Noncentral Skewed HEAVY Model for High-Dimensional Realized Returns and Covariances Matrices

TARAS BODNAR[a], YAREMA OKHRIN[b], NESTOR PAROLYA[c] AND MICHAEL STOLLENWERK[d]

[a] *Department of Mathematics, Stockholm University, 10691 Stockholm, Sweden*

[b] *Department of Statistics, Augsburg University, Germany*

e-mail: yarema.okhrin@wiwi.uni-augsburg.de

[c] *Department of Applied Mathematics, Delft University of Technology, The Netherlands*

e-mail: n.parolya@tudelft.nl

[d] *Faculty of Economics and Social Studies, Heidelberg University, Bergheimer Strasse 58, 69115 Heidelberg, Germany*

michael.stollenwerk@awi.uni-heidelberg.de

Abstract

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

JEL subject classifications:

Keywords:

1 Introduction

2 Noncentral HEAVY GAS skew tF Model

Assumptions:

(A1) The vector of asset returns $\mathbf{y}_t \in \mathbb{R}^p$ conditionally on the observation set $\mathcal{F}_{t-1}$ is assumed to follow a normalized multivariate skew Student's t distribution with density given by

$$f_y(\mathbf{y}_t|\mathbf{\Sigma}_t, \mathcal{F}_{t-1}; \nu_0) = 2 f_{T_{\nu_0}}(\mathbf{y}_t|\mathbf{\Sigma}_t, \mathcal{F}_{t-1}) F_{T_{\nu_0+p}}\left(\frac{\boldsymbol{\alpha}^\top \mathbf{y}_t}{(\nu_0 - 2 + \mathbf{y}_t^\top \mathbf{\Sigma}_t^{-1}\mathbf{y}_t)^{1/2}}\sqrt{\nu_0 + p}\right), \tag{2.1}$$

where the function $F_{T_{\nu_0}}(\cdot)$ denotes the c.d.f. of the standard univariate t-distribution with ν_0 degrees of freedom and $f_{T_{\nu_0}}(\cdot)$ is density of the (standardized) multivariate central t distribution with ν_0 degrees of freedom given by

$$f_{T_{\nu_0}}(\mathbf{y}_t|\mathbf{\Sigma}_t, \mathcal{F}_{t-1}) = \frac{\Gamma[(\nu_0+p)/2]}{\Gamma[\nu_0/2][\pi(\nu_0-2)]^{p/2}} |\mathbf{\Sigma}_t|^{-1/2} \left(1 + \frac{\mathbf{y}_t^\top \mathbf{\Sigma}_t^{-1}\mathbf{y}_t}{\nu_0 - 2}\right)^{-(\nu_0+p)/2}. \tag{2.2}$$

The vector $\boldsymbol{\alpha}$ stays for the skewness parameters.

Basic properties of multivariate skew t:

(i) $\boldsymbol{\alpha} = \mathbf{0}$ implies the central multivariate t-distribution with ν_0 degrees of freedom.

(ii) $\lim_{\nu_0 \to \infty} f_{T_{\nu_0}}(\mathbf{y}_t) = 2\phi_{\mathbf{0},\mathbf{\Sigma}_t}(\mathbf{y}_t|\mathbf{\Sigma}_t, \mathcal{F}_{t-1})\Phi(\boldsymbol{\alpha}^\top \mathbf{y}_t)$, i.e., the multivariate skew t tends to the multivariate skew normal as $\nu_0 \to \infty$.

(iii) Let $\sigma_{t,jj}$ is the j-th diagonal element of $\mathbf{\Sigma}_t$. Then $\sigma_{t,jj}^{-1} y_{t,j}^2 \sim F(1, \nu_0)$ and it is independent of $\boldsymbol{\alpha}$ and of $\mathcal{F}_{t-1}$.

(iv) the quadratic form $p^{-1}\mathbf{y}_t^\top \mathbf{\Sigma}_t^{-1}\mathbf{y}_t \sim F(p, \nu_0)$ and it is independent of $\boldsymbol{\alpha}$ and of $\mathcal{F}_{t-1}$.

(v) At last, the multivariate skew t-distribution with ν_0 degrees of freedom obeys the following stochastic representation

$$\mathbf{y}_t \overset{d}{\sim} \frac{\mathbf{x}_t}{\sqrt{W_t/\nu_0}}$$

with $\mathbf{x}_t|\mathbf{\Sigma}_t, \mathcal{F}_{t-1} \sim \mathcal{SN}(\mathbf{0}, \mathbf{\Sigma}_t, \boldsymbol{\alpha})$ (multivariate skew normal) and $W_t \sim \chi^2_{\nu_0}$ independent of $\mathbf{x}_t$. Or, using the stochastic representation of $\mathbf{x}_t$, we get

$$\mathbf{y}_t \overset{d}{\sim} \frac{1}{\sqrt{W_t/\nu_0}}\left((\mathbf{\Sigma}_t^{-1} + \boldsymbol{\alpha}\boldsymbol{\alpha}^\top)^{-1/2}\mathbf{z}_t + \frac{\mathbf{\Sigma}_t\boldsymbol{\alpha}}{(1+\boldsymbol{\alpha}^\top\mathbf{\Sigma}_t\boldsymbol{\alpha})^{1/2}}|u_t|\right), \tag{2.3}$$

where $\mathbf{z}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, $u_t \sim \mathcal{N}(0,1)$ and W_t are independent.

(vi) The first two moments are given by

$$\mathbb{E}(\mathbf{y}_t|\mathbf{\Sigma}_t, \mathcal{F}_{t-1}) = \sqrt{\frac{\nu_0}{\pi}}\frac{\Gamma\left(\frac{\nu_0-1}{2}\right)}{\Gamma\left(\frac{\nu_0}{2}\right)}\frac{\mathbf{\Sigma}_t\boldsymbol{\alpha}}{(1+\boldsymbol{\alpha}^\top\mathbf{\Sigma}_t\boldsymbol{\alpha})^{1/2}}, \quad \nu_0 > 1 \tag{2.4}$$

$$\mathbb{C}\mathrm{ov}(\mathbf{y}_t|\mathbf{\Sigma}_t, \mathcal{F}_{t-1}) = \frac{\nu_0}{\nu_0-2}\mathbf{\Sigma}_t - \frac{\nu_0}{\pi}\left(\frac{\Gamma\left(\frac{\nu_0-1}{2}\right)}{\Gamma\left(\frac{\nu_0}{2}\right)}\right)^2\frac{\mathbf{\Sigma}_t\boldsymbol{\alpha}\boldsymbol{\alpha}^\top\mathbf{\Sigma}_t}{1+\boldsymbol{\alpha}^\top\mathbf{\Sigma}_t\boldsymbol{\alpha}}, \quad \nu_0 > 2\,. \tag{2.5}$$

(vii) Thus, the multivariate skew t can be seen as a generalization of a non-central multivariate t-distribution.

(A2) The realized covariance matrix RC_t conditionally on $\mathcal{F}_{t-1}$ is assumed to follow a rescaled matrix variate non-central F distribution with degrees of freedom ν_1, ν_2 and non-centrality matrix $\mathbf{\Sigma}_t^{-1}\mathbf{\Omega}_t$ with density given by

$$
\begin{aligned}
&f_{RC}(RC_t|\mathbf{\Sigma}_t,\mathbf{\Omega}_t,\mathcal{F}_{t-1};\nu_1,\nu_2)\\
&= K(\nu_1,\nu_2)\left|\frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t^{-1}\right|^{\nu_1/2}\frac{|RC_t|^{(\nu_1-p-1)/2}}{\left|\mathbf{I}+\frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t^{-1}RC_t\right|^{(\nu_1+\nu_2)/2}}\\
&\quad\times e^{\mathrm{tr}\left(-\frac{1}{2}\mathbf{\Sigma}_t^{-1}\mathbf{\Omega}_t\right)}{}_1F_1\left(\frac{\nu_1+\nu_2}{2};\frac{\nu_1}{2};\frac{1}{2}\frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t^{-1}\mathbf{\Omega}_t\mathbf{\Sigma}_t^{-1}RC_t[\mathbf{I}+\frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t^{-1}RC_t]^{-1}\right)
\end{aligned}
\tag{2.6}
$$

$$
\begin{aligned}
&= K(\nu_1,\nu_2)\left(\frac{\nu_1}{\nu_2-p-1}\right)^{p\nu_1/2}|RC_t|^{-(\nu_2+p+1)/2}|\mathbf{\Sigma}_t|^{(\nu_1+2\nu_2)/2}\left|\mathbf{\Sigma}_tRC_t^{-1}\mathbf{\Sigma}_t+\frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t\right|^{-(\nu_1+\nu_2)/2}\\
&\quad\times e^{\mathrm{tr}\left(-\frac{1}{2}\mathbf{\Sigma}_t^{-1}\mathbf{\Omega}_t\right)}{}_1F_1\left(\frac{\nu_1+\nu_2}{2};\frac{\nu_1}{2};\frac{1}{2}\frac{\nu_1}{\nu_2-p-1}\mathbf{\Omega}_t[\mathbf{\Sigma}_tRC_t^{-1}\mathbf{\Sigma}_t+\frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t]^{-1}\right)
\end{aligned}
\tag{2.7}
$$

