Pruned Skewed Kalman Filter and Smoother: With Application to the Yield Curve
Online Appendix

Gaygysyz Guljanov[a], Willi Mutschler[b], Mark Trede[a]

[a] *Center for Quantitative Economics, University of Münster.*

[b] *School of Business and Economics, University of Tübingen.*

Abstract

This the online appendix which shows the derivation of smoothing step of skewed Kalman filter.

1. Introduction and notations

Derivation of the smoothing step is a tedious task, also in terms of notations. In order to carry out this task as neatly as possible, we will proceed as follows. We first set notations and abbreviations. Afterwards, in a separate section, we show some useful derivations which we use later on. Then the derivation of the smoothing step proceeds with finding smoothed variables of periods $T-1$, $T-2$, $T-3$ and $T-4$, recursively. Here T denotes the last time period. From what we learn by deriving the above four periods manually, we devise general formulas for any time period. The more compact formulas for any time period is then given in the last section.

Now, let us start with with notations and abbreviations.

$$J_t = \Sigma_{t|t} G' \Sigma_{t+1|t}^{-1}$$
$$K_t = \Sigma_{t+1|t} F' (F \Sigma_{t+1|t} F' + \Sigma_\varepsilon)^{-1} (y_{t+1} - F \mu_{t+1|t}).$$

For any t, we know the prediction step distribution (see Rezaie & Eidsvik (2014))

$$x_{t+1|t} \sim CSN_{p, q_t + q_\eta}(\mu_{t+1|t}, \Sigma_{t+1|t}, \Gamma_{t+1|t}, \nu_{t+1|t}, \Delta_{t+1|t})$$

Email addresses: gaygysyz.guljanov@wiwi.uni-muenster.de (Gaygysyz Guljanov), willi@mutschler.eu (Willi Mutschler), mark.trede@uni-muenster.de (Mark Trede)

where

$$
\begin{aligned}
\mu_{t+1|t} &= G\mu_{t|t} + \mu_\eta \\
\Sigma_{t+1|t} &= G\Sigma_{t|t}G' + \Sigma_\eta \\
\Gamma_{t+1|t} &= \begin{pmatrix} \Gamma_{t|t}J_t \\ \Gamma_\eta \Sigma_\eta \Sigma_{t+1|t}^{-1} \end{pmatrix} \\
\nu_{t+1|t} &= \begin{pmatrix} \nu_{t|t} \\ \nu_\eta \end{pmatrix} \\
\Delta_{t+1|t} &= \begin{pmatrix} \Delta_{t|t} + \Gamma_{t|t}\Sigma_{t|t}\Gamma_{t|t}' - \Gamma_{t|t}J_t G\Sigma_{t|t}\Gamma_{t|t}' & \Gamma_{t|t}J_t\Sigma_\eta\Gamma_\eta' \\ \Gamma_\eta\Sigma_\eta J_t'\Gamma_{t|t}' & \Delta_\eta + \Gamma_\eta\Sigma_\eta\Gamma_\eta' - \Gamma_\eta\Sigma_\eta\Sigma_{t+1|t}^{-1}\Sigma_\eta\Gamma_\eta' \end{pmatrix}
\end{aligned}
$$

and the update step distribution (see Rezaie & Eidsvik (2014))

$$
x_{t+1|t+1} \sim CSN_{p,q_t}(\mu_{t+1|t+1}, \Sigma_{t+1|t+1}, \Gamma_{t+1|t+1}, \nu_{t+1|t+1}, \Delta_{t+1|t+1})
$$

where

$$
\begin{aligned}
\mu_{t+1|t+1} &= \mu_{t+1|t} + K_t \\
\Sigma_{t+1|t+1} &= \Sigma_{t+1|t} - \Sigma_{t+1|t}F'(F\Sigma_{t+1|t}F' + \Sigma_\varepsilon)^{-1}F\Sigma_{t+1|t} \\
\Gamma_{t+1|t+1} &= \Gamma_{t+1|t} \\
\nu_{t+1|t+1} &= \nu_{t+1|t} - \Gamma_{t+1|t}K_t \\
\Delta_{t+1|t+1} &= \Delta_{t+1|t}.
\end{aligned}
$$

2. Preliminary theorems and lemmas

Theorem 1.

$$
x_t|x_{t+1}, D_t \sim CSN(\mu_{dt}, \Sigma_{dt}, \Gamma_{dt}, \nu_{dt}, \Delta_{dt})
$$

with

$$
\begin{aligned}
\mu_{dt} &= \mu_{t|t} + \Sigma_{t|t} G' \Sigma_{t+1|t}^{-1} (x_{t+1} - G\mu_{t|t} - \mu_\eta) \\
\Sigma_{dt} &= \Sigma_{t|t} - \Sigma_{t|t} G' \Sigma_{t+1|t}^{-1} G \Sigma_{t|t} \\
\Gamma_{dt} &= \begin{pmatrix} \Gamma_{t|t} \\ -\Gamma_\eta G \end{pmatrix} \\
\nu_{dt} &= \begin{pmatrix} \nu_{t|t} \\ \nu_\eta \end{pmatrix} - \begin{pmatrix} \Gamma_{t|t}\Sigma_{t|t}G'\Sigma_{t+1|t}^{-1} \\ \Gamma_\eta - \Gamma_\eta G \Sigma_{t|t} G' \Sigma_{t+1|t}^{-1} \end{pmatrix} (x_{t+1} - G\mu_{t|t} - \mu_\eta) \\
&= \begin{pmatrix} \nu_{t|t} - \Gamma_{t|t}\Sigma_{t|t}G'\Sigma_{t+1|t}^{-1}(x_{t+1} - G\mu_{t|t} - \mu_\eta) \\ \nu_\eta - \Gamma_\eta + \Gamma_\eta G \Sigma_{t|t} G' \Sigma_{t+1|t}^{-1}(x_{t+1} - G\mu_{t|t} - \mu_\eta) \end{pmatrix} \\
\Delta_{dt} &= \Delta_c \\
&= \begin{pmatrix} \Delta_{t|t} & 0 \\ 0 & \Delta_\eta \end{pmatrix}.
\end{aligned}
$$

Proof. Our starting point is the equation

$$
\begin{pmatrix} x_{t|t} \\ x_{t+1|t} \end{pmatrix} = \begin{pmatrix} I \\ G \end{pmatrix} x_{t|t} + \begin{pmatrix} 0 \\ \eta_t \end{pmatrix}.
$$

which is equivalent to

$$
\begin{pmatrix} x_{t|t} \\ x_{t+1|t} \end{pmatrix} = \underbrace{\begin{pmatrix} I & 0 \\ G & I \end{pmatrix}}_{\equiv A} \begin{pmatrix} x_{t|t} \\ \eta_t \end{pmatrix}.
$$

