_t) \\ +2(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b})\varsigma_t(\boldsymbol{\theta}) + 6(\mathbf{b}'\boldsymbol{\varepsilon}_t)\varsigma_t(\boldsymbol{\theta}) \\ -2N\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b} - 4N(\mathbf{b}'\boldsymbol{\varepsilon}_t) - 2N\varsigma_t(\boldsymbol{\theta}) + 4(\mathbf{b}'\boldsymbol{\varepsilon}_t)^2 \\ -\gamma^2 + (3/2)\varsigma_t^2(\boldsymbol{\theta}) + (N^2/2) + (1/3) \end{array}\right]\eta^2 + O(\eta^3). \end{aligned} \tag{46}$$

If we introduce (39), (40), (43), (44), (45) and (46) in (31) we can check that the elements with a pole at $\eta = 0$ cancel out, while the remaining terms yield

$$\lim_{\eta\to 0^-}\frac{\partial l(\mathbf{y}_t|\, I_{t-1};\boldsymbol{\phi})}{\partial \eta} = -\lim_{\eta\to 0^+}\frac{\partial l(\mathbf{y}_t|\, I_{t-1};\boldsymbol{\phi})}{\partial \eta}.$$

3.4.3 Third case: $\psi \to 0$

To obtain this limit, we rely on (62), since in this case the order of the Bessel function remains fixed while its argument goes to infinity.

In particular, we can write for small ψ

$$R_\nu(\gamma) = 1 + \left(\frac{1}{2} + \nu\right)\psi + \left(\frac{3}{8} + \nu + \frac{\nu^2}{2}\right)\psi^2 + O(\psi^3), \tag{47}$$

$$D_{\nu+1}(\gamma) = 1 + \psi + \psi^2 - \left(\frac{\nu^2}{2} + 2\nu + \frac{7}{8}\right)\psi^3 + O(\psi^4), \tag{48}$$

$$c_t(\boldsymbol{\phi}) = 1 - \mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\psi + [2(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b})^2 - \mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}]\psi^2 + O(\psi^3). \tag{49}$$

Also using (62), together with (47) and (49), it is possible to show that the following two Taylor expansions hold:

$$\begin{aligned} f(I_T, \boldsymbol{\phi}) &= 1 + \left(\frac{N}{2} - \frac{\varsigma_t(\boldsymbol{\theta})}{2} - \mathbf{b}'\boldsymbol{\varepsilon}_t + 1\right)\psi + O(\psi^2), & (50)\\ g(I_T, \boldsymbol{\phi}) &= 1 + \left(\mathbf{b}'\boldsymbol{\varepsilon}_t + \frac{\varsigma_t(\boldsymbol{\theta})}{2} - \frac{N}{2}\right)\psi \\ &\quad + \begin{pmatrix} \mathbf{b}'\boldsymbol{\varepsilon}_t + \frac{\varsigma_t(\boldsymbol{\theta})}{2} - \frac{N}{4} - (\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b})(\mathbf{b}'\boldsymbol{\varepsilon}_t) \\ -\frac{(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b})\varsigma_t(\boldsymbol{\theta})}{2} - \frac{(\mathbf{b}'\boldsymbol{\varepsilon}_t)\varsigma_t(\boldsymbol{\theta})}{2} + \frac{N(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b})}{2} \\ -\frac{\varsigma_t^2(\boldsymbol{\theta})}{8} + \frac{N^2}{8} \end{pmatrix}\psi^2 + O(\psi^3). & (51) \end{aligned}$$

If we introduce these results in (10) and take the limit $\psi \to 0$, we can show that the resulting limit is the score of the Gaussian log-likelihood. Similarly it is possible to show that $\lim_{\psi\to 0}[\partial l(\mathbf{y}_t | I_{t-1}; \boldsymbol{\phi})/\partial \mathbf{b}] = \mathbf{0}$. In addition, we can also obtain from (62) that

$$\begin{aligned} \lim_{\psi\to 0} \frac{\partial \log R_\nu(\gamma)}{\partial \eta} &= 0, \\ \lim_{\psi\to 0} \frac{\partial \log K_\nu(\gamma)}{\partial \eta} &= 0, \end{aligned}$$

and

$$E[\log \xi_t | Y_T; \boldsymbol{\phi}] = -\log\psi + O(\psi),$$

which imply that $\lim_{\psi\to 0}[\partial l(\mathbf{y}_t | I_{t-1}; \boldsymbol{\phi})/\partial \eta] = 0$.