$$
\begin{aligned}
&= \frac{K(\nu_1,\nu_2)}{\Gamma_p[(\nu_1+\nu_2)/2]}\left(\frac{\nu_1}{\nu_2-p-1}\right)^{p\nu_1/2}|RC_t|^{-(\nu_2+p+1)/2}|\mathbf{\Sigma}_t|^{(\nu_1+2\nu_2)/2}e^{\mathrm{tr}\left(-\frac{1}{2}\mathbf{\Sigma}_t^{-1}\mathbf{\Omega}_t\right)}\\
&\quad\times\int\limits_{\mathbf{V}>\mathbf{O}}\exp\left(-\frac{1}{2}\mathrm{tr}((\mathbf{\Sigma}_tRC_t^{-1}\mathbf{\Sigma}_t+\frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t)\mathbf{V})\right)|\mathbf{V}|^{\frac{\nu_1+\nu_2-p-1}{2}}{}_0F_1\left(\frac{\nu_1}{2};\frac{1}{4}\frac{\nu_1}{\nu_2-p-1}\mathbf{\Omega}_t\mathbf{V}\right)\mathrm{d}\mathbf{V}
\end{aligned}
\tag{2.8}
$$

$$
\begin{aligned}
&= \frac{K(\nu_1,\nu_2)}{\Gamma_p[(\nu_1+\nu_2)/2]}\left(\frac{\nu_1}{\nu_2-p-1}\right)^{p\nu_1/2}|RC_t|^{-(\nu_2+p+1)/2}|\mathbf{\Sigma}_t|^{(\nu_1+2\nu_2)/2}\left|\mathbf{\Sigma}_tRC_t^{-1}\mathbf{\Sigma}_t+\frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t\right|^{-(\nu_1+\nu_2)/2}\\
&\quad\times 2^{-(\nu_1+\nu_2)/2}|\mathbf{\Omega}_t|^{-(\nu_1+\nu_2)/2}e^{\mathrm{tr}\left(-\frac{1}{2}\mathbf{\Sigma}_t^{-1}\mathbf{\Omega}_t\right)}\int\limits_{\mathbf{V}>\mathbf{O}}\exp\left(-\frac{1}{2}\mathrm{tr}(\mathbf{\Omega}_t^{-1}\mathbf{V})\right)|\mathbf{V}|^{\frac{\nu_1+\nu_2-p-1}{2}}\\
&\quad\times{}_0F_1\left(\frac{\nu_1}{2};\frac{1}{4}\frac{\nu_1}{\nu_2-p-1}[\mathbf{\Sigma}_tRC_t^{-1}\mathbf{\Sigma}_t+\frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t]^{-1}\mathbf{V}\right)\mathrm{d}\mathbf{V},
\end{aligned}
\tag{2.9}
$$

where $K(\nu_1,\nu_2)=\frac{\Gamma_p[(\nu_1+\nu_2)/2]}{\Gamma_p[\nu_1/2]\Gamma_p[\nu_2/2]}$ with $\Gamma_p(x)=\pi^{p(p-1)/4}\prod_{i=1}^{p}\Gamma(x-(i-1)/2)$ for $x>\frac{p-1}{2}$ is the multivariate Gamma function. The function ${}_1F_1(\cdot)$ is the matrix variate hypergeometric function

$$
{}_qF_k(a_1,\ldots,a_k;b_1,\ldots,b_k;\mathbf{A})=\sum_{i=0}^{\infty}\sum_{l}\frac{(a_1)_l\ldots(a_q)_l}{(b_1)_l\ldots(b_k)_l}\frac{C_l(\mathbf{A})}{i!}\,, \tag{2.10}
$$

where $\mathbf{A}$ is a $p\times p$ symmetric matrix, $\sum_l$ denotes the summation over all partitions l, $(c)_l=\frac{\Gamma_p(c,l)}{\Gamma_p(c)}$ with $\Gamma_p(c,l)=\pi^{p(p-1)/4}\prod_{j=1}^{p}\Gamma(c+i_j-(j-1)/2)$ for $c>(p-1)/2-i_m$ and $l=(i-1,\ldots,i_m)$ with $i_1\geq i_2\geq\ldots i_m\geq 0$ and $\sum_j^p i_j=i$; $C_l(\mathbf{A})$ stands for the zonal polynomial.

We used Gupta & Nagar (2000), Theorem 1.6.2 to get from (2.7) to (2.8),

$$
\begin{aligned}
&{}_1F_1\left(\frac{\nu_1+\nu_2}{2};\frac{\nu_1}{2};\frac{1}{2}\frac{\nu_1}{\nu_2-p-1}\mathbf{\Omega}_t[\mathbf{\Sigma}_tRC_t^{-1}\mathbf{\Sigma}_t+\frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t]^{-1}\right)\\
&=\frac{1}{\Gamma_p[(\nu_1+\nu_2)/2]}\left|\mathbf{\Sigma}_tRC_t^{-1}\mathbf{\Sigma}_t+\frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t\right|^{\frac{\nu_1+\nu_2}{2}}\\
&\times\int\limits_{\mathbf{V}>\mathbf{O}}\exp\left(-\frac{1}{2}\mathrm{tr}((\mathbf{\Sigma}_tRC_t^{-1}\mathbf{\Sigma}_t+\frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t)\mathbf{V})\right)|\mathbf{V}|^{\frac{\nu_1+\nu_2-p-1}{2}}{}_0F_1\left(\frac{\nu_1}{2};\frac{1}{4}\frac{\nu_1}{\nu_2-p-1}\mathbf{\Omega}_t\mathbf{V}\right)\mathrm{d}\mathbf{V},
\end{aligned}
$$

and similarly to get from (2.7) to (2.9). (2.8) is useful to derive the score w.r.t. $\mathbf{\Sigma}_t$, (2.9) is useful to derive the score w.r.t. $\mathbf{\Omega}_t$.

Basic properties of non-central matrix variate F distribution:

(i) Obviously, if $\mathbf{\Omega}_t \equiv \mathbf{O}$ the distribution becomes the central one and the density function simplifies to

$$f_{RC}(RC_t|\mathbf{\Sigma}_t, \mathcal{F}_{t-1}; \nu_1, \nu_2) = K(\nu_1, \nu_2) \frac{\left|\frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t^{-1}\right|^{\nu_1/2} |RC_t|^{(\nu_1-p-1)/2}}{\left|\mathbf{I} + \frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t^{-1} RC_t\right|^{(\nu_1+\nu_2)/2}}. \tag{2.11}$$

(ii) For $\nu_2 \to \infty$ it collapses to a non-central Wishart distribution with ν_1 degrees of freedom.

(iii) Let $\mathbf{\Omega}_t = \mathbf{\Lambda}_t \mathbf{\Lambda}_t^\top$ where $\mathbf{\Lambda}_t$ is a $p \times \nu_1$ matrix. The matrix F distribution obeys the following stochastic representation

$$RC_t \overset{d}{\sim} \frac{\nu_2 - p - 1}{\nu_1} \mathbf{\Sigma}_t^{1/2} \mathbf{V}_t^{1/2} \mathbf{T}_t^{-1} \mathbf{V}_t^{1/2} \mathbf{\Sigma}_t^{1/2}, \tag{2.12}$$

where $\mathbf{V}_t = (\mathbf{X}_t + \mathbf{\Sigma}_t^{-1/2}\mathbf{\Lambda}_t)(\mathbf{X}_t + \mathbf{\Sigma}_t^{-1/2}\mathbf{\Lambda}_t)^\top$ has a standard non-central Wishart distribution with ν_1 degrees of freedom and non-centrality matrix $\mathbf{\Sigma}_t^{-1/2}\mathbf{\Lambda}_t\mathbf{\Lambda}_t^\top\mathbf{\Sigma}_t^{-1/2}$, $\mathbf{X}_t \sim \mathcal{N}_{p,\nu_1}(\mathbf{O}, \mathbf{I}_p \otimes \mathbf{I}_{\nu_1})$ is matrix variate standard normal, $\mathbf{V}_t^{1/2}$ denotes a symmetric square root of $\mathbf{V}_t$ and $\mathbf{T}_t$ is a standard Wishart matrix with ν_2 degrees of freedom.

(iv) The mean of $RC_t|\mathbf{\Sigma}_t, \mathcal{F}_{t-1}$ is equal to $\mathbf{\Sigma}_t + \frac{1}{\nu_1}\mathbf{\Omega}_t$.