Firstly, let us find the joint distribution of

$$
\begin{pmatrix} x_{t|t} \\ \eta_t \end{pmatrix} \sim \text{CSN}(\mu_j,\ \Sigma_j,\ \Gamma_j,\ \nu_j,\ \Delta_j)
$$

where,

$$\mu_j = \begin{bmatrix} \mu_{t|t} \\ \mu_\eta \end{bmatrix}$$

$$\Sigma_j = \begin{bmatrix} \Sigma_{t|t} & 0 \\ 0 & \Sigma_\eta \end{bmatrix}$$

$$\Gamma_j = \begin{bmatrix} \Gamma_{t|t} & 0 \\ 0 & \Gamma_\eta \end{bmatrix}$$

$$\nu_j = \begin{bmatrix} \nu_{t|t} \\ \nu_\eta \end{bmatrix}$$

$$\Delta_j = \begin{bmatrix} \Delta_{t|t} & 0 \\ 0 & \Delta_\eta \end{bmatrix}$$

Then, take the linear transformation of the above joint distribution with matrix A.

$$A \begin{pmatrix} x_{t|t} \\ \eta_t \end{pmatrix} \sim \text{CSN}(\mu_A,\ \Sigma_A,\ \Gamma_A,\ \nu_A,\ \Delta_A)$$

where,

$$\mu_A = A\mu_j = \begin{bmatrix} \mu_{t|t} \\ G\mu_{t|t} + \mu_\eta \end{bmatrix}$$

$$\Sigma_A = A\Sigma_j A' = \begin{bmatrix} \Sigma_{t|t} & \Sigma_{t|t}G' \\ G\Sigma_{t|t} & G\Sigma_{t|t}G' + \Sigma_\eta \end{bmatrix}$$

$$\Gamma_A = \Gamma_j A^{-1} = \begin{bmatrix} \Gamma_{t|t} & 0 \\ 0 & \Gamma_\eta \end{bmatrix} \begin{bmatrix} I & 0 \\ -G & I \end{bmatrix} = \begin{bmatrix} \Gamma_{t|t} & 0 \\ -\Gamma_\eta G & \Gamma_\eta \end{bmatrix}$$

$$\nu_A = \nu_j = \begin{bmatrix} \nu_{t|t} \\ \nu_\eta \end{bmatrix}$$

$$\Delta_A = \Delta_j = \begin{bmatrix} \Delta_{t|t} & 0 \\ 0 & \Delta_\eta \end{bmatrix}$$

Using the rules for conditional distributions (property 5 of the paper) we obtain

$$x_t | x_{t+1}, D_t \sim CSN(\mu_{dt}, \Sigma_{dt}, \Gamma_{dt}, \nu_{dt}, \Delta_{dt})$$

□

Using the notations above and some simplifications, we can rewrite out theorem as follows:

$$x_t|x_{t+1}, D_t \sim CSN(\mu_{dt}, \Sigma_{dt}, \Gamma_{dt}, \nu_{dt}, \Delta_{dt})$$

with

$$\begin{aligned}
\mu_{dt} &= \mu_{t|t} + J_t(x_{t+1} - \mu_{t+1|t}) \\
\Sigma_{dt} &= \Sigma_{t|t} - J_t G \Sigma_{t|t} \\
\Gamma_{dt} &= \begin{pmatrix} \Gamma_{t|t} \\ -\Gamma_\eta G \end{pmatrix} \\
\nu_{dt} &= \nu_{t+1|t} - \Gamma_{t+1|t}(x_{t+1} - \mu_{t+1|t}) \\
\Delta_{dt} &= \begin{pmatrix} \Delta_{t|t} & 0 \\ 0 & \Delta_\eta \end{pmatrix}.
\end{aligned}$$

Theorem 2. *The following equality holds,*

$$\Phi(\Gamma_{t+1|t+1}(x_{t+1} - \mu_{t+1|t+1}); \nu_{t+1|t+1}, \Delta_{t+1|t+1}) = \Phi(0; \nu_{dt}, \Delta_{dt} + \Gamma_{dt}\Sigma_{dt}\Gamma'_{dt})$$

or, equivalently,

$$\Phi(0; \nu_{t+1|t+1} - \Gamma_{t+1|t+1}(x_{t+1} - \mu_{t+1|t+1}), \Delta_{t+1|t+1}) = \Phi(0; \nu_{dt}, \Delta_{dt} + \Gamma_{dt}\Sigma_{dt}\Gamma'_{dt}).$$

Proof. We rewrite each expression in more basic terms. For the left hand side we obtain

$$\begin{aligned}
\Gamma_{t+1|t+1}(x_{t+1} - \mu_{t+1|t+1}) &= \Gamma_{t+1|t}(x_{t+1} - \mu_{t+1|t} - \Sigma_{t+1|t}F'(F\Sigma_{t+1|t}F' + \Sigma_\varepsilon)^{-1}(y_{t+1} - F\mu_{t+1|t})) \\
\nu_{t+1|t+1} &= \nu_{t+1|t} - \Gamma_{t+1|t}\Sigma_{t+1|t}F'(F\Sigma_{t+1|t}F' + \Sigma_\varepsilon)^{-1}(y_{t+1} - F\mu_{t+1|t}) \\
\Delta_{t+1|t+1} &= \begin{pmatrix} \Delta_{t|t} + \Gamma_{t|t}\Sigma_{t|t}\Gamma'_{t|t} - \Gamma_{t|t}\Sigma_{t|t}G'\Sigma^{-1}_{t+1|t}G\Sigma_{t|t}\Gamma'_{t|t} & -\Gamma_{t|t}\Sigma_{t|t}G'\Sigma^{-1}_{t+1|t}\Sigma_\eta\Gamma'_\eta \\ (-\Gamma_{t|t}\Sigma_{t|t}G'\Sigma^{-1}_{t+1|t}\Sigma_\eta\Gamma'_\eta)' & \Delta_\eta + \Gamma_\eta\Sigma_\eta\Gamma'_\eta - \Gamma_\eta\Sigma_\eta\Sigma^{-1}_{t+1|t}\Sigma_\eta\Gamma'_\eta \end{pmatrix}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\nu_{t+1|t+1} - \Gamma_{t+1|t+1}(x_{t+1} - \mu_{t+1|t+1}) &= \nu_{t+1|t} - \Gamma_{t+1|t}(x_{t+1} - \mu_{t+1|t}) \\
&= \begin{pmatrix} \nu_{t|t} \\ \nu_\eta \end{pmatrix} - \begin{pmatrix} \Gamma_{t|t}\Sigma_{t|t}G'\Sigma^{-1}_{t+1|t} \\ \Gamma_\eta\Sigma_\eta\Sigma^{-1}_{t+1|t} \end{pmatrix}(x_{t+1} - \mu_{t+1|t}) \\
&= \begin{pmatrix} \nu_{t|t} - \Gamma_{t|t}\Sigma_{t|t}G'\Sigma^{-1}_{t+1|t}(x_{t+1} - \mu_{t+1|t}) \\ \nu_\eta - \Gamma_\eta\Sigma_\eta\Sigma^{-1}_{t+1|t}(x_{t+1} - \mu_{t+1|t}) \end{pmatrix}
\end{aligned}$$

and for the right hand side we get

$$\nu_{dt} = \begin{pmatrix} \nu_{t|t} - \Gamma_{t|t}\Sigma_{t|t}G'\Sigma_{t+1|t}^{-1}(x_{t+1} - G\mu_{t|t} - \mu_\eta) \\ \nu_\eta - \Gamma_\eta + \Gamma_\eta G\Sigma_{t|t}G'\Sigma_{t+1|t}^{-1}(x_{t+1} - G\mu_{t|t} - \mu_\eta) \end{pmatrix}$$

$$\Delta_{dt} + \Gamma_{dt}\Sigma_{dt}\Gamma_{dt}' = \begin{pmatrix} \Delta_{t|t} & 0 \\ 0 & \Delta_\eta \end{pmatrix} + \begin{pmatrix} \Gamma_{t|t} \\ -\Gamma_\eta G \end{pmatrix} (\Sigma_{t|t} - \Sigma_{t|t}G'\Sigma_{t+1|t}^{-1}G\Sigma_{t|t}) \begin{pmatrix} \Gamma_{t|t} \\ -\Gamma_\eta G \end{pmatrix}'$$

Closer inspection shows that both expressions are the same.