As for the score with respect to ψ, we can obtain from (62) that

$$\frac{\partial \log R_\nu(\gamma)}{\partial \psi} = \frac{1}{2} + \nu + \left(\frac{1}{2} + \nu\right)\psi + O(\psi^2) \tag{52}$$

and

$$\frac{\partial \log K_\nu(\gamma)}{\partial \psi} = \frac{1}{\psi^2} + \frac{1}{2\psi} + \frac{3}{8} + \frac{\nu^2}{2} + O(\psi). \tag{53}$$

Then, if we rewrite (12) as

$$\begin{aligned}\frac{\partial l\left(\mathbf{y}_t\right|I_{t-1};\boldsymbol{\phi})}{\partial\psi} &= \varrho_1-\frac{f\left(I_T,\boldsymbol{\phi}\right)}{2}\varrho_2-\frac{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}{2}g\left(I_T,\boldsymbol{\phi}\right)\varrho_3\\ &\quad +g\left(I_T,\boldsymbol{\phi}\right)\frac{R_\nu\left(\gamma\right)}{\psi^2},\end{aligned}\tag{54}$$

where

$$\begin{aligned}\varrho_1 &= \frac{N}{2}\frac{\partial\log R_\nu\left(\gamma\right)}{\partial\psi}+\left(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}-\frac{1}{2c_t(\boldsymbol{\phi})}\right)\frac{\partial c_t(\boldsymbol{\phi})}{\partial\psi}\\ &\quad +\frac{1}{2\eta\psi\left(1-\psi\right)}-\frac{\partial\log K_\nu\left(\gamma\right)}{\partial\psi}+\frac{N}{2\psi\left(1-\psi\right)},\end{aligned}$$

$$\varrho_2=\left[\frac{\partial\log R_\nu\left(\gamma\right)}{\partial\psi}+\frac{1}{\psi\left(1-\psi\right)}\right]\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})\mathbf{p}_t+\frac{\partial c_t(\boldsymbol{\phi})}{\partial\psi}\left[\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}-\frac{\left(\mathbf{b}'\boldsymbol{\varepsilon}_t\right)^2}{c_t^2(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\right],$$

and

$$\varrho_3=-\frac{c_t(\boldsymbol{\phi})}{\psi\left(1-\psi\right)}+\frac{\partial c_t(\boldsymbol{\phi})}{\partial\psi}-c_t(\boldsymbol{\phi})\frac{\partial\log R_\nu\left(\gamma\right)}{\partial\psi},$$

we can use (47), (49), (52) and (53) to obtain the following expansions:

$$\begin{aligned}\varrho_1 &= \frac{-1}{\psi^2}+\frac{N-2\nu-1}{2\psi}+\frac{3N}{4}+\left(\frac{N}{2}-1\right)\nu\\ &\quad -(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b})^2+\frac{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}{2}-\frac{3}{8}-\frac{\nu^2}{2}+O(\psi),\end{aligned}\tag{55}$$

$$\begin{aligned}\varrho_2 &= \frac{\varsigma_t(\boldsymbol{\theta})+2(\mathbf{b}'\boldsymbol{\varepsilon}_t)+(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b})}{\psi}+\frac{3(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b})}{2}\\ &\quad +3(\mathbf{b}'\boldsymbol{\varepsilon}_t)+\frac{3\varsigma_t(\boldsymbol{\theta})}{2}+(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b})\nu+2(\mathbf{b}'\boldsymbol{\varepsilon}_t)\nu\\ &\quad +\varsigma_t(\boldsymbol{\theta})\nu-2(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b})^2+2(\mathbf{b}'\boldsymbol{\varepsilon}_t)^2+O(\psi),\end{aligned}\tag{56}$$

$$\varrho_3=\frac{-1}{\psi}-\frac{3}{2}-\nu+O(\psi).\tag{57}$$

If we introduce (50), (51) (55), (56) and (57), in (54), we can check that the elements with a pole at $\psi=0$ cancel out, while the remaining terms yield

$$\lim_{\psi\to 0}\frac{\partial l\left(\mathbf{y}_t\right|I_{t-1};\boldsymbol{\phi})}{\partial\psi}=\frac{1}{2}\lim_{\eta\to 0^+}\frac{\partial l\left(\mathbf{y}_t\right|I_{t-1};\boldsymbol{\phi})}{\partial\eta}.$$

4 Modified Bessel function of the third kind

The modified Bessel function of the third kind with order ν, which we denote as $K_\nu\left(\cdot\right)$, is closely related to the modified Bessel function of the first kind $I_\nu\left(\cdot\right)$, as

$$K_\nu\left(x\right)=\frac{\pi}{2}\frac{I_{-\nu}\left(x\right)-I_\nu\left(x\right)}{\sin\left(\pi\nu\right)}.\tag{58}$$