(A3) Dynamics for $\mathbf{\Sigma}_t$ and $\mathbf{\Omega}_t$:

$$\mathbf{\Sigma}_t = \mathbf{\Xi} + \alpha \mathbf{S}_{t-1} + \beta \mathbf{\Sigma}_{t-1} \quad \text{GAS recursion} \tag{2.13}$$

$$\mathbf{\Omega}_t = \mathbf{M} RC_{t-1} \mathbf{M}^\top \quad \text{WAR specification} \tag{2.14}$$

The matrix $\mathbf{M}$ could be assumed of some simple structure (diagonal, sparse etc.); $\mathbf{\Xi}$ is the unconditional covariance matrix and $\mathbf{S}_t$ is the scaled score matrix.

(A4) Independence of error processes:
We assume that all five error processes $\{\mathbf{z}_t\}$, $\{u_t\}$, $\{W_t\}$, $\{\mathbf{V}_t\}$ (or $\{\mathbf{X}_t\}$), and $\{\mathbf{T}_t\}$ in (2.3) and (2.12) are mutually independent. Assumption (A4) in particularly implies that the return vector $\mathbf{y}_t$ and the realized covariance matrix RC_t conditionally on $\mathcal{F}_{t-1}$ are independent.

2.1 t-noncentral-spiked-F scalar (p, q, r) model

The model of the t-noncentral-spiked-F (p, q, r) model dynamics (Eq. 2.9 and 2.10) are given by

$$\mathbf{\Sigma}_t = \mathbf{\Xi} + \sum_{i=1}^{p} \alpha_i \mathbf{S}_{\mathbf{\Sigma}_{t-i}} + \sum_{j=1}^{q} \beta_i \mathbf{\Sigma}_{t-i}$$

$$\mathbf{\Omega}_t = \sum_{i=1}^{r} \mathbf{M}_i \mathbf{S}_{\mathbf{\Omega}_{t-i}} \mathbf{M}_i',$$

where $\mathbf{\Xi}$ is a positive definite p by p matrix, α_i and β_i are scalars and $\mathbf{M}_i = \mathbf{m}_i\mathbf{m}_i'$ and $\mathbf{m}_i$ is a p by 1 vector, such that $\mathbf{\Omega}_t$ is rank 1 and the noncentral-F becomes the spiked noncentral-F.

Since (assuming stationarity),

$$\begin{aligned}
\mathbf{E}[RC_t] &= \mathbf{E}[\Sigma_t] + \frac{1}{\nu_1}\mathbf{E}[\Omega_t] \\
&= \frac{\mathbf{\Xi}}{(1-\sum_{j=1}^{q}\beta_i)} + \frac{1}{\nu_1}\mathbf{E}\left[\sum_{i=1}^{r}\mathbf{M}_i RC_{t-i}\mathbf{M}_i'\right] \\
&= \frac{\mathbf{\Xi}}{(1-\sum_{j=1}^{q}\beta_i)} + \frac{1}{\nu_1}\sum_{i=1}^{r}\mathbf{M}_i\mathbf{E}\left[RC_t\right]\mathbf{M}_i' \\
&\Leftrightarrow \\
\mathbf{\Xi} &= \left(1-\sum_{i=1}^{q}\beta_i\right)\left(\mathbf{E}[RC_t] - \frac{1}{\nu_1}\sum_{i=1}^{r}\mathbf{M}_i\mathbf{E}[RC_t]\mathbf{M}_i'\right),
\end{aligned}$$

we can use

$$\hat{\mathbf{\Xi}} = \left(1-\sum_{i=1}^{q}\beta_i\right)\left(\frac{1}{T}\sum_{t=1}^{T}RC_t - \frac{\sum_{j=1}^{r}\mathbf{M}_j\frac{1}{T}\sum_{t=1}^{T}RC_t\mathbf{M}_j'}{\nu_1}\right).$$

The the only parameters to estimate via ML are α_i, β_i and $\mathbf{m}_i$, such that the estimation is fast and the model is applicable in higher dimensions.

2.2 t-noncentral-F full (p, q, r) model

The model of the t-noncentral-spiked-F (p, q, r) full model dynamics (Eq. 2.9 and 2.10) are given by

$$\begin{aligned}
\mathbf{\Sigma}_t &= \mathbf{\Xi} + \sum_{i=1}^{p}\mathbf{A}_i RC_{t-i}\mathbf{A}_i' + \sum_{j=1}^{q}\mathbf{B}_j\mathbf{\Sigma}_{t-i}\mathbf{B}_j' \\
\mathbf{\Omega}_t &= \sum_{i=1}^{r}\mathbf{M}_i RC_{t-i}\mathbf{M}_i,
\end{aligned}$$

where $\mathbf{A}_i$, $\mathbf{B}_i$ and $\mathbf{M}_i$ are p by p matrices and $\mathbf{\Xi}$ is a positive definite p by p matrix. All parameters are estimated via one-step ML.

3 Likelihood function and score matrix

Let $\text{vech}(\mathbf{A}) = (a_{11}, ..., a_{k1}, ..., a_{ii}, ..., a_{ki}, ...a_{kk})^\top$ for an arbitrary symmetric matrix $\mathbf{A} = (a_{ij})$. In the following we make use of the operator **vec** defined by $\text{vec}(\mathbf{A}) = (a_{11}, ..., a_{k1}, ..., a_{1i}, ..., a_{ki}, ...a_{kk})^\top$. For the properties of the operators **vech** and **vec** we refer to Harville (1997). Moreover, let $\mathbf{D}_k$ be a $k^2 \times k(k+1)/2$ duplication matrix such that $\mathbf{D}_k\text{vech}(A) = \text{vec}(A)$ and $\mathbf{D}_k^+ = (\mathbf{D}_k^\top\mathbf{D}_k)^{-1}\mathbf{D}_k^\top$ with the property $\mathbf{D}_k^+\text{vec}(A) = \text{vech}(A)$. Finally, the symbol $\otimes$ stand for the Kronecker product.

In the derivation of the likelihood we use following two results (see, Harville (1997, p. 368))

$$\frac{\partial(\text{vec}(\mathbf{A}^{-1}))^\top}{\partial\text{vech}(\mathbf{A})} = -\mathbf{D}_k^\top(\mathbf{A}^{-1}\otimes\mathbf{A}^{-1})\mathbf{D}_k^{+\top}\mathbf{D}_k^\top\,. \tag{2.15}$$

and

$$\frac{\partial\log|\mathbf{A}|}{\partial\text{vech}(\mathbf{A})} = \mathbf{D}_k^\top\text{vec}(\mathbf{A}^{-1})\,, \tag{2.16}$$

where (2.16) follows from Magnus and Neudecker (1999, p. 171) since

$$\mathrm{d}\log|\mathbf{A}| = \mathrm{tr}(\mathbf{A}^{-1}\,\mathrm{d}\mathbf{A}) = \mathrm{vec}(\mathbf{A}^{-1})^\top\ \mathrm{dvec}(\mathbf{A}) = \mathrm{vec}(\mathbf{A}^{-1})^\top\mathbf{D}_k\ \mathrm{dvech}(\mathbf{A})$$

The conditional log-likelihood of the considered model is given by

$$\mathcal{L}_t = \mathcal{L}_{t;1} + \mathcal{L}_{t;2} \tag{2.17}$$

with

$$\mathcal{L}_{t;1} = \log f_y(\mathbf{y}_t|\boldsymbol{\Sigma}_t, \mathcal{F}_{t-1};\nu_0)$$

and

$$\mathcal{L}_{t;2} = \log f_{RC}(RC_t|\boldsymbol{\Sigma}_t, \boldsymbol{\Omega}_t, \mathcal{F}_{t-1};\nu_1,\nu_2)$$