□

Theorem 3.

$$\nu_{T|T}^{top} = \nu_{t+1|t} - \Gamma_{t+1|t}(\mu_{t+1|T} - \mu_{t+1|t}) \tag{1}$$

where $\nu_{T|T}^{top}$ *are the top* $t+2$ *elements of* $\nu_{T|T}$ *such that the dimensions fit.*

Proof. Note that

$$K_t = \mu_{t+1|t+1} - \mu_{t+1|t}$$

For period $T-1$: It can be easily seen that this equation holds:

$$\nu_{T|T} = \nu_{T|T-1} - \Gamma_{T|T-1}(\mu_{T|T} - \mu_{T|T-1})$$

For period $T-2$: We will start with equation above

$$\begin{aligned} \nu_{T|T}^{top} &= \nu_{T|T-1} - \Gamma_{T|T-1}(\mu_{T|T} - \mu_{T|T-1}) \\ &= \begin{pmatrix} \nu_{T-1|T-1} \\ \nu_\eta \end{pmatrix} - \begin{pmatrix} \Gamma_{T-1|T-1}J_{T-1} \\ \Gamma_\eta\Sigma_\eta\Sigma_{T|T-1}^{-1} \end{pmatrix} (\mu_{T|T} - \mu_{T|T-1}) \end{aligned}$$

We can now take the first row and the corresponding $\nu_{T|T}^{top}$:

$$\begin{aligned} \nu_{T|T}^{top} &= \nu_{T-1|T-1} - \Gamma_{T-1|T-1}J_{T-1}(\mu_{T|T} - \mu_{T|T-1}) \\ &= \nu_{T-1|T-2} - \Gamma_{T-1|T-2}K_{T-2} - \Gamma_{T-1|T-2}J_{T-1}(\mu_{T|T} - \mu_{T|T-1}) \\ &= \nu_{T-1|T-2} - \Gamma_{T-1|T-2}\Big[K_{T-2} + J_{T-1}(\mu_{T|T} - \mu_{T|T-1})\Big] \\ &= \nu_{T-1|T-2} - \Gamma_{T-1|T-2}\Big[\mu_{T-1|T-1} - \mu_{T-1|T-2} + J_{T-1}(\mu_{T|T} - \mu_{T|T-1})\Big] \\ &= \nu_{T-1|T-2} - \Gamma_{T-1|T-2}\Big[\mu_{T-1|T} - \mu_{T-1|T-2}\Big] \end{aligned}$$

For period $T-3$: We will start with equation above

$$\begin{aligned}
\nu_{T|T}^{top} &= \nu_{T-1|T-2} - \Gamma_{T-1|T-2}(\mu_{T-1|T} - \mu_{T-1|T-2}) \\
&= \begin{pmatrix} \nu_{T-2|T-2} \\ \nu_\eta \end{pmatrix} - \begin{pmatrix} \Gamma_{T-2|T-2} J_{T-2} \\ \Gamma_\eta \Sigma_\eta \Sigma_{T-1|T-2}^{-1} \end{pmatrix} (\mu_{T-1|T} - \mu_{T-1|T-2})
\end{aligned}$$

We can now take the first row and the corresponding $\nu_{T|T}^{top}$:

$$\begin{aligned}
\nu_{T|T}^{top} &= \nu_{T-2|T-2} - \Gamma_{T-2|T-2} J_{T-1}(\mu_{T-1|T} - \mu_{T-1|T-2}) \\
&= \nu_{T-2|T-3} - \Gamma_{T-2|T-3} K_{T-3} - \Gamma_{T-2|T-3} J_{T-2}(\mu_{T-1|T} - \mu_{T-1|T-2}) \\
&= \nu_{T-2|T-3} - \Gamma_{T-2|T-3}\Big[K_{T-3} + J_{T-2}(\mu_{T-1|T} - \mu_{T-1|T-2})\Big] \\
&= \nu_{T-2|T-3} - \Gamma_{T-2|T-3}\Big[\mu_{T-2|T-2} - \mu_{T-2|T-3} + J_{T-2}(\mu_{T-1|T} - \mu_{T-1|T-2})\Big] \\
&= \nu_{T-2|T-3} - \Gamma_{T-2|T-3}\Big[\mu_{T-2|T} - \mu_{T-2|T-3}\Big]
\end{aligned}$$

For period (any) t: We will assume the following holds for $t+1$:

$$\nu_{T|T}^{top} = \nu_{t+2|t+1} - \Gamma_{t+2|t+1}(\mu_{t+2|T} - \mu_{t+2|t+1})$$

We will start with equation above

$$\begin{aligned}
\nu_{T|T}^{top} &= \nu_{t+2|t+1} - \Gamma_{t+2|t+1}(\mu_{t+2|T} - \mu_{t+2|t+1}) \\
&= \begin{pmatrix} \nu_{t+1|t+1} \\ \nu_\eta \end{pmatrix} - \begin{pmatrix} \Gamma_{t+1|t+1} J_{t+1} \\ \Gamma_\eta \Sigma_\eta \Sigma_{t+2|t+1}^{-1} \end{pmatrix} (\mu_{t+2|T} - \mu_{t+2|t+1})
\end{aligned}$$

We can now take the first row and the corresponding $\nu_{T|T}^{top}$:

$$\begin{aligned}
\nu_{T|T}^{top} &= \nu_{t+1|t+1} - \Gamma_{t+1|t+1} J_{t+1}(\mu_{t+2|T} - \mu_{t+2|t+1}) \\
&= \nu_{t+1|t} - \Gamma_{t+1|t} K_t - \Gamma_{t+1|t} J_{t+1}(\mu_{t+2|T} - \mu_{t+2|t+1}) \\
&= \nu_{t+1|t} - \Gamma_{t+1|t}\Big[K_t + J_{t+1}(\mu_{t+2|T} - \mu_{t+2|t+1})\Big] \\
&= \nu_{t+1|t} - \Gamma_{t+1|t}\Big[\mu_{t+1|t+1} - \mu_{t+1|t} + J_{t+1}(\mu_{t+2|T} - \mu_{t+2|t+1})\Big] \\
&= \nu_{t+1|t} - \Gamma_{t+1|t}\Big[\mu_{t+1|T} - \mu_{t+1|t}\Big]
\end{aligned}$$