Some basic properties of $K_\nu(\cdot)$, taken from Abramowitz and Stegun (1965), are

$$\begin{aligned} K_\nu(x) &= K_{-\nu}(x), \qquad (59)\\ K_{\nu+1}(x) &= 2\nu x^{-1} K_\nu(x) + K_{\nu-1}(x) \end{aligned}$$

and

$$\begin{aligned} \frac{\partial K_\nu(x)}{\partial x} &= -\nu x^{-1} K_\nu(x) - K_{\nu-1}(x), \\ \frac{\partial K_\nu(x)}{\partial x} &= -K_{\nu+1}(x) + \nu x^{-1} K_\nu(x). \end{aligned}$$

Hence,

$$\frac{\partial \log K_\nu(x)}{\partial x} = -R_\nu(x) + \nu x^{-1}, \qquad (60)$$

where $R_\nu(x) = K_{\nu+1}(x)/K_\nu(x)$. For small values of the argument x, and ν fixed, it holds that

$$K_\nu(x) \simeq \frac{1}{2}\Gamma(\nu)\left(\frac{1}{2}x\right)^{-\nu}. \qquad (61)$$

Similarly, for ν fixed, $|x|$ large and $m = 4\nu^2$, the following asymptotic expansion is valid

$$K_\nu(x) = \sqrt{\frac{\pi}{2x}} e^{-x} \left\{ 1 + \frac{m\text{-}1}{8x} + \frac{(m\text{-}1)(m\text{-}9)}{2!\,(8x)^2} + \frac{(m\text{-}1)(m\text{-}9)(m\text{-}25)}{3!\,(8x)^3} + \cdots \right\}. \qquad (62)$$

Finally, for large values of x and ν we have that

$$\begin{aligned} K_\nu(x) &\simeq \sqrt{\frac{\pi}{2\nu}} \frac{\exp(-\nu l^{-1})}{l^{-1/2}} \left[\frac{(x/\nu)}{1+l^{-1}}\right]^{-\nu} \\ &\times \left[\begin{array}{c} 1\text{-}\frac{3l\text{-}5l^3}{24\nu} + \frac{81l^2\text{-}462l^4+385l^6}{1152\nu^2} \\ -\frac{30375l^3-369603l^5+765765l^7-425425l^9}{414720\nu^3} + \cdots \end{array} \right], \qquad (63) \end{aligned}$$

where $\nu > 0$ and $l = \left[1 + (x/\nu)^2\right]^{-\frac{1}{2}}$. Both (62) and (63) are convergent infinite series. The rule followed by higher order terms can be obtained in Abramowitz and Stegun (1965, page 378).

Although the existing literature does not discuss how to obtain numerically reliable derivatives of $K_\nu(x)$ with respect to its order, our experience suggests the following conclusions:

- For $\nu \leq 10$ and $|x| > 12$, the derivative of (62) with respect to ν gives a better approximation than the direct derivative of $K_\nu(x)$, which is in fact very unstable.
- For $\nu > 10$, the derivative of (63) with respect to ν works better than the direct derivative of $K_\nu(x)$.

• Otherwise, the direct derivative of the original function works well.

We can express such a derivative as a function of $I_\nu(x)$ by using (58) as:

$$\frac{\partial K_\nu(x)}{\partial \nu} = \frac{\pi}{2\sin(\nu\pi)}\left[\frac{\partial I_{-\nu}(x)}{\partial \nu} - \frac{\partial I_\nu(x)}{\partial \nu}\right] - \pi\cot(\nu\pi) K_\nu(x)$$

However, this formula becomes numerically unstable when ν is near any non-negative integer $n = 0, 1, 2, \cdots$ due to the sine that appears in the denominator. In our experience, it is much better to use the following Taylor expansion for small $|\nu - n|$:

$$\frac{\partial K_\nu(x)}{\partial \nu} = \left.\frac{\partial K_\nu(x)}{\partial \nu}\right|_{\nu=n} + \left.\frac{\partial^2 K_\nu(x)}{\partial \nu^2}\right|_{\nu=n}(\nu - n)$$
$$+ \left.\frac{\partial^3 K_\nu(x)}{\partial \nu^3}\right|_{\nu=n}(\nu - n)^2 + \left.\frac{\partial^4 K_\nu(x)}{\partial \nu^4}\right|_{\nu=n}(\nu - n)^3,$$

where for *integer* ν:

$$\frac{\partial K_\nu(x)}{\partial \nu} = \frac{1}{4\cos(\pi n)}\left[\frac{\partial^2 I_{-\nu}(x)}{\partial \nu^2} - \frac{\partial^2 I_\nu(x)}{\partial \nu^2}\right] + \pi^2\left[I_{-\nu}(x) - I_\nu(x)\right],$$