3.1 Skew-t score

3.1.1 Score w.r.t. $\boldsymbol{\Sigma}_t$

Then, we compute

$$\begin{aligned}
\frac{\partial\mathcal{L}_{t;1}}{\partial\mathrm{vech}(\boldsymbol{\Sigma}_t)} &= -\frac{1}{2}\frac{\partial\log|\boldsymbol{\Sigma}_t|}{\partial\mathrm{vech}(\boldsymbol{\Sigma}_t)} - \frac{\nu_0+p}{2}\frac{\partial\log\left(1+\frac{\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t}{\nu_0-2}\right)}{\partial\mathrm{vech}(\boldsymbol{\Sigma}_t)} \\
&+ \frac{\partial\log\left(F_{T_{\nu_0+p}}\left(\frac{\boldsymbol{\alpha}^\top\mathbf{y}_t}{(\nu_0-2+\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t)^{1/2}}\sqrt{\nu_0+p}\right)\right)}{\partial\mathrm{vech}(\boldsymbol{\Sigma}_t)} \\
&= -\frac{1}{2}\mathbf{D}_p^\top\mathrm{vec}(\boldsymbol{\Sigma}_t^{-1}) - \frac{\nu_0+p}{2}\left(1+\frac{\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t}{\nu_0-2}\right)^{-1}\frac{1}{\nu_0-2}\frac{\partial(\mathrm{vec}(\boldsymbol{\Sigma}_t^{-1}))^\top}{\partial\mathrm{vech}(\boldsymbol{\Sigma}_t)}(\mathbf{y}_t\otimes\mathbf{y}_t) \\
&- \frac{1}{2}\frac{f_{T_{\nu_0+p}}\left(\frac{\boldsymbol{\alpha}^\top\mathbf{y}_t}{(\nu_0-2+\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t)^{1/2}}\sqrt{\nu_0+p}\right)}{F_{T_{\nu_0+p}}\left(\frac{\boldsymbol{\alpha}^\top\mathbf{y}_t}{(\nu_0-2+\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t)^{1/2}}\sqrt{\nu_0+p}\right)}\frac{\sqrt{\nu_0+p}\boldsymbol{\alpha}^\top\mathbf{y}_t}{(\nu_0-2+\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t)^{3/2}}\frac{\partial(\mathrm{vec}(\boldsymbol{\Sigma}_t^{-1}))^\top}{\partial\mathrm{vech}(\boldsymbol{\Sigma}_t)}(\mathbf{y}_t\otimes\mathbf{y}_t) \\
&= -\frac{1}{2}\mathbf{D}_p^\top\mathrm{vec}(\boldsymbol{\Sigma}_t^{-1}) + \frac{1}{2}\frac{\sqrt{\nu_0+p}}{\sqrt{\nu_0-2}}\left(1+\frac{\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t}{\nu_0-2}\right)^{-1} \\
&\times \left(\frac{\sqrt{\nu_0+p}}{\sqrt{\nu_0-2}} + \frac{f_{T_{\nu_0+p}}\left(\frac{\boldsymbol{\alpha}^\top\mathbf{y}_t}{(\nu_0-2+\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t)^{1/2}}\sqrt{\nu_0+p}\right)}{F_{T_{\nu_0+p}}\left(\frac{\boldsymbol{\alpha}^\top\mathbf{y}_t}{(\nu_0-2+\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t)^{1/2}}\sqrt{\nu_0+p}\right)}\frac{\boldsymbol{\alpha}^\top\mathbf{y}_t/(\nu_0-2)}{\sqrt{1+\frac{\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t}{\nu_0-2}}}\right) \\
&\times \mathbf{D}_p^\top(\boldsymbol{\Sigma}_t^{-1}\otimes\boldsymbol{\Sigma}_t^{-1})\mathbf{D}_p^{+\top}\mathbf{D}_p^\top(\mathbf{y}_t\otimes\mathbf{y}_t) \\
&= -\frac{1}{2}\mathbf{D}_p^\top\mathrm{vec}(\boldsymbol{\Sigma}_t^{-1}) + \frac{1}{2}\frac{\sqrt{\nu_0+p}}{\sqrt{\nu_0-2}}\left(1+\frac{\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t}{\nu_0-2}\right)^{-1} \\
&\times \left(\frac{\sqrt{\nu_0+p}}{\sqrt{\nu_0-2}} + \frac{f_{T_{\nu_0+p}}\left(\frac{\boldsymbol{\alpha}^\top\mathbf{y}_t}{(\nu_0-2+\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t)^{1/2}}\sqrt{\nu_0+p}\right)}{F_{T_{\nu_0+p}}\left(\frac{\boldsymbol{\alpha}^\top\mathbf{y}_t}{(\nu_0-2+\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t)^{1/2}}\sqrt{\nu_0+p}\right)}\frac{\boldsymbol{\alpha}^\top\mathbf{y}_t/(\nu_0-2)}{\sqrt{1+\frac{\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t}{\nu_0-2}}}\right)\mathbf{D}_p^\top(\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t\otimes\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t),
\end{aligned}$$

where in the last equality we used the following properties:

(i) $\mathbf{D}_k\mathbf{D}_k^+ = \frac{1}{2}(\mathbf{I}_{k^2}+\mathbf{K}_k)$,

(ii) $\mathbf{K}_k = \mathbf{K}_k^\top = \mathbf{K}_k^{-1}$,

(iii) $\mathbf{K}_k(\mathbf{A} \otimes \mathbf{a}) = (\mathbf{a} \otimes \mathbf{A})$ for an arbitrary $k \times n$ matrix $\mathbf{A}$ and an arbitrary $k \times 1$ vector $\mathbf{a}$.

Now removing vech operators and denoting

$$\mathbf{Y}_t = \frac{1}{2}\boldsymbol{\Sigma}_t^{-1}\left(\mathbf{y}_t\mathbf{y}_t^\top \omega_t - \boldsymbol{\Sigma}_t\right)\boldsymbol{\Sigma}_t^{-1} \tag{2.18}$$

with

$$\omega_t = \frac{\sqrt{\nu_0+p}}{\sqrt{\nu_0-2}}\left(1+\frac{\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t}{\nu_0-2}\right)^{-1}\left(\frac{\sqrt{\nu_0+p}}{\sqrt{\nu_0-2}} + \frac{f_{T_{\nu_0+p}}\left(\frac{\boldsymbol{\alpha}^\top\mathbf{y}_t}{(\nu_0-2+\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t)^{1/2}}\sqrt{\nu_0+p}\right)}{F_{T_{\nu_0+p}}\left(\frac{\boldsymbol{\alpha}^\top\mathbf{y}_t}{(\nu_0-2+\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t)^{1/2}}\sqrt{\nu_0+p}\right)}\frac{\boldsymbol{\alpha}^\top\mathbf{y}_t/(\nu_0-2)}{\sqrt{1+\frac{\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t}{\nu_0-2}}}\right)$$

we obtain

$$\frac{\partial\mathcal{L}_{t;1}}{\partial\boldsymbol{\Sigma}_t} = 2\mathbf{Y}_t - \mathbf{Y}_t \circ \mathbf{I}\,.$$

3.1.2 Score w.r.t. $\boldsymbol{\alpha}_t$ (if needed and if we will decide to replace $\boldsymbol{\alpha}$ by $\boldsymbol{\alpha}_t$)

After replacing $\boldsymbol{\alpha}$ by $\boldsymbol{\alpha}_t$ we get

$$\begin{aligned}\frac{\partial\mathcal{L}_{t;1}}{\partial\boldsymbol{\alpha}_t} &= -\frac{1}{2}\frac{\partial\log|\boldsymbol{\Sigma}_t|}{\partial\boldsymbol{\alpha}_t} - \frac{\nu_0+p}{2}\frac{\partial\log\left(1+\frac{\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t}{\nu_0-2}\right)}{\partial\boldsymbol{\alpha}_t}\\ &+ \frac{\partial\log\left(F_{T_{\nu_0+p}}\left(\frac{\boldsymbol{\alpha}_t^\top\mathbf{y}_t}{(\nu_0-2+\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t)^{1/2}}\sqrt{\nu_0+p}\right)\right)}{\partial\boldsymbol{\alpha}_t}\\ &= \frac{f_{T_{\nu_0+p}}\left(\frac{\boldsymbol{\alpha}_t^\top\mathbf{y}_t}{(\nu_0-2+\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t)^{1/2}}\sqrt{\nu_0+p}\right)}{F_{T_{\nu_0+p}}\left(\frac{\boldsymbol{\alpha}_t^\top\mathbf{y}_t}{(\nu_0-2+\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t)^{1/2}}\sqrt{\nu_0+p}\right)}\frac{\sqrt{\nu_0+p}}{\sqrt{\nu_0-2+\mathbf{y}_t^\top\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t}}\mathbf{y}_t\end{aligned}$$

3.2 Noncentral-F Score

3.2.1 Score w.r.t. $\boldsymbol{\Sigma}_t$

For derivation of the noncentral matrix F score w.r.t. $\boldsymbol{\Sigma}_t$, p.d.f. representation (2.8) is used. The Log-Likelihood function w.r.t $\boldsymbol{\Sigma}$ is

$$\begin{aligned}\mathcal{L}_{t;2} = &\frac{\nu_1+2\nu_2}{2}\log\left(|\boldsymbol{\Sigma}_t|\right) - \frac{1}{2}\mathrm{tr}(\boldsymbol{\Sigma}_t^{-1}\boldsymbol{\Omega}_t)\\ &+\log\left(\int_{\mathbf{V}>\mathbf{O}}\exp\left(-\frac{1}{2}\mathrm{tr}((\boldsymbol{\Sigma}_t R C_t^{-1}\boldsymbol{\Sigma}_t + \frac{\nu_1}{\nu_2-p-1}\boldsymbol{\Sigma}_t)\mathbf{V})\right)|\mathbf{V}|^{\frac{\nu_1+\nu_2-p-1}{2}}\,{}_0F_1\left(\frac{\nu_1}{2};\frac{1}{4}\frac{\nu_1}{\nu_2-p-1}\boldsymbol{\Omega}_t\mathbf{V}\right)\mathrm{d}\mathbf{V}\right).\end{aligned} \tag{2.19}$$