$\square$

Lemma 1. *For any t, the following holds,*

$$\Phi\Bigg[\Gamma_{dt}(x_t-\mu_{dt});\nu_{dt},\Delta_{dt}\Bigg]=\Phi\Bigg[\begin{pmatrix}\Gamma_{t|t} & 0\\ -\Gamma_\eta G & \Gamma_\eta\end{pmatrix}\left(\begin{pmatrix}x_t\\ x_{t+1}\end{pmatrix}-\begin{pmatrix}\mu_{t|T}\\ \mu_{t+1|T}\end{pmatrix}\right); \quad \nu_{t+1|t}-\Gamma_{t+1|t}(\mu_{t+1|T}-\mu_{t+1|t}), \quad \begin{pmatrix}\Delta_{t|t} & 0\\ 0 & \Delta_\eta\end{pmatrix}\Bigg] \tag{2}$$

where,

$$\mu_{t|T}\equiv\mu_{t|t}+\Sigma_{t|t}G'\Sigma^{-1}_{t+1|t}(\mu_{t+1|T}-\mu_{t+1|t})$$

which will coincide with first parameter of the distribution of $X_{t|T}$.

Proof. Note that $\Gamma_\eta G J_t-\Gamma_\eta=-\Gamma_\eta\Sigma_\eta\Sigma^{-1}_{t+1|t}$.

$$\Phi\Bigg[\Gamma_{dt}(x_t-\mu_{dt});\nu_{dt},\Delta_{dt}\Bigg]$$

$$=\Phi\left[\begin{pmatrix}\Gamma_{t|t}\\ -\Gamma_\eta G\end{pmatrix}(x_t-\mu_{t|t})-\begin{pmatrix}\Gamma_{t|t}J_t\\ -\Gamma_\eta G J_t\end{pmatrix}(x_{t+1}-\mu_{t+1|t});\nu_{t+1|t}-\begin{pmatrix}\Gamma_{t|t}J_t\\ \Gamma_\eta\Sigma_\eta\Sigma^{-1}_{t+1|t}\end{pmatrix}(x_{t+1}-\mu_{t+1|t}),\begin{pmatrix}\Delta_{t|t} & 0\\ 0 & \Delta_\eta\end{pmatrix}\right]$$

$$=\Phi\Bigg[\Gamma_{t|t}(x_t-\mu_{t|t});\nu_{t|t},\Delta_{t|t}\Bigg]\times\Phi\Bigg[-\Gamma_\eta G(x_t-\mu_{t|t})+\Gamma_\eta(x_{t+1}-\mu_{t+1|t});\nu_\eta,\Delta_\eta\Bigg]$$

Let us expand these cdfs

$$\Phi\Bigg[\Gamma_{t|t}(x_t-\mu_{t|t});\nu_{t|t},\Delta_{t|t}\Bigg]\times\Phi\Bigg[-\Gamma_\eta G(x_t-\mu_{t|t})+\Gamma_\eta(x_{t+1}-\mu_{t+1|t});\nu_\eta,\Delta_\eta\Bigg]$$

$$=\Phi\Bigg[\Gamma_{t|t}(x_t-\mu_{t|t}-J_t(\mu_{t+1|T}-\mu_{t+1|t}));\ \nu_{t|t}-\Gamma_{t|t}J_t(\mu_{t+1|T}-\mu_{t+1|t}),\ \Delta_{t|t}\Bigg]$$

$$\times\Phi\Bigg[-\Gamma_\eta G(x_t-\mu_{t|t}-J_t(\mu_{t+1|T}-\mu_{t+1|t})+J_t(\mu_{t+1|T}-\mu_{t+1|t}))+\Gamma_\eta(x_{t+1}-\mu_{t+1|T}+\mu_{t+1|T}-\mu_{t+1|t});\ \nu_\eta,\ \Delta_\eta\Bigg]$$

$$=\Phi\Bigg[\Gamma_{t|t}(x_t-\mu_{t|T});\ \nu_{t|t}-\Gamma_{t|t}J_t(\mu_{t+1|T}-\mu_{t+1|t}),\ \Delta_{t|t}\Bigg]$$

$$\times\Phi\Bigg[-\Gamma_\eta G(x_t-\mu_{t|T})+\Gamma_\eta(x_{t+1}-\mu_{t+1|T})-\Gamma_\eta G J_t(\mu_{t+1|T}-\mu_{t+1|t})+\Gamma_\eta(\mu_{t+1|T}-\mu_{t+1|t});\ \nu_\eta,\ \Delta_\eta\Bigg]$$

$$=\Phi\Bigg[\Gamma_{t|t}(x_t-\mu_{t|T});\ \nu_{t|t}-\Gamma_{t|t}J_t(\mu_{t+1|T}-\mu_{t+1|t}),\ \Delta_{t|t}\Bigg]$$

$$\times\Phi\Bigg[-\Gamma_\eta G(x_t-\mu_{t|T})+\Gamma_\eta(x_{t+1}-\mu_{t+1|T});\ \nu_\eta+(\Gamma_\eta G J_t-\Gamma_\eta)(\mu_{t+1|T}-\mu_{t+1|t}),\ \Delta_\eta\Bigg]$$

$$=\Phi\left[\begin{pmatrix}\Gamma_{t|t} & 0\\ -\Gamma_\eta G & \Gamma_\eta\end{pmatrix}\left(\begin{pmatrix}x_t\\ x_{t+1}\end{pmatrix}-\begin{pmatrix}\mu_{t|T}\\ \mu_{t+1|T}\end{pmatrix}\right);\ \nu_{t+1|t}-\Gamma_{t+1|t}(\mu_{t+1|T}-\mu_{t+1|t}),\ \begin{pmatrix}\Delta_{t|t} & 0\\ 0 & \Delta_\eta\end{pmatrix}\right]$$

$\square$

Lemma 2. *As to the product of two normal pdfs, we obtain for any t*

$$\phi\Big[x_t;\ \mu_{dt},\ \Sigma_{dt}\Big] \times \phi\Big[x_{t+1};\ \mu_{t+1|T},\ \Sigma_{t+1|T}\Big] = \phi\Bigg[\begin{pmatrix} x_t \\ x_{t+1} \end{pmatrix};\ \begin{pmatrix} \mu_{t|t} + J_t(\mu_{t+1|T} - \mu_{t+1|t}) \\ \mu_{t+1|T} \end{pmatrix},\ \begin{pmatrix} \Sigma_{t|t} + J_t(\Sigma_{t+1|T} - \Sigma_{t+1|t})J_t' & J_t\Sigma_{t+1|T} \\ \Sigma_{t+1|T}J_t' & \Sigma_{t+1|T} \end{pmatrix}\Bigg] \tag{3}$$

where,

$$\mu_{t|T} \equiv \mu_{t|t} + \Sigma_{t|t}G'\Sigma_{t+1|t}^{-1}(\mu_{t+1|T} - \mu_{t+1|t})$$
$$\Sigma_{t|T} \equiv \Sigma_{t|t} + \Sigma_{t|t}G'\Sigma_{t+1|t}^{-1}(\Sigma_{t+1|T} - \Sigma_{t+1|t})\Sigma_{t+1|t}^{-1}G\Sigma_{t|t}$$

which will coincide with the first two parameters of the distribution of $X_{t|T}$.