$$\frac{\partial^2 K_\nu(x)}{\partial \nu^2} = \frac{1}{6\cos(\pi n)}\left[\frac{\partial^3 I_{-\nu}(x)}{\partial \nu^3} - \frac{\partial^3 I_\nu(x)}{\partial \nu^3}\right] + \frac{\pi^2}{3\cos(\pi n)}\left[\frac{\partial I_{-\nu}(x)}{\partial \nu} - \frac{\partial I_\nu(x)}{\partial \nu}\right] - \frac{\pi^2}{3}K_n(x),$$

$$\frac{\partial^3 K_\nu(x)}{\partial \nu^3} = \frac{1}{8\cos(\pi n)}\left\{\left[\frac{\partial^4 I_{-\nu}(x)}{\partial \nu^4} - \frac{\partial^4 I_\nu(x)}{\partial \nu^4}\right]\right.$$
$$\left. -4\pi^2\left[\frac{\partial^2 I_{-\nu}(x)}{\partial \nu^2} - \frac{\partial^2 I_\nu(x)}{\partial \nu^2}\right] - 12\pi^4\left[I_{-\nu}(x) - I_\nu(x)\right]\right\} + 3\pi^2\frac{\partial K_n(x)}{\partial \nu},$$

and

$$\frac{\partial^4}{\partial \nu^4}K_\nu(x) = \frac{1}{8\cos(\pi n)}\left\{\frac{3}{2}\left[\frac{\partial^5 I_{-\nu}(x)}{\partial \nu^5} - \frac{\partial^5 I_\nu(x)}{\partial \nu^5}\right]\right.$$
$$\left. -10\pi^2\left[\frac{\partial^3 I_{-\nu}(x)}{\partial \nu^3} - \frac{\partial^3 I_\nu(x)}{\partial \nu^3}\right] -4\pi^4\left[\frac{\partial I_{-\nu}(x)}{\partial \nu} - \frac{\partial I_\nu(x)}{\partial \nu}\right]\right\} + 6\pi^2\frac{\partial^2 K_n(x)}{\partial \nu^2} - \pi^4 K_n(x).$$

Let $\psi^{(i)}(\cdot)$ denote the polygamma function (see Abramowitz and Stegun, 1965). The first five derivatives of $I_\nu(x)$ for any real ν are as follows:

$$\frac{\partial I_\nu(x)}{\partial \nu} = I_\nu(x)\log\left(\frac{x}{2}\right) - \left(\frac{x}{2}\right)^\nu \sum_{k=0}^{\infty}\frac{Q_1(\nu + k + 1)}{k!}\left(\frac{1}{4}x^2\right)^k,$$

where

$$Q_1(z) = \begin{cases} \psi(z)/\Gamma(z) & \text{if } z > 0 \\ \pi^{-1}\Gamma(1 - z)\left[\psi(1 - z)\sin(\pi z) - \pi\cos(\pi z)\right] & \text{if } z \leq 0 \end{cases}$$

$$\frac{\partial^2 I_\nu(x)}{\partial \nu^2} = 2\log\left(\frac{x}{2}\right)\frac{\partial I_\nu(x)}{\partial \nu} - I_\nu(x)\left[\log\left(\frac{x}{2}\right)\right]^2 - \left(\frac{x}{2}\right)^\nu \sum_{k=0}^{\infty}\frac{Q_2(\nu + k + 1)}{k!}\left(\frac{1}{4}x^2\right)^k,$$

where

$$Q_2(z) = \begin{cases} \left[\psi'(z) - \psi^2(z)\right]/\Gamma(z) \;\text{ if } z > 0 \\ \pi^{-1}\Gamma(1-z)\left[\pi^2 - \psi'(1-z) - \left[\psi(1-z)\right]^2\right]\sin(\pi z) \\ +2\Gamma(1-z)\psi(1-z)\cos(\pi z) \;\text{ if } z \leq 0 \end{cases}$$

$$\begin{aligned} \frac{\partial^3 I_\nu(x)}{\partial \nu^3} &= 3\log\left(\frac{x}{2}\right)\frac{\partial^2 I_\nu(x)}{\partial \nu^2} - 3\left[\log\left(\frac{x}{2}\right)\right]^2\frac{\partial I_\nu(x)}{\partial \nu} + \left[\log\left(\frac{x}{2}\right)\right]^3 I_\nu(x) \\ &\quad - \left(\frac{x}{2}\right)^\nu \sum_{k=0}^{\infty}\frac{Q_3(\nu+k+1)}{k!}\left(\frac{1}{4}x^2\right)^k, \end{aligned}$$