Then the score obtains as

$$
\begin{aligned}
&\frac{\partial \mathcal{L}_{t;2}}{\partial \mathrm{vech}(\mathbf{\Sigma}_t)} \\
&= \left(\frac{\nu_1}{2}+\nu_2\right)\frac{\partial \log|\mathbf{\Sigma}_t|}{\partial \mathrm{vech}(\mathbf{\Sigma}_t)} - \frac{1}{2}\frac{\partial \mathrm{tr}\left(\mathbf{\Sigma}_t^{-1}\mathbf{\Omega}_t\right)}{\partial \mathrm{vech}(\mathbf{\Sigma}_t)} \\
&+ \frac{\partial \log\left(\int_{\mathbf{V}>\mathbf{O}} \exp\left(-\frac{1}{2}\mathrm{tr}((\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t + \frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t)\mathbf{V})\right)|\mathbf{V}|^{\frac{\nu_1+\nu_2-p-1}{2}} {}_0F_1\left(\frac{\nu_1}{2};\frac{1}{4}\frac{\nu_1}{\nu_2-p-1}\mathbf{\Omega}_t\mathbf{V}\right)\mathrm{d}\mathbf{V}\right)}{\partial \mathrm{vech}(\mathbf{\Sigma}_t)}
\end{aligned} \tag{2.20}
$$

$$
\begin{aligned}
&= \left(\frac{\nu_1}{2}+\nu_2\right)\mathbf{D}_p^\top \mathrm{vec}(\mathbf{\Sigma}_t^{-1}) + \frac{1}{2}\mathbf{D}_p^\top(\mathbf{\Sigma}_t^{-1}\otimes\mathbf{\Sigma}_t^{-1})\mathrm{vec}(\mathbf{\Omega}_t) \\
&+ \frac{1}{\int_{\mathbf{V}>\mathbf{O}} \exp\left(-\frac{1}{2}\mathrm{tr}((\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t + \frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t)\mathbf{V})\right)|\mathbf{V}|^{\frac{\nu_1+\nu_2-p-1}{2}} {}_0F_1\left(\frac{\nu_1}{2};\frac{1}{4}\frac{\nu_1}{\nu_2-p-1}\mathbf{\Omega}_t\mathbf{V}\right)\mathrm{d}\mathbf{V})} \\
&\times \frac{\partial\left(\int_{\mathbf{V}>\mathbf{O}} \exp\left(-\frac{1}{2}\mathrm{tr}((\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t + \frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t)\mathbf{V})\right)|\mathbf{V}|^{\frac{\nu_1+\nu_2-p-1}{2}} {}_0F_1\left(\frac{\nu_1}{2};\frac{1}{4}\frac{\nu_1}{\nu_2-p-1}\mathbf{\Omega}_t\mathbf{V}\right)\mathrm{d}\mathbf{V}\right)}{\partial \mathrm{vech}(\mathbf{\Sigma}_t)}
\end{aligned} \tag{2.21}
$$

$$
\begin{aligned}
&= \left(\frac{\nu_1}{2}+\nu_2\right)\mathbf{D}_p^\top \mathrm{vec}(\mathbf{\Sigma}_t^{-1}) + \frac{1}{2}\mathbf{D}_p^\top(\mathbf{\Sigma}_t^{-1}\otimes\mathbf{\Sigma}_t^{-1})\mathrm{vec}(\mathbf{\Omega}_t) \\
&+ \frac{1}{\int_{\mathbf{V}>\mathbf{O}} \exp\left(-\frac{1}{2}\mathrm{tr}((\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t + \frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t)\mathbf{V})\right)|\mathbf{V}|^{\frac{\nu_1+\nu_2-p-1}{2}} {}_0F_1\left(\frac{\nu_1}{2};\frac{1}{4}\frac{\nu_1}{\nu_2-p-1}\mathbf{\Omega}_t\mathbf{V}\right)\mathrm{d}\mathbf{V})} \\
&\times \int_{\mathbf{V}>\mathbf{O}} \frac{\partial \exp\left(-\frac{1}{2}\mathrm{tr}((\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t + \frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t)\mathbf{V})\right)}{\partial \mathrm{vech}(\mathbf{\Sigma}_t)}|\mathbf{V}|^{\frac{\nu_1+\nu_2-p-1}{2}} {}_0F_1\left(\frac{\nu_1}{2};\frac{1}{4}\frac{\nu_1}{\nu_2-p-1}\mathbf{\Omega}_t\mathbf{V}\right)\mathrm{d}\mathbf{V}
\end{aligned} \tag{2.22}
$$

$$
\begin{aligned}
&= \left(\frac{\nu_1}{2}+\nu_2\right)\mathbf{D}_p^\top \mathrm{vec}(\mathbf{\Sigma}_t^{-1}) + \frac{1}{2}\mathbf{D}_p^\top(\mathbf{\Sigma}_t^{-1}\otimes\mathbf{\Sigma}_t^{-1})\mathrm{vec}(\mathbf{\Omega}_t) \\
&- \frac{1}{2}\mathbf{D}_p^\top \frac{\int_{\mathbf{V}>\mathbf{O}} \mathbf{Z}(\mathbf{V})|\mathbf{V}|^{\frac{\nu_2}{2}} \exp\left(-\frac{1}{2}\mathrm{tr}((\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t + \frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t)\mathbf{V})\right)|\mathbf{V}|^{\frac{\nu_1-p-1}{2}} {}_0F_1\left(\frac{\nu_1}{2};\frac{1}{4}\frac{\nu_1}{\nu_2-p-1}\mathbf{\Omega}_t\mathbf{V}\right)\mathrm{d}\mathbf{V}}{\int_{\mathbf{V}>\mathbf{O}} |\mathbf{V}|^{\frac{\nu_2}{2}} \exp\left(-\frac{1}{2}\mathrm{tr}((\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t + \frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t)\mathbf{V})\right)|\mathbf{V}|^{\frac{\nu_1-p-1}{2}} {}_0F_1\left(\frac{\nu_1}{2};\frac{1}{4}\frac{\nu_1}{\nu_2-p-1}\mathbf{\Omega}_t\mathbf{V}\right)\mathrm{d}\mathbf{V}}
\end{aligned} \tag{2.23}
$$

$$
= \left(\frac{\nu_1}{2}+\nu_2\right)\mathbf{D}_p^\top \mathrm{vec}(\mathbf{\Sigma}_t^{-1}) + \frac{1}{2}\mathbf{D}_p^\top(\mathbf{\Sigma}_t^{-1}\otimes\mathbf{\Sigma}_t^{-1})\mathrm{vec}(\mathbf{\Omega}_t) - \frac{1}{2}\mathbf{D}_p^\top \mathbf{Z}\left(\frac{\mathrm{E}\left[\mathbf{V}|\mathbf{V}|^{\frac{\nu_2}{2}}\right]}{\mathrm{E}[|\mathbf{V}|^{\frac{\nu_2}{2}}]}\right) \tag{2.24}
$$

$$
\approx \left(\frac{\nu_1}{2}+\nu_2\right)\mathbf{D}_p^\top \mathrm{vec}(\mathbf{\Sigma}_t^{-1}) + \frac{1}{2}\mathbf{D}_p^\top(\mathbf{\Sigma}_t^{-1}\otimes\mathbf{\Sigma}_t^{-1})\mathrm{vec}(\mathbf{\Omega}_t) - \frac{1}{2}\mathbf{D}_p^\top \mathbf{Z}\left(\frac{(\bar{\mathbf{\Sigma}}+\frac{\bar{\mathbf{\Omega}}}{\nu_1})|\mathbf{I}+\frac{\bar{\mathbf{\Sigma}}^{-1}\bar{\mathbf{\Omega}}}{\nu_1}|^{\frac{\nu_2}{2}}(\nu_1+\nu_2)}{e^{\mathrm{tr}\left(-\frac{1}{2}\bar{\mathbf{\Sigma}}^{-1}\bar{\mathbf{\Omega}}\right)}{}_1F_1\left((\nu_1+\nu_2)/2,\,\frac{1}{2}\nu_1,\,\frac{1}{2}\bar{\mathbf{\Sigma}}^{-1}\bar{\mathbf{\Omega}}\right)}\right),
$$

see Appendix for last step

where in (2.23)

$$
\begin{aligned}
&\frac{\partial\left(-\frac{1}{2}\mathrm{tr}((\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t + \frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t)\mathbf{V})\right)}{\partial \mathrm{vech}(\mathbf{\Sigma}_t)} \\
&= -\frac{1}{2}\mathbf{D}_p^\top\left((\mathbf{V}\otimes RC_t^{-1})\mathrm{vec}(\mathbf{\Sigma}_t) + (RC_t^{-1}\otimes\mathbf{V})\mathrm{vec}(\mathbf{\Sigma}_t) + \frac{\nu_1}{\nu_2-p-1}\mathrm{vec}(\mathbf{V})\right) && (2.25) \\
&\equiv -\frac{1}{2}\mathbf{D}_p^\top \mathbf{Z}(\mathbf{V}) && (2.26)
\end{aligned}
$$