Proof. Recall the conditioning theorem for Normal distribution. Let $X \sim N_p(\psi, \Omega)$ be partitioned into X_1 of length p_1 and X_2 of length p_2, such that $X = (X_1', X_2')'$. The parameters are partitioned accordingly,

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \quad \Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix},$$

Then,

$$X_{1|2} = (X_1|X_2 = x_2) \sim N_{p_1}(\psi_{1|2}, \Omega_{1|2})$$

with $\psi_{1|2} = \psi_1 + \Omega_{12}\Omega_{22}^{-1}(x_2 - \psi_2)$ and, $\Omega_{1|2} = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}$.

Now, let us write left hand side of equation 3 in more basic terms and add/subtract some of its expressions,

$$
\begin{aligned}
&\phi\Big[x_t;\ \mu_{t|t} + J_t(x_{t+1} - \mu_{t+1|t}),\ \Sigma_{t|t} - J_t G \Sigma_{t|t}\Big] \times \phi\Big[x_{t+1};\ \mu_{t+1|T},\ \Sigma_{t+1|T}\Big] \\
&= \phi\Big[x_t; \\
&\qquad \mu_{t|t} + J_t \Sigma_{t+1|t} \Sigma_{t+1|t}^{-1} (x_{t+1} - \mu_{t+1|T} + \mu_{t+1|T} - \mu_{t+1|t}), \\
&\qquad \Sigma_{t|t} - J_t \Sigma_{t+1|t} \underbrace{\Sigma_{t+1|t}^{-1} G \Sigma_{t|t}}_{=J_t'} + J_t \Sigma_{t+1|T} \Sigma_{t+1|T}^{-1} \Sigma_{t+1|T} J_t' - \underbrace{J_t \Sigma_{t+1|T}}_{\text{analogous to } \Omega_{12}} \Sigma_{t+1|T}^{-1} \Sigma_{t+1|T} J_t' \Big] \\
&\times \phi\Big[x_{t+1};\ \mu_{t+1|T},\ \Sigma_{t+1|T}\Big] \\
&= \phi\Big[x_t; \\
&\qquad \mu_{t|t} + J_t(\mu_{t+1|T} - \mu_{t+1|t}) + \underbrace{J_t \Sigma_{t+1|t}}_{\text{analogous to } \Omega_{12}} \underbrace{\Sigma_{t+1|t}^{-1}}_{\Omega_{22}^{-1}} (x_{t+1} - \mu_{t+1|T}), \\
&\qquad \Sigma_{t|t} - J_t \Sigma_{t+1|t} J_t' + J_t \Sigma_{t+1|T} J_t' - J_t \Sigma_{t+1|T} \Sigma_{t+1|T}^{-1} \Sigma_{t+1|T} J_t' \Big] \\
&\times \phi\Big[x_{t+1};\ \mu_{t+1|T},\ \Sigma_{t+1|T}\Big]
\end{aligned}
$$

At this point, it is trivial to get the equation of lemma, if we use the reverse of the conditioning theorem for Normal distribution.

□

Lemma 3. *The equality below holds,*

$$\Delta_{dt} + \Gamma_{dt} \Sigma_{dt} \Gamma_{dt}' = \Delta_{t+1|t+1}. \tag{4}$$

Proof. We will write $\Delta_{dt} + \Gamma_{dt}\Sigma_{dt}\Gamma'_{dt}$ in more basic terms and evaluate:

$$\begin{aligned}
\Delta_{dt} + \Gamma_{dt}\Sigma_{dt}\Gamma'_{dt} &= \begin{pmatrix} \Delta_{t|t} & 0 \\ 0 & \Delta_\eta \end{pmatrix} + \begin{pmatrix} \Gamma_{t|t} \\ -\Gamma_\eta G \end{pmatrix} (\Sigma_{t|t} - J_t G \Sigma_{t|t}) \begin{pmatrix} \Gamma'_{t|t} & -G'\Gamma'_\eta \end{pmatrix} \\
&= \begin{pmatrix} \Delta_{t|t} & 0 \\ 0 & \Delta_\eta \end{pmatrix} + \begin{pmatrix} \Gamma_{t|t}(\Sigma_{t|t} - J_t G\Sigma_{t|t})\Gamma'_{t|t} & -\Gamma_{t|t}(\Sigma_{t|t} - J_t G\Sigma_{t|t})G'\Gamma'_\eta \\ -\Gamma_\eta G(\Sigma_{t|t} - J_t G\Sigma_{t|t})\Gamma'_{t|t} & \Gamma_\eta G(\Sigma_{t|t} - J_t G\Sigma_{t|t})G'\Gamma'_\eta \end{pmatrix} \\
&= \begin{pmatrix} \Delta_{t|t} + \Gamma_{t|t}\Sigma_{t|t}\Gamma'_{t|t} - \Gamma_{t|t}J_t G\Sigma_{t|t}\Gamma'_{t|t} & -\Gamma_{t|t}\Sigma_{t|t}G'\Gamma'_\eta + \Gamma_{t|t}\Sigma_{t|t}G'\Sigma^{-1}_{t+1|t} \underbrace{G\Sigma_{t|t}G'}_{=\Sigma_{t+1|t}-\Sigma_\eta} \Gamma'_\eta \\ -\Gamma_\eta G\Sigma_{t|t}\Gamma'_{t|t} + \Gamma_\eta \underbrace{G\Sigma_{t|t}G'}_{=\Sigma_{t+1|t}-\Sigma_\eta} \Sigma^{-1}_{t+1|t} G\Sigma_{t|t}\Gamma'_{t|t} & \Delta_\eta + \Gamma_\eta \underbrace{G\Sigma_{t|t}G'}_{=\Sigma_{t+1|t}-\Sigma_\eta} \Gamma'_\eta - \Gamma_\eta \underbrace{G\Sigma_{t|t}G'}_{=\Sigma_{t+1|t}-\Sigma_\eta} \Sigma^{-1}_{t+1|t} \underbrace{G\Sigma_{t|t}G'}_{=\Sigma_{t+1|t}-\Sigma_\eta} \Gamma'_\eta \end{pmatrix} \\
&= \begin{pmatrix} \Delta_{t|t} + \Gamma_{t|t}\Sigma_{t|t}\Gamma'_{t|t} - \Gamma_{t|t}J_t G\Sigma_{t|t}\Gamma'_{t|t} & \Gamma_{t|t}J_t\Sigma_\eta\Gamma'_\eta \\ \Gamma_\eta\Sigma_\eta J'_t\Gamma'_{t|t} & \Delta_\eta + \Gamma_\eta\Sigma_\eta\Gamma'_\eta - \Gamma_\eta\Sigma_\eta\Sigma^{-1}_{t+1|t}\Sigma_\eta\Gamma'_\eta \end{pmatrix} \\
&= \Delta_{t+1|t} \\
&= \Delta_{t+1|t+1}
\end{aligned}$$

□

3. Smoothing formulas for period T-1

In the final period, $x_{T|T}$ is both the filtered and the smoothed distribution. This is where the backward recursion kicks in.