where

$$Q_3(z) = \begin{cases} \left[\psi^3(z) - 3\psi(z)\psi'(z) + \psi''(z)\right]/\Gamma(z) \;\text{ if } z > 0 \\ \pi^{-1}\Gamma(1-z)\left\{\psi^3(1-z) - 3\psi(1-z)\left[\pi^2 - \psi'(1-z)\right] + \psi''(1-z)\right\}\sin(\pi z) \\ +\Gamma(1-z)\left\{\pi^2 - 3\left[\psi^2(1-z) + \psi'(1-z)\right]\right\}\cos(\pi z) \;\text{ if } z \leq 0 \end{cases}$$

$$\begin{aligned} \frac{\partial^4 I_\nu(x)}{\partial \nu^4} &= 4\log\left(\frac{x}{2}\right)\frac{\partial^3 I_\nu(x)}{\partial \nu^3} - 6\left[\log\left(\frac{x}{2}\right)\right]^2\frac{\partial^2 I_\nu(x)}{\partial \nu^2} + 4\left[\log\left(\frac{x}{2}\right)\right]^3\frac{\partial I_\nu(x)}{\partial \nu} \\ &\quad - \left[\log\left(\frac{x}{2}\right)\right]^4 I_\nu(x) - \left(\frac{x}{2}\right)^\nu \sum_{k=0}^{\infty}\frac{Q_4(\nu+k+1)}{k!}\left(\frac{1}{4}x^2\right)^k, \end{aligned}$$

where

$$Q_4(z) = \begin{cases} \left[-\psi^4(z) + 6\psi^2(z)\psi'(z) - 4\psi(z)\psi''(z) - 3\left[\psi'(z)\right]^2 + \psi'''(z)\right]/\Gamma(z) \;\text{ if } z > 0 \\ \pi^{-1}\Gamma(1-z)\left\{-\psi^4(1-z) + 6\pi^2\psi^2(1-z) - 6\psi^2(1-z)\psi'(1-z)\right. \\ -4\psi(1-z)\psi''(1-z) - 3\left[\psi'(1-z)\right]^2 + 6\pi^2\psi'(1-z) \\ \left.-\psi'''(1-z) - \pi^4\right\}\sin(\pi z) + \Gamma(1-z)\,4\psi^3(1-z) - 4\pi^2\psi(1-z) \\ +12\psi(1-z)\psi'(1-z) + 4\psi''(1-z)\cos(\pi z) \;\text{ if } z \leq 0 \end{cases}$$

and finally,

$$\begin{aligned} \frac{\partial^5 I_\nu(x)}{\partial \nu^5} &= 5\log\left(\frac{x}{2}\right)\frac{\partial^4 I_\nu(x)}{\partial \nu^4} - 10\left[\log\left(\frac{x}{2}\right)\right]^2\frac{\partial^3 I_\nu(x)}{\partial \nu^3} + 10\left[\log\left(\frac{x}{2}\right)\right]^3\frac{\partial^2 I_\nu(x)}{\partial \nu^2} \\ &-5\left[\log\left(\frac{x}{2}\right)\right]^4\frac{\partial I_\nu(x)}{\partial \nu} + \left[\log\left(\frac{x}{2}\right)\right]^5 I_\nu(x) - \left(\frac{x}{2}\right)^\nu \sum_{k=0}^{\infty}\frac{Q_5(\nu+k+1)}{k!}\left(\frac{1}{4}x^2\right)^k, \end{aligned}$$

where

$$Q_5(z) = \begin{cases} \left\{\psi^5(z) - 10\psi^3(z)\psi'(z) + 10\psi^2(z)\psi''(z) + 15\psi(z)\left[\psi'(z)\right]^2\right. \\ \left.-5\psi(z)\psi'''(z) - 10\psi'(z)\psi''(z) + \psi^{(iv)}(z)\right\}/\Gamma(z) \;\text{ if } z > 0 \\ \pi^{-1}\Gamma(1-z)f_a(z)\sin(\pi z) + \Gamma(1-z)f_b(z)\cos(\pi z) \;\text{ if } z \leq 0 \end{cases}$$

with

$$\begin{aligned} f_a(z) &= \psi^5(1-z) - 10\pi^2\psi^3(1-z) + 10\psi^3(1-z)\psi'(1-z) + 10\psi^2(1-z)\psi''(1-z) \\ &+15\psi(1-z)\left[\psi'(1-z)\right]^2 + 5\psi(1-z)\psi'''(1-z) + 5\pi^4\psi(1-z) \\ &-30\pi^2\psi(1-z)\psi'(1-z) + 10\psi'(1-z)\psi''(1-z) - 10\pi^2\psi''(1-z) + \psi^{(iv)}(1-z), \end{aligned}$$

and

$$f_b(z) = -5\psi^4(1-z) + 10\pi^2\psi^2(1-z) - 30\psi^2(1-z)\psi'(1-z)$$
$$-20\psi(1-z)\psi''(1-z) - 15\left[\psi'(1-z)\right]^2 + 10\pi^2\psi'(1-z) - 5\psi'''(1-z) - \pi^4.$$