and the expectation in (2.24) are w.r.t. $\mathbf{V} \sim$ noncentral Wishart, since

$$
\exp\left(-\frac{1}{2}\mathrm{tr}((\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t + \frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t)\mathbf{V})\right)|\mathbf{V}|^{\frac{\nu_1-p-1}{2}} {}_0F_1\left(\frac{\nu_1}{2};\frac{1}{4}\frac{\nu_1}{\nu_2-p-1}\mathbf{\Omega}_t\mathbf{V}\right)
$$

is the kernel of the p.d.f. of the noncentral Wishart distribution with covariance matrix

$$
\bar{\mathbf{\Sigma}} = \left(\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t + \frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t\right)^{-1},
$$

non-centrality matrix

$$
\bar{\mathbf{\Omega}} = \left(\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t + \frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t\right)^{-1}\frac{\nu_1}{\nu_2-p-1}\mathbf{\Omega}_t\left(\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t + \frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t\right)^{-1}
$$

and

$$\mathbf{E}[\mathbf{V}] = \nu_1 \bar{\mathbf{\Sigma}} + \bar{\mathbf{\Omega}} \equiv \mathbf{Q}.$$

3.2.2 Score w.r.t. $\mathbf{\Omega}_t$

For derivation of the noncentral matrix F score w.r.t. $\mathbf{\Omega}_t$, p.d.f. representation (2.9) is used. The Log-Likelihood function w.r.t $\mathbf{\Omega}_t$ is

$$\begin{aligned}\mathcal{L}_{t;2} = & -\frac{\nu_1+\nu_2}{2}\log|\mathbf{\Omega}_t| - \frac{1}{2}\mathrm{tr}\left(\mathbf{\Sigma}_t^{-1}\mathbf{\Omega}_t\right) \\ & + \log\left(\int_{\mathbf{V}>\mathbf{O}} \exp\left(-\frac{1}{2}\mathrm{tr}(\mathbf{\Omega}_t^{-1}\mathbf{V})\right)|\mathbf{V}|^{\frac{\nu_1+\nu_2-p-1}{2}} {}_0F_1\left(\frac{\nu_1}{2};\frac{1}{4}\frac{\nu_1}{\nu_2-p-1}[\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t + \frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t]^{-1}\mathbf{V}\right)\mathrm{d}\mathbf{V}\right).\end{aligned} \tag{2.27}$$

Then the score obtains as

$$\begin{aligned}\frac{\partial\mathcal{L}_{t;2}}{\partial\mathrm{vech}(\mathbf{\Omega}_t)} \\ = & -\frac{\nu_1+\nu_2}{2}\frac{\partial\log|\mathbf{\Omega}_t|}{\partial\mathrm{vech}(\mathbf{\Omega}_t)} - \frac{1}{2}\frac{\partial\mathrm{tr}\left(\mathbf{\Sigma}_t^{-1}\mathbf{\Omega}_t\right)}{\partial\mathrm{vech}(\mathbf{\Omega}_t)} \\ & + \frac{\partial\log\left(\int_{\mathbf{V}>\mathbf{O}}\exp\left(-\frac{1}{2}\mathrm{tr}(\mathbf{\Omega}_t^{-1}\mathbf{V})\right)|\mathbf{V}|^{\frac{\nu_1+\nu_2-p-1}{2}}{}_0F_1\left(\frac{\nu_1}{2};\frac{1}{4}\frac{\nu_1}{\nu_2-p-1}[\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t+\frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t]^{-1}\mathbf{V}\right)\mathrm{d}\mathbf{V}\right)}{\partial\mathrm{vech}(\mathbf{\Omega}_t)}\end{aligned} \tag{2.28}$$

$$\begin{aligned} = & -\frac{\nu_1+\nu_2}{2}\mathbf{D}_p^\top\mathrm{vec}(\mathbf{\Omega}_t^{-1}) - \frac{1}{2}\mathbf{D}_p^\top\mathrm{vec}(\mathbf{\Sigma}_t^{-1}) \\ & + \frac{1}{\int_{\mathbf{V}>\mathbf{O}}\exp\left(-\frac{1}{2}\mathrm{tr}(\mathbf{\Omega}_t^{-1}\mathbf{V})\right)|\mathbf{V}|^{\frac{\nu_1+\nu_2-p-1}{2}}{}_0F_1\left(\frac{\nu_1}{2};\frac{1}{4}\frac{\nu_1}{\nu_2-p-1}[\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t+\frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t]^{-1}\mathbf{V}\right)\mathrm{d}\mathbf{V}} \\ & \times\frac{\partial\int_{\mathbf{V}>\mathbf{O}}\exp\left(-\frac{1}{2}\mathrm{tr}(\mathbf{\Omega}_t^{-1}\mathbf{V})\right)|\mathbf{V}|^{\frac{\nu_1+\nu_2-p-1}{2}}{}_0F_1\left(\frac{\nu_1}{2};\frac{1}{4}\frac{\nu_1}{\nu_2-p-1}[\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t+\frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t]^{-1}\mathbf{V}\right)\mathrm{d}\mathbf{V}}{\partial\mathrm{vech}(\mathbf{\Omega}_t)}\end{aligned} \tag{2.29}$$

$$\begin{aligned} = & -\frac{\nu_1+\nu_2}{2}\mathbf{D}_p^\top\mathrm{vec}(\mathbf{\Omega}_t^{-1}) - \frac{1}{2}\mathbf{D}_p^\top\mathrm{vec}(\mathbf{\Sigma}_t^{-1}) \\ & + \frac{\int_{\mathbf{V}>\mathbf{O}}\frac{\partial\exp\left(-\frac{1}{2}\mathrm{tr}(\mathbf{\Omega}_t^{-1}\mathbf{V})\right)}{\partial\mathrm{vech}(\mathbf{\Omega}_t)}|\mathbf{V}|^{\frac{\nu_1+\nu_2-p-1}{2}}{}_0F_1\left(\frac{\nu_1}{2};\frac{1}{4}\frac{\nu_1}{\nu_2-p-1}[\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t+\frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t]^{-1}\mathbf{V}\right)\mathrm{d}\mathbf{V}}{\int_{\mathbf{V}>\mathbf{O}}\exp\left(-\frac{1}{2}\mathrm{tr}(\mathbf{\Omega}_t^{-1}\mathbf{V})\right)|\mathbf{V}|^{\frac{\nu_1+\nu_2-p-1}{2}}{}_0F_1\left(\frac{\nu_1}{2};\frac{1}{4}\frac{\nu_1}{\nu_2-p-1}[\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t+\frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t]^{-1}\mathbf{V}\right)\mathrm{d}\mathbf{V}}\end{aligned} \tag{2.30}$$

$$\begin{aligned} = & -\frac{\nu_1+\nu_2}{2}\mathbf{D}_p^\top\mathrm{vec}(\mathbf{\Omega}_t^{-1}) - \frac{1}{2}\mathbf{D}_p^\top\mathrm{vec}(\mathbf{\Sigma}_t^{-1}) \\ & + \frac{1}{2}\mathbf{D}_p^\top(\mathbf{\Omega}_t^{-1}\otimes\mathbf{\Omega}_t^{-1}) \\ & \times\frac{\int_{\mathbf{V}>\mathbf{O}}\mathrm{vec}(\mathbf{V})|\mathbf{V}|^{\frac{\nu_2}{2}}\exp\left(-\frac{1}{2}\mathrm{tr}(\mathbf{\Omega}_t^{-1}\mathbf{V})\right)|\mathbf{V}|^{\frac{\nu_1-p-1}{2}}{}_0F_1\left(\frac{\nu_1}{2};\frac{1}{4}\frac{\nu_1}{\nu_2-p-1}[\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t+\frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t]^{-1}\mathbf{V}\right)\mathrm{d}\mathbf{V}}{\int_{\mathbf{V}>\mathbf{O}}|\mathbf{V}|^{\frac{\nu_2}{2}}\exp\left(-\frac{1}{2}\mathrm{tr}(\mathbf{\Omega}_t^{-1}\mathbf{V})\right)|\mathbf{V}|^{\frac{\nu_1-p-1}{2}}{}_0F_1\left(\frac{\nu_1}{2};\frac{1}{4}\frac{\nu_1}{\nu_2-p-1}[\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t+\frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t]^{-1}\mathbf{V}\right)\mathrm{d}\mathbf{V}}\end{aligned} \tag{2.31}$$