First, consider the penultimate period $T-1$. The conditional density of $x_{T-1}|x_T, D_{T-1}$ is (see theorem 1)

$$\frac{\Phi(\Gamma_{dT-1}(x_{T-1} - \mu_{dT-1}); \nu_{dT-1}, \Delta_{dT-1})}{\Phi(0; \nu_{dT-1}, \Delta_{dT-1} + \Gamma_{dT-1}\Sigma_{dT-1}\Gamma'_{dT-1})} \phi(x_{T-1}; \mu_{dT-1}, \Sigma_{dT-1})$$

where both μ_{dT-1} and ν_{dT-1} are functions of x_T. To find the distribution of $x_{T-1}|D_T$ we average the density of $x_{T-1}|x_T, D_{T-1}$ over $x_T|D_T$,

$$\begin{aligned}
\int \frac{\Phi(\Gamma_{dT-1}(x_{T-1} - \mu_{dT-1}); \nu_{dT-1}, \Delta_{dT-1})}{\Phi(0; \nu_{dT-1}, \Delta_{dT-1} + \Gamma_{dT-1}\Sigma_{dT-1}\Gamma'_{dT-1})} \frac{\Phi(\Gamma_{T|T}(x_T - \mu_{T|T}); \nu_{T|T}, \Delta_{T|T})}{\Phi(0; \nu_{T|T}, \Delta_{T|T} + \Gamma_{T|T}\Sigma_{T|T}\Gamma'_{T|T})} \\
\times \phi(x_{T-1}; \mu_{dT-1}, \Sigma_{dT-1})\phi(x_T; \mu_{T|T}, \Sigma_{T|T}) dx_T.
\end{aligned}$$

Due to theorem 2, the top right cdf and the bottom left cdf are identical,

$$\Phi(\Gamma_{T|T}(x_T - \mu_{T|T}); \nu_{T|T}, \Delta_{T|T}) = \Phi(0; \nu_{dT-1}, \Delta_{dT-1} + \Gamma_{dT-1}\Sigma_{dT-1}\Gamma'_{dT-1}).$$

Having cancelled the cdfs, the integral simplifies to

$$\int \frac{\Phi(\Gamma_{dT-1}(x_{T-1}-\mu_{dT-1});\nu_{dT-1},\Delta_{dT-1})}{\Phi(0;\nu_{T|T},\Delta_{T|T}+\Gamma_{T|T}\Sigma_{T|T}\Gamma'_{T|T})} \times \phi(x_{T-1};\mu_{dT-1},\Sigma_{dT-1})\phi(x_T;\mu_{T|T},\Sigma_{T|T})dx_T.$$

We first look at the joint distribution of x_{T-1} and x_T (given D_T) and then marginalize. Let μ_{jT-1}, Σ_{jT-1}, Γ_{jT-1}, ν_{jT-1}, Δ_{jT-1} denote the parameters of the joint distribution. Obviously, (after using lemma 2)

$$\mu_{jT-1} = \begin{pmatrix} \mu_{T-1|T-1} + J_{T-1}(\mu_{T|T} - \mu_{T|T-1}) \\ \mu_{T|T} \end{pmatrix}$$
$$\Sigma_{jT-1} = \begin{pmatrix} \Sigma_{T-1|T-1} + J_{T-1}(\Sigma_{T|T} - \Sigma_{T|T-1})J'_{T-1} & J_{T-1}\Sigma_{T|T} \\ \Sigma_{T|T}J'_{T-1} & \Sigma_{T|T} \end{pmatrix}$$

Less obvious, but still relatively straightforward (use the lemma 1 and the lemma 3),

$$\Gamma_{jT-1} = \begin{pmatrix} \Gamma_{T-1|T-1} & 0 \\ -\Gamma_\eta G & \Gamma_\eta \end{pmatrix}$$
$$\nu_{jT-1} = \nu_{T|T}$$
$$\Delta_{jT-1} = \begin{pmatrix} \Delta_{T-1|T-1} & 0 \\ 0 & \Delta_\eta \end{pmatrix}.$$

The marginal distribution of $x_{T-1}|D_T$ is (obtained using lemma 2.3.1 of Genton (2004))

$$\mu_{T-1|T} = \mu_{T-1|T-1} + J_{T-1}(\mu_{T|T} - \mu_{T|T-1})$$
$$\Sigma_{T-1|T} = \Sigma_{T-1|T-1} + J_{T-1}(\Sigma_{T|T} - \Sigma_{T|T-1})J'_{T-1}$$
$$\Gamma_{T-1|T} = \begin{pmatrix} \Gamma_{T-1|T-1} \\ N_{T-1} \end{pmatrix}$$
$$\nu_{T-1|T} = \nu_{T|T}$$
$$\Delta_{T-1|T} = \begin{pmatrix} \Delta_{T-1|T-1} & 0 \\ 0 & \tilde{\Delta}_{T-1} \end{pmatrix}$$

where

$$N_{T-1} = -\Gamma_\eta G + \Gamma_\eta M_{T-1}$$
$$M_{T-1} = \Sigma_{T|T}J'_{T-1}\Sigma^{-1}_{T-1|T}$$
$$L_{T-1} = \Sigma_{T|T} - M_{T-1}\Sigma_{T-1|T}M'_{T-1}$$

and

$$\tilde{\Delta}_{T-1} = \Delta_\eta + \Gamma_\eta L_{T-1}\Gamma'_\eta.$$

4. Smoothing formulas for period T-2

Now consider period $T-2$. The conditional density of $x_{T-2}|x_{T-1}, D_{T-2}$ is

$$\frac{\Phi(\Gamma_{dT-2}(x_{T-2}-\mu_{dT-2});\nu_{dT-2},\Delta_{dT-2})}{\Phi(0;\nu_{dT-2},\Delta_{dT-2}+\Gamma_{dT-2}\Sigma_{dT-2}\Gamma'_{dT-2})}\phi(x_{T-2};\mu_{dT-2},\Sigma_{dT-2})$$

where μ_{dT-2} and ν_{dT-2} are functions of x_{T-1}. To find the distribution of $x_{T-2}|D_T$ we again derive the joint distribution of x_{T-2} and x_{T-1} given D_T,

$$\frac{\Phi(\Gamma_{dT-2}(x_{T-2}-\mu_{dT-2});\nu_{dT-2},\Delta_{dT-2})}{\Phi(0;\nu_{dT-2},\Delta_{dT-2}+\Gamma_{dT-2}\Sigma_{dT-2}\Gamma'_{dT-2})}\frac{\Phi(\Gamma_{T-1|T}(x_{T-1}-\mu_{T-1|T});\nu_{T-1|T},\Delta_{T-1|T})}{\Phi(0;\nu_{T-1|T},\Delta_{T-1|T}+\Gamma_{T-1|T}\Sigma_{T-1|T}\Gamma'_{T-1|T})}$$
$$\times\,\phi(x_{T-2};\mu_{dT-2},\Sigma_{dT-2})\phi(x_{T-1};\mu_{T-1|T},\Sigma_{T-1|T}).$$