The ratio of two Bessel functions $R_\nu(x)$ deserves particular attention. Jørgensen (1982) shows that

$$R'_\nu(x) = R^2_\nu(x) - (2\nu+1)x^{-1}R_\nu(x) - 1. \tag{64}$$

Once again, (61) can be used to show that

$$R_\nu(x) \simeq 2\nu x^{-1} \tag{65}$$

for small x and fixed $\nu > 0$.

5 Moments of the GIG distribution

If $X \sim GIG(\nu, \delta, \gamma)$, its density function will be

$$\frac{(\gamma/\delta)^\nu}{2K_\nu(\delta\gamma)} x^{\nu-1} \exp\left[-\frac{1}{2}\left(\frac{\delta^2}{x} + \gamma^2 x\right)\right],$$

where $K_\nu(\cdot)$ is the modified Bessel function of the third kind and $\delta, \gamma \geq 0$, $\nu \in \mathbb{R}$, $x > 0$. Two important properties of this distribution are $X^{-1} \sim GIG(-\nu, \gamma, \delta)$ and $(\gamma/\delta)X \sim GIG\left(\nu, \sqrt{\gamma\delta}, \sqrt{\gamma\delta}\right)$. For our purposes, the most useful moments of X when $\delta\gamma > 0$ are

$$E\left(X^k\right) = \left(\frac{\delta}{\gamma}\right)^k \frac{K_{\nu+k}(\delta\gamma)}{K_\nu(\delta\gamma)} \tag{66}$$
$$E(\log X) = \log\left(\frac{\delta}{\gamma}\right) + \frac{\partial}{\partial\nu}K_\nu(\delta\gamma). \tag{67}$$

The GIG nests some well-known important distributions, such as the gamma ($\nu > 0$, $\delta = 0$), the reciprocal gamma ($\nu < 0$, $\gamma = 0$) or the inverse Gaussian ($\nu = -1/2$). Importantly, all the moments of this distribution are finite, except in the reciprocal gamma case, in which (66) becomes infinite for $k \geq |\nu|$. A complete discussion on this distribution can be found in Jørgensen (1982), who also presents several useful Gaussian approximations based on the following limits:

$$\sqrt{\delta\gamma}[(\gamma x/\delta) - 1] \overset{\delta\gamma\to\infty}{\to} N(0,1)$$
$$\sqrt{\delta\gamma}\log(\gamma x/\delta) \overset{\delta\gamma\to\infty}{\to} N(0,1)$$
$$\frac{\gamma^2}{2\sqrt{\nu}}\left[x - \frac{2\nu}{\gamma^2}\right] \overset{\nu\to+\infty}{\to} N(0,1)$$
$$\frac{-2\nu^{3/2}}{\delta^2}\left[x + \frac{\delta^2}{2\nu}\right] \overset{\nu\to-\infty}{\to} N(0,1)$$

6 Power of the normality tests

We can determine the power of the sup test by rewriting it as a quadratic form in

$$\begin{bmatrix} 2/[N(N+2)] & \mathbf{0}' \\ \mathbf{0} & \hat{\mathbf{\Sigma}}^{-1}/[2(N+2)] \end{bmatrix}$$

evaluated at $\bar{m}_T(\tilde{\boldsymbol{\theta}}_T) = [\bar{m}_{kT}(\tilde{\boldsymbol{\theta}}_T), \bar{m}'_{sT}(\tilde{\boldsymbol{\theta}}_T)]'$, where $\tilde{\boldsymbol{\theta}}_T$ must be interpreted as a PML estimator of $\boldsymbol{\theta}_0 = (\boldsymbol{\mu}'_0, vech'(\mathbf{\Sigma}_0))'$ under the alternative of *GH* innovations. Hence, its asymptotic distribution will be given by the robust formulae provided by Bollerslev and Wooldridge (1992), which, in terms of the Gaussian score can be written as

$$\sqrt{T}\left[\tilde{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\right] = \mathcal{A}^{-1}(\boldsymbol{\theta}_0)\sqrt{T}\bar{s}_{\boldsymbol{\theta}T}(\boldsymbol{\theta}_0, 0, 0, \mathbf{0}) + o_p(1),$$

where

$$\mathcal{A}(\boldsymbol{\phi}_0) = \frac{\partial \boldsymbol{\mu}'}{\partial \boldsymbol{\theta}}\mathbf{\Sigma}^{-1}\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}} + \frac{1}{2}\frac{\partial vec'\mathbf{\Sigma}}{\partial \boldsymbol{\theta}}\left[\mathbf{\Sigma}^{-1} \otimes \mathbf{\Sigma}^{-1}\right]\frac{\partial vec\mathbf{\Sigma}}{\partial \boldsymbol{\theta}}.$$