$$= -\frac{\nu_1+\nu_2}{2}\mathbf{D}_p^\top\mathrm{vec}(\mathbf{\Omega}_t^{-1}) - \frac{1}{2}\mathbf{D}_p^\top\mathrm{vec}(\mathbf{\Sigma}_t^{-1}) + \frac{1}{2}\mathbf{D}_p^\top(\mathbf{\Omega}_t^{-1}\otimes\mathbf{\Omega}_t^{-1})\mathrm{vec}\left(\frac{\mathrm{E}\left[\mathbf{V}|\mathbf{V}|^{\frac{\nu_2}{2}}\right]}{\mathrm{E}[|\mathbf{V}|^{\frac{\nu_2}{2}}]}\right) \tag{2.32}$$

$$\approx -\frac{\nu_1+\nu_2}{2}\mathbf{D}_p^\top\mathrm{vec}(\mathbf{\Omega}_t^{-1}) - \frac{1}{2}\mathbf{D}_p^\top\mathrm{vec}(\mathbf{\Sigma}_t^{-1}) + \frac{1}{2}\mathbf{D}_p^\top(\mathbf{\Omega}_t^{-1}\otimes\mathbf{\Omega}_t^{-1})\mathrm{vec}\left(\frac{(\bar{\mathbf{\Sigma}}+\frac{\bar{\mathbf{\Omega}}}{\nu_1})\left|\mathbf{I}+\frac{\bar{\mathbf{\Sigma}}^{-1}\bar{\mathbf{\Omega}}}{\nu_1}\right|^{\frac{\nu_2}{2}}(\nu_1+\nu_2)}{e^{\mathrm{tr}\left(-\frac{1}{2}\bar{\mathbf{\Sigma}}^{-1}\bar{\mathbf{\Omega}}\right)}{}_1F_1\left((\nu_1+\nu_2)/2,\frac{1}{2}\nu_1,\frac{1}{2}\bar{\mathbf{\Sigma}}^{-1}\bar{\mathbf{\Omega}}\right)}\right),$$

see Appendix for last step

where the expectation in (2.32) are w.r.t. $\mathbf{V}\sim$ noncentral Wishart, since

$$\exp\left(-\frac{1}{2}\mathrm{tr}(\mathbf{\Omega}_t^{-1}\mathbf{V})\right)|\mathbf{V}|^{\frac{\nu_1-p-1}{2}}{}_0F_1\left(\frac{\nu_1}{2};\frac{1}{4}\frac{\nu_1}{\nu_2-p-1}[\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t+\frac{\nu_1}{\nu_2-p-1}\mathbf{\Sigma}_t]^{-1}\mathbf{V}\right)$$

is the kernel of the p.d.f. of the noncentral Wishart distribution with covariance matrix

$$\bar{\mathbf{\Sigma}} = \mathbf{\Omega}_t,$$

non-centrality matrix

$$\bar{\mathbf{\Omega}} = \frac{\nu_1}{\nu_2 - p - 1}\mathbf{\Omega}_t[\mathbf{\Sigma}_t RC_t^{-1}\mathbf{\Sigma}_t + \frac{\nu_1}{\nu_2 - p - 1}\mathbf{\Sigma}_t]^{-1}\mathbf{\Omega}_t$$

and

$$\mathbf{E}[\mathbf{V}] = \nu_1\bar{\mathbf{\Sigma}} + \bar{\mathbf{\Omega}} \equiv \mathbf{Q}.$$

3.3 Total Score

Finally, the score matrix is calculated by

$$\nabla_{\mathbf{\Sigma}_t} = \frac{\partial \mathcal{L}_t}{\partial \mathbf{\Sigma}_t} = \frac{\partial \mathcal{L}_{t;1}}{\partial \mathbf{\Sigma}_t} + \frac{\partial \mathcal{L}_{t;2}}{\partial \mathbf{\Sigma}_t}\,, \tag{2.33}$$

and

$$\nabla_{\mathbf{\Omega}_t} = \frac{\partial \mathcal{L}_t}{\partial \mathbf{\Omega}_t} = \frac{\partial \mathcal{L}_{t;1}}{\partial \mathbf{\Omega}_t} + \frac{\partial \mathcal{L}_{t;2}}{\partial \mathbf{\Omega}_t}\,. \tag{2.34}$$

We then scale $\nabla_{\mathbf{\Sigma}_t}$ as in Opschoor et al.,

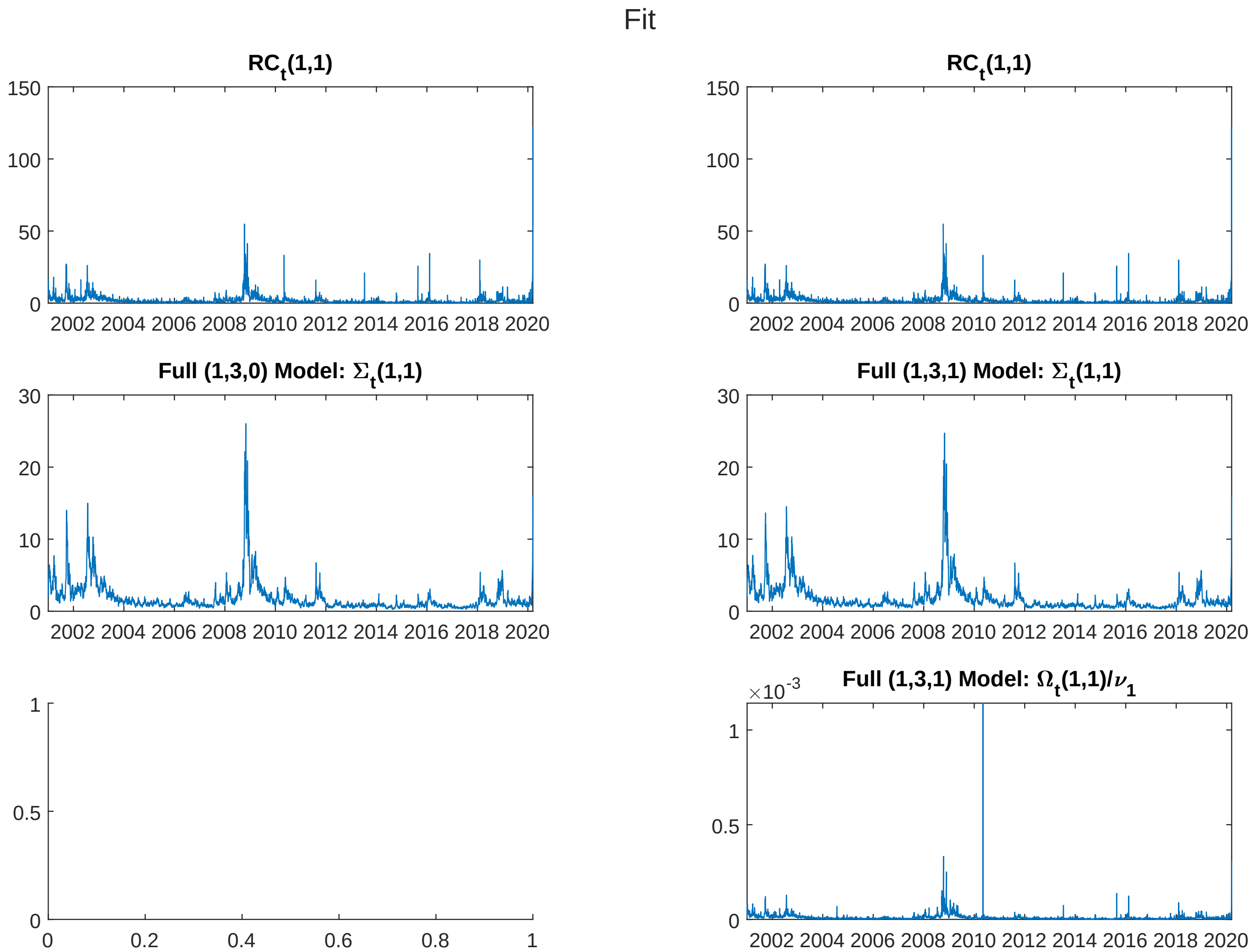
$$\mathbf{S}_{\mathbf{\Sigma}_t} = \nabla_{\mathbf{\Sigma}_t} \tag{2.35}$$

4 Empirical Application

$r = 0$ is the baseline Opschoor t-F model. Data is $p = 5$ (Assets: BA/HD/JPM/PFE/PG as in Opschoor Table 4 Panel A.1) from Jan 2001 to Mar 2020 for $T \approx 4500$.