According to theorem 3 and lemma 1, for $t=T-2$, the top left cdf can be rewritten as

$$\begin{aligned}&\Phi(\Gamma_{dT-2}(x_{T-2}-\mu_{dT-2});\nu_{dT-2},\Delta_{dT-2})\\ &=\Phi\left(\begin{pmatrix}\Gamma_{T-2|T-2} & 0\\ -\Gamma_\eta G & \Gamma_\eta\end{pmatrix}\left(\begin{pmatrix}x_{T-2}\\ x_{T-1}\end{pmatrix}-\begin{pmatrix}\mu_{T-2|T}\\ \mu_{T-1|T}\end{pmatrix}\right);\nu^{top}_{T|T},\begin{pmatrix}\Delta_{T-2|T-2} & 0\\ 0 & \Delta_\eta\end{pmatrix}\right)\end{aligned}$$

The top right cdf can be factorized as

$$\begin{aligned}\Phi(\Gamma_{T-1|T}(x_{T-1}-\mu_{T-1|T});\nu_{T-1|T},\Delta_{T-1|T}) &= \Phi\left(\begin{pmatrix}\Gamma_{T-1|T-1}\\ N_{T-1}\end{pmatrix}(x_{T-1}-\mu_{T-1|T});\nu_{T|T},\begin{pmatrix}\Delta_{T-1|T-1} & 0\\ 0 & \tilde{\Delta}_{T-1}\end{pmatrix}\right)\\ &=\Phi(\Gamma_{T-1|T-1}(x_{T-1}-\mu_{T-1|T});\nu^{top}_{T|T},\Delta_{T-1|T-1}) \qquad (5)\\ &\quad\times\Phi(N_{T-1}(x_{T-1}-\mu_{T-1|T});\nu^{btm}_{T|T},\tilde{\Delta}_{T-1}) \qquad (6)\end{aligned}$$

where $\nu_{T|T}$ is suitably partitioned into $\nu^{top}_{T|T}$ and $\nu^{btm}_{T|T}$.

Next, turn to the bottom left cdf. Using theorem 2 and theorem 3 it can be shown that

$$\begin{aligned}&\Phi(0;\nu_{dT-2},\Delta_{dT-2}+\Gamma_{dT-2}\Sigma_{dT-2}\Gamma'_{dT-2})\\ &=\Phi(\Gamma_{T-1|T-1}(x_{T-1}-\mu_{T-1|T-1});\nu_{T-1|T-1},\Delta_{T-1|T-1})\\ &=\Phi(\Gamma_{T-1|T-1}(x_{T-1}-\mu_{T-1|T}+\mu_{T-1|T}-\mu_{T-1|T-1});\nu_{T-1|T-1},\Delta_{T-1|T-1})\\ &=\Phi(\Gamma_{T-1|T-1}(x_{T-1}-\mu_{T-1|T});\nu_{T-1|T-1}-\Gamma_{T-1|T-1}(\mu_{T-1|T}-\mu_{T-1|T-1}),\Delta_{T-1|T-1})\end{aligned}$$

and

$$\begin{aligned}&\nu_{T-1|T-1}-\Gamma_{T-1|T-1}(\mu_{T-1|T}-\mu_{T-1|T-1})\\ &=\nu_{T-1|T-2}-\Gamma_{T-1|T-2}K_{T-2}-\Gamma_{T-1|T-2}(\mu_{T-1|T}-\mu_{T-1|T-2}-K_{T-2})\\ &=\nu_{T-1|T-2}-\Gamma_{T-1|T-2}(\mu_{T-1|T}-\mu_{T-1|T-2})\\ &=\nu^{top}_{T|T}\end{aligned}$$

As a result, we have

$$\begin{aligned}&\Phi(0;\nu_{dT-2},\Delta_{dT-2}+\Gamma_{dT-2}\Sigma_{dT-2}\Gamma'_{dT-2})\\&=\Phi(\Gamma_{T-1|T-1}(x_{T-1}-\mu_{T-1|T});\nu^{top}_{T|T},\Delta_{T-1|T-1})\end{aligned}$$

Therefore, the first part of equation 5 and bottom left cdf cancel each other out.

The remaining part of the top right cdf can be merged into the top left cdf. This results in the numerator

$$\Phi\left(\begin{pmatrix}\Gamma_{T-2|T-2} & 0\\ -\Gamma_\eta G & \Gamma_\eta\\ 0 & N_{T-1}\end{pmatrix}\begin{pmatrix}x_{T-2}-\mu_{T-2|T}\\ x_{T-1}-\mu_{T-1|T}\end{pmatrix};\nu_{T|T},\begin{pmatrix}\Delta_{T-2|T-2} & 0 & 0\\ 0 & \Delta_\eta & 0\\ 0 & 0 & \tilde{\Delta}_{T-1}\end{pmatrix}\right)$$

Let μ_{jT-2} etc denote the parameters of the joint distribution,

$$\begin{aligned}\mu_{jT-2}&=\begin{pmatrix}\mu_{T-2|T-2}+J_{T-2}(\mu_{T-1|T}-\mu_{T-1|T-2})\\ \mu_{T-1|T}\end{pmatrix}\\ \Sigma_{jT-2}&=\begin{pmatrix}\Sigma_{T-2|T-2}+J_{T-2}(\Sigma_{T-1|T}-\Sigma_{T-1|T-2})J'_{T-2} & J_{T-2}\Sigma_{T-1|T}\\ \Sigma_{T-1|T}J'_{T-2} & \Sigma_{T-1|T}\end{pmatrix}\\ \Gamma_{jT-2}&=\begin{pmatrix}\Gamma_{T-2|T-2} & 0\\ -\Gamma_\eta G & \Gamma_\eta\\ 0 & N_{T-1}\end{pmatrix}\\ \nu_{jT-2}&=\nu_{T|T}\\ \Delta_{jT-2}&=\begin{pmatrix}\Delta_{T-2|T-2} & 0 & 0\\ 0 & \Delta_\eta & 0\\ 0 & 0 & \tilde{\Delta}_{T-1}\end{pmatrix}.\end{aligned}$$