Hence, the usual Taylor expansion around the true parameter values yields

$$\sqrt{T}\bar{m}_T\left(\tilde{\boldsymbol{\theta}}_T\right) = \begin{bmatrix} -\mathcal{B}(\boldsymbol{\theta}_0)\mathcal{A}^{-1}(\boldsymbol{\theta}_0) & \mathbf{I}_{N+1} \end{bmatrix}\sqrt{T}\begin{bmatrix} \bar{s}_{\boldsymbol{\theta}T}(\boldsymbol{\theta}_0, 0, 0, \mathbf{0}) \\ \bar{m}_T(\boldsymbol{\theta}_0) \end{bmatrix} + o_p(1), \qquad (68)$$

where $\mathcal{B}(\boldsymbol{\theta}_0) = -E[\partial \bar{m}_T(\boldsymbol{\theta}_0)/\partial \boldsymbol{\theta}']$

Fortunately, $\mathcal{A}(\boldsymbol{\phi}_0)$, $\mathcal{B}(\boldsymbol{\theta}_0)$, as well as the mean and variance of $\bar{\mathbf{s}}_{\boldsymbol{\theta}t}(\boldsymbol{\theta}_0)$ and $\bar{m}_T(\boldsymbol{\theta}_0)$ under the alternative can be computed analytically by using the location-scale mixture of normals interpretation of the *GH* distribution. In particular, we can write

$$\boldsymbol{\varepsilon}^*_t = c(\boldsymbol{\phi})\mathbf{b}(h_t - 1) + \sqrt{h_t}\mathbf{A}\mathbf{r}_t,$$

$$\varsigma_t = \boldsymbol{\varepsilon}^{*\prime}_t\boldsymbol{\varepsilon}^*_t = c^2(\boldsymbol{\phi})(h_t - 1)^2\mathbf{b}'\mathbf{b} + 2c(\boldsymbol{\phi})\sqrt{h_t}(h_t - 1)\mathbf{b}'\mathbf{A}\mathbf{r}_t + h_t\mathbf{r}'_t\mathbf{A}'\mathbf{A}\mathbf{r}_t,$$

with $h_t = \xi_t^{-1}\gamma/R_\nu(\gamma)$, and

$$\mathbf{A} = \left[\mathbf{I}_N + \frac{c(\boldsymbol{\phi}, \nu, \gamma) - 1}{\mathbf{b}'\mathbf{b}}\mathbf{b}\mathbf{b}'\right]^{\frac{1}{2}},$$

where $\mathbf{r}_t|\mathbf{z}_t, I_{t-1} \sim N(0, \mathbf{I}_N)$ and $\xi_t|\mathbf{z}_t, I_{t-1} \sim GIG[.5\eta^{-1}, \psi^{-1}(1-\psi), 1]$ are mutually independent. But since both ξ_t and $\mathbf{r}_t$ are *iid*, then $\boldsymbol{\varepsilon}^*_t$ and $\varsigma_t = \boldsymbol{\varepsilon}^{*\prime}_t\boldsymbol{\varepsilon}^*_t$ will also be *iid*. As a result, given that all the moments of normal and *GIG* random variables are finite (except when $\psi = 1$, in which case some moments may become unbounded for large enough η; see Jørgensen, 1982), we can apply the Lindeberg-Lévy Central Limit Theorem to show that the asymptotic distribution of $\sqrt{T}\bar{m}_T(\tilde{\boldsymbol{\theta}}_T)$ is $N[m(\boldsymbol{\theta}_0, \eta, \psi, \mathbf{b}), V(\boldsymbol{\theta}_0, \eta, \psi, \mathbf{b})]$, where the required expressions can be computed from (68). In particular, we can use Magnus (1986) to evaluate the moments of quadratic forms of normals, such as $\mathbf{r}_t'\mathbf{A}'\mathbf{A}\mathbf{r}_t$.

Finally, we can use Koerts and Abrahamse's (1969) implementation of Imhof's procedure for evaluating the probability that a quadratic form of normals is less than a given value (see also Farebrother, 1990).