(p, q, r)	$dim(\theta)$	$\mathcal{L}(\hat{\theta})$	BIC	$\sum_{i=1}^{q} \beta_i$	NaN	e^*_{11}	e^*_{21}	e^*_{31}	e^*_{41}	e^*_{51}	e^*_{22}	e^*_{32}	e^*_{42}	e^*_{52}	e^*_{33}	e^*_{43}	e^*_{53}	e^*_{44}	e^*_{54}	e^*_{55}
t-noncentral-spiked-F scalar (p,q,r) model. Estimation with intercept targeting.																				
$(1,1,0)$	20	-20788	41746	0.988																
$(1,1,1)$	25	-20712	41636	0.987																
$(1,1,2)$	30	-20669	41593	0.987																
$(1,2,0)$	21	-20519	41215	0.988																
$(1,2,1)$	26	-20451	41123	0.988																
$(1,2,2)$	31	-20403	41069	0.987																
$(1,3,0)$	22	-20400	40987	0.988																
$(1,3,1)$	27	-20334	40898	0.989																
$(1,3,2)$	32	-20289	40849	0.988																
$(2,1,0)$	21	-20430	41039	0.991																
$(2,1,1)$	26	-20364	40949	0.991																
$(2,1,2)$	31	-20317	40898	0.991																
$(2,2,0)$	22	-20113	40412	0.999																
$(2,2,1)$	27	-20031	40292	0.999																
$(2,2,2)$	32	-20007	40286	0.999																
$(2,3,0)$	23	-20062	40318	1																
$(2,3,1)$	28	-19981	40200	1																
$(2,3,2)$	33	-19952	40184	1																
$(3,1,0)$	22	-20320	40827	0.993																
$(3,1,1)$	27	-20252	40732	0.993																
$(3,1,2)$	32	-20210	40692	0.993																
$(3,2,0)$	23	-20055	40306	1																
$(3,2,1)$	28	-19975	40187	1																
$(3,2,2)$	33	-19945	40170	1																
$(3,3,0)$	24	-20055	40314	1																
$(3,3,1)$	29	-19975	40195	1																
$(3,3,2)$	34	-19942	40172	1																

t-noncentral-F full (p,q,r) model. Full Likelihood Optimization.


Fit
RC_t(1,1)
RC_t(1,1)
Full (1,3,0) Model: Σ_t(1,1)
Full (1,3,1) Model: Σ_t(1,1)
Full (1,3,1) Model: Ω_t(1,1)/ν_1
×10^-3
2002 2004 2006 2008 2010 2012 2014 2016 2018 2020

5 Appendix

Let's derive $\mathrm{E}\left[\mathbf{V}|\mathbf{V}|^h\right]$ if $\mathbf{V} \sim W(\mathbf{\Sigma}, \nu)$.

First, note that if $\mathbf{S} \sim W(\nu + 2h, \mathbf{\Sigma})$, then we know that

$$\mathrm{E}[\mathbf{S}] = \int\limits_{\mathbf{S}>\mathbf{0}} \mathbf{S} \frac{|\mathbf{S}|^{\frac{1}{2}((\nu+2h)-p-1)} e^{-\frac{1}{2}\mathrm{tr}((\mathbf{\Sigma})^{-1}\mathbf{S})}}{2^{((\nu+2h)p)/2}|\mathbf{\Sigma}|^{(\nu+2h)/2}\Gamma_p((\nu+2h)/2)} d\mathbf{S} = (\nu+2h)\mathbf{\Sigma}. \tag{2.36}$$

Thus

$$\int\limits_{\mathbf{S}>\mathbf{0}} \mathbf{S}|\mathbf{S}|^{\frac{1}{2}(\nu+2h-p-1)} e^{-\frac{1}{2}\mathrm{tr}((\mathbf{\Sigma})^{-1}\mathbf{S})} d\mathbf{S} = (\nu+2h)\mathbf{\Sigma} 2^{((\nu+2h)p)/2}|\mathbf{\Sigma}|^{(\nu+2h)/2}\Gamma_p((\nu+2h)/2). \tag{2.37}$$

Using (2.37) we can now derive $\mathrm{E}\left[\mathbf{V}|\mathbf{V}|^h\right]$ if $\mathbf{V} \sim W(\mathbf{\Sigma}, \nu)$,

$$\begin{aligned}
\mathrm{E}[\mathbf{V}|\mathbf{V}|^h] &= \int\limits_{\mathbf{V}>\mathbf{0}} \mathbf{V}|\mathbf{V}|^h \frac{|\mathbf{V}|^{\frac{1}{2}(\nu-p-1)} e^{-\frac{1}{2}\mathrm{tr}((\mathbf{\Sigma})^{-1}\mathbf{V})}}{2^{(\nu p)/2}|\mathbf{\Sigma}|^{\nu/2}\Gamma_p(\nu/2)} d\mathbf{V} && (2.38)\\
&= \frac{1}{2^{(\nu p)/2}|\mathbf{\Sigma}|^{\nu/2}\Gamma_p(\nu/2)} \int\limits_{\mathbf{V}>\mathbf{0}} \mathbf{V}|\mathbf{V}|^{\frac{1}{2}(\nu+2h-p-1)} e^{-\frac{1}{2}\mathrm{tr}((\mathbf{\Sigma})^{-1}\mathbf{V})} d\mathbf{V} && (2.39)\\
&\overset{2.37}{=} \frac{1}{2^{(\nu p)/2}|\mathbf{\Sigma}|^{\nu/2}\Gamma_p(\nu/2)} (\nu+2h)\mathbf{\Sigma} 2^{((\nu+2h)p)/2}|\mathbf{\Sigma}|^{(\nu+2h)/2}\Gamma_p((\nu+2h)/2) && (2.40)\\
&= \mathbf{\Sigma}|\mathbf{\Sigma}|^h 2^{hp}(\nu+2h)\frac{\Gamma_p((\nu+2h)/2)}{\Gamma_p(\nu/2)}. && (2.41)
\end{aligned}$$

Now, if $\mathbf{V} \sim$ noncentral-$W(\mathbf{\Sigma}, \mathbf{\Omega}, \nu)$

$$\frac{\mathrm{E}\left[\mathbf{V}|\mathbf{V}|^{\frac{\nu_2}{2}}\right]}{\mathrm{E}[|\mathbf{V}|^{\frac{\nu_2}{2}}]} = \frac{\mathrm{E}\left[\mathbf{V}|\mathbf{V}|^{\frac{\nu_2}{2}}\right]}{\frac{2^{(\nu_2/2)p}\Gamma_p((\nu_1+\nu_2)/2)|\mathbf{\Sigma}|^{\frac{\nu_2}{2}}}{\Gamma_p(\nu_1/2)} e^{\mathrm{tr}\left(-\frac{1}{2}\mathbf{\Sigma}^{-1}\mathbf{\Omega}_t\right)} {}_1F_1\left(\frac{1}{2}\nu_1 + \frac{\nu_2}{2}, \frac{1}{2}\nu_1, \frac{1}{2}\mathbf{\Sigma}^{-1}\mathbf{\Omega}_t\right)}$$

Gupta and Nagar (Theorem 3.5.6, p. 119))

$$\approx \frac{\left(\mathbf{\Sigma} + \frac{\mathbf{\Omega}}{\nu_1}\right)\left|\mathbf{\Sigma} + \frac{\mathbf{\Omega}}{\nu_1}\right|^{\frac{\nu_2}{2}} 2^{(\nu_2/2)p}(\nu_1+\nu_2)\frac{\Gamma_p((\nu_1+\nu_2)/2)}{\Gamma_p(\nu_1/2)}}{\frac{2^{(\nu_2/2)p}\Gamma_p((\nu_1+\nu_2)/2)|\mathbf{\Sigma}|^{\frac{\nu_2}{2}}}{\Gamma_p(\nu_1/2)} e^{\mathrm{tr}\left(-\frac{1}{2}\mathbf{\Sigma}^{-1}\mathbf{\Omega}_t\right)} {}_1F_1\left(\frac{1}{2}\nu_1 + \frac{\nu_2}{2}, \frac{1}{2}\nu_1, \frac{1}{2}\mathbf{\Sigma}^{-1}\mathbf{\Omega}_t\right)}$$

$\mathrm{E}\left[\mathbf{V}|\mathbf{V}|^{\frac{\nu_2}{2}}\right]$ is derived in (2.41) for $\mathbf{V} \sim W(\mathbf{\Sigma} + \mathbf{\Omega}/\nu, \nu) \approx$ noncentral-$W(\mathbf{\Sigma}, \mathbf{\Omega}, \nu)$

$$= \frac{(\mathbf{\Sigma} + \frac{\mathbf{\Omega}}{\nu_1})\left|\mathbf{I} + \frac{\mathbf{\Sigma}^{-1}\mathbf{\Omega}}{\nu_1}\right|^{\frac{\nu_2}{2}}(\nu_1+\nu_2)}{e^{\mathrm{tr}\left(-\frac{1}{2}\mathbf{\Sigma}^{-1}\mathbf{\Omega}\right)} {}_1F_1\left((\nu_1+\nu_2)/2, \frac{1}{2}\nu_1, \frac{1}{2}\mathbf{\Sigma}^{-1}\mathbf{\Omega}\right)}. \tag{2.42}$$