To obtain the power of the KT test, we will use the following alternative formulation

$$\frac{KT}{T} = \frac{2}{N(N+2)}\bar{m}_{kT}^2(\tilde{\boldsymbol{\theta}}_T) \cdot \mathbf{1}\left(\bar{m}_{kT}\left(\tilde{\boldsymbol{\theta}}_T\right) \geq 0\right) + \frac{1}{2(N+2)}\bar{m}_{sT}'(\tilde{\boldsymbol{\theta}}_T)\hat{\boldsymbol{\Sigma}}^{-1}\bar{m}_{sT}(\tilde{\boldsymbol{\theta}}_T).$$

Hence, the distribution function of the KT statistic can be expressed as

$$\Pr\left(\frac{KT}{T} < x\right) = \int_{-\infty}^{\infty} \Pr\left(\frac{KT}{T} < x \middle| \bar{m}_{kt} = l\right) f_k(l)\, dl, \tag{69}$$

where $f_k(\cdot)$ is the pdf of the distribution of the kurtosis component. But since the joint asymptotic distribution of $\sqrt{T}\bar{m}_T(\tilde{\boldsymbol{\theta}}_T)$ is normal, so that the conditional distribution of $\sqrt{T}\bar{m}_{sT}(\tilde{\boldsymbol{\theta}}_T)$ given $\sqrt{T}\bar{m}_{kT}(\tilde{\boldsymbol{\theta}}_T)$ will also be normal, the KT test can also be written as a quadratic form of normals for each value of the kurtosis component. As a result, we can use Imhof's procedure again to evaluate

$$\begin{aligned} & \Pr\left[\frac{1}{2(N+2)}\bar{m}_{sT}(\tilde{\boldsymbol{\theta}}_T)\hat{\boldsymbol{\Sigma}}^{-1}\bar{m}_{sT}(\tilde{\boldsymbol{\theta}}_T) < x - \frac{2}{N(N+2)}l^2 \cdot \mathbf{1}(l \geq 0) \middle| \bar{m}_{kt} = l\right] \\ = & \Pr\left(\frac{KT}{T} < x \middle| \bar{m}_{kt} = l\right). \end{aligned}$$

Once we know this conditional probability, we can evaluate the integral in (69) by numerical integration with a standard quadrature algorithm.

References

Aas, K., X. Dimakos, and I. Haff (2005). Risk estimation using the multivariate normal inverse gaussian distribution. *Journal of Risk 8*, 39–60.

Abramowitz, M. and A. Stegun (1965). *Handbook of mathematical functions.* New York: Dover Publications.

Barndorff-Nielsen, O. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Proc. R. Soc. 353*, 401–419.

Barndorff-Nielsen, O. and N. Shephard (2001). Normal modified stable processes. *Theory of Probability and Mathematical Statistics 65*, 1–19.

Bollerslev, T. and J. Wooldridge (1992). Quasi maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews 11*, 143–172.

Cajigas, J. and G. Urga (2007). A risk management analysis using the AGDCC model with asymmetric multivariate Laplace distribution of innovations. mimeo CBS.

Chen, Y., W. Härdle, and S. Jeong (2008). Nonparametric risk management with Generalized Hyperbolic distributions. *Journal of the American Statistical Association 103*, 910–923.

Farebrother, R. (1990). The distribution of a quadratic form in normal variables (Algorithm AS 256.3). *Applied Statistics 39*, 294–309.

Fiorentini, G., E. Sentana, and G. Calzolari (2003). Maximum likelihood estimation and inference in multivariate conditionally heteroskedastic dynamic regression models with Student t innovations. *Journal of Business and Economic Statistics 21*, 532–546.

Jørgensen, B. (1982). *Statistical properties of the generalized inverse Gaussian distribution.* New York: Springer-Verlag.

Koerts, J. and A. P. J. Abrahamse (1969). *On the theory and application of the general linear model.* Rotterdam: Rotterdam University Press.

Madan, D. B. and F. Milne (1991). Option pricing with V.G. martingale components. *Mathematical Finance 1*, 39–55.

Magnus, J. R. (1986). The exact moments of a ratio of quadratic forms in normal variables. *Annales d'Économie et de Statistique 4*, 95–109.

Magnus, J. R. and H. Neudecker (1988). *Matrix differential calculus with applications in statistics and econometrics.* Chichester: John Wiley and Sons.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika 57*, 519–530.

Mencía, J. and E. Sentana (2009). Multivariate location-scale mixtures of normals and mean-variance-skewness portfolio allocation. *Journal of Econometrics 153*, 105–121.

Yañez, S., M. C. Jaramillo, and J. C. Correa (1999). Una revisión de medidas multivariadas de asimetría y kurtosis para pruebas de multinormalidad. *Revista Colombiana de Estadística 22*, 5–16.