The marginal distribution of $x_{T-2}|D_T$ is

$$\begin{aligned}\mu_{T-2|T}&=\mu_{T-2|T-2}+J_{T-2}(\mu_{T-1|T}-\mu_{T-1|T-2})\\ \Sigma_{T-2|T}&=\Sigma_{T-2|T-2}+J_{T-2}(\Sigma_{T-1|T}-\Sigma_{T-1|T-2})J'_{T-2}\\ \Gamma_{T-2|T}&=\begin{pmatrix}\Gamma_{T-2|T-2}\\ N_{T-2}\\ N_{T-1}M_{T-2}\end{pmatrix}\\ \nu_{T-2|T}&=\nu_{T|T}\\ \Delta_{T-2|T}&=\begin{pmatrix}\Delta_{T-2|T-2} & 0\\ 0 & \tilde{\Delta}_{T-2}\end{pmatrix}.\end{aligned}$$

with

$$
\begin{aligned}
N_{T-2} &= -\Gamma_\eta G + \Gamma_\eta M_{T-2} \\
M_{T-2} &= \Sigma_{T-1|T} J'_{T-2} \Sigma^{-1}_{T-2|T} \\
L_{T-2} &= \Sigma_{T-1|T} - M_{T-2}\Sigma_{T-2|T} M'_{T-2}
\end{aligned}
$$

and

$$
\tilde{\Delta}_{T-2} = \begin{pmatrix} \Delta_\eta & 0 \\ 0 & \tilde{\Delta}_{T-1} \end{pmatrix} + \begin{pmatrix} \Gamma_\eta \\ N_{T-1} \end{pmatrix} L_{T-2} \begin{pmatrix} \Gamma_\eta \\ N_{T-1} \end{pmatrix}'.
$$

5. Smoothing formulas for period T-3

Consider period $T-3$. The conditional density of $x_{T-3}|x_{T-2}, D_{T-3}$ is

$$
\frac{\Phi(\Gamma_{dT-3}(x_{T-3} - \mu_{dT-3}); \nu_{dT-3}, \Delta_{dT-3})}{\Phi(0; \nu_{dT-3}, \Delta_{dT-3} + \Gamma_{dT-3}\Sigma_{dT-3}\Gamma'_{dT-3})} \phi(x_{T-3}; \mu_{dT-3}, \Sigma_{dT-3})
$$

where μ_{dT-3} and ν_{dT-3} are functions of x_{T-2}. To find the distribution of $x_{T-3}|D_T$ we again derive the joint distribution of x_{T-3} and x_{T-2} given D_T,

$$
\begin{aligned}
&\frac{\Phi(\Gamma_{dT-3}(x_{T-3} - \mu_{dT-3}); \nu_{dT-3}, \Delta_{dT-3})}{\Phi(0; \nu_{dT-3}, \Delta_{dT-3} + \Gamma_{dT-3}\Sigma_{dT-3}\Gamma'_{dT-3})} \frac{\Phi(\Gamma_{T-2|T}(x_{T-2} - \mu_{T-2|T}); \nu_{T-2|T}, \Delta_{T-2|T})}{\Phi(0; \nu_{T-2|T}, \Delta_{T-2|T} + \Gamma_{T-2|T}\Sigma_{T-2|T}\Gamma'_{T-2|T})} \\
&\qquad \times \phi(x_{T-3}; \mu_{dT-3}, \Sigma_{dT-3})\phi(x_{T-2}; \mu_{T-2|T}, \Sigma_{T-2|T}).
\end{aligned}
$$

According to theorem 3 and lemma 1 for $t = T-3$, the top left cdf can be rewritten as

$$
\begin{aligned}
&\Phi(\Gamma_{dT-3}(x_{T-3} - \mu_{dT-3}); \nu_{dT-3}, \Delta_{dT-3}) \\
&= \Phi\left(\begin{pmatrix} \Gamma_{T-3|T-3} & 0 \\ -\Gamma_\eta G & \Gamma_\eta \end{pmatrix} \left(\begin{pmatrix} x_{T-3} \\ x_{T-2} \end{pmatrix} - \begin{pmatrix} \mu_{T-3|T} \\ \mu_{T-2|T} \end{pmatrix} \right); \nu^{top}_{T|T}, \begin{pmatrix} \Delta_{T-3|T-3} & 0 \\ 0 & \Delta_\eta \end{pmatrix} \right)
\end{aligned}
$$

The top right cdf can be factorized as

$$
\begin{aligned}
\Phi(\Gamma_{T-2|T}(x_{T-2} - \mu_{T-2|T}); \nu_{T-2|T}, \Delta_{T-2|T}) &= \Phi\left(\begin{pmatrix} \Gamma_{T-2|T-2} \\ N_{T-2} \\ N_{T-1}M_{T-2} \end{pmatrix} (x_{T-2} - \mu_{T-2|T}); \nu_{T|T}, \begin{pmatrix} \Delta_{T-2|T-2} & 0 \\ 0 & \tilde{\Delta}_{T-2} \end{pmatrix} \right) \\
&= \Phi(\Gamma_{T-2|T-2}(x_{T-2} - \mu_{T-2|T}); \nu^{top}_{T|T}, \Delta_{T-2|T-2}) \\
&\quad \times \Phi\left(\begin{pmatrix} N_{T-2} \\ N_{T-1}M_{T-2} \end{pmatrix} (x_{T-2} - \mu_{T-2|T}); \nu^{btm}_{T|T}, \tilde{\Delta}_{T-2} \right)
\end{aligned}
$$

where $\nu_{T|T}$ is suitably partitioned into $\nu^{top}_{T|T}$ and $\nu^{btm}_{T|T}$.

Next, turn to the bottom left cdf. Using theorem 2 and theorem 3, it can be shown that as we did in the last section

$$\begin{aligned}&\Phi(0; \nu_{dT-3}, \Delta_{dT-3} + \Gamma_{dT-3}\Sigma_{dT-3}\Gamma'_{dT-3})\\&= \Phi(\Gamma_{T-2|T-2}(x_{T-2} - \mu_{T-2|T}); \nu^{top}_{T|T}, \Delta_{T-2|T-2})\end{aligned}$$

Therefore, the first part of factorization of top right cdf and bottom left cdf cancel each other out. The remaining part of the top right cdf can be merged into the top left cdf. This results in the numerator

$$\Phi\left(\begin{pmatrix}\Gamma_{T-3|T-3} & 0\\ -\Gamma_\eta G & \Gamma_\eta\\ 0 & N_{T-2}\\ 0 & N_{T-1}M_{T-2}\end{pmatrix}\begin{pmatrix}x_{T-3} - \mu_{T-3|T}\\ x_{T-2} - \mu_{T-2|T}\end{pmatrix}; \nu_{T|T}, \begin{pmatrix}\Delta_{T-3|T-3} & 0 & 0\\ 0 & \Delta_\eta & 0\\ 0 & 0 & \tilde{\Delta}_{T-2}\end{pmatrix}\right)$$

Let μ_{jT-3} etc denote the parameters of the joint distribution,

$$\begin{aligned}\mu_{jT-3} &= \begin{pmatrix}\mu_{T-3|T-3} + J_{T-3}(\mu_{T-2|T} - \mu_{T-2|T-3})\\ \mu_{T-2|T}\end{pmatrix}\\ \Sigma_{jT-3} &= \begin{pmatrix}\Sigma_{T-3|T-3} + J_{T-3}(\Sigma_{T-2|T} - \Sigma_{T-2|T-3})J'_{T-3} & J_{T-3}\Sigma_{T-2|T}\\ \Sigma_{T-2|T}J'_{T-3} & \Sigma_{T-2|T}\end{pmatrix}\\ \Gamma_{jT-3} &= \begin{pmatrix}\Gamma_{T-3|T-3} & 0\\ -\Gamma_\eta G & \Gamma_\eta\\ 0 & N_{T-2}\\ 0 & N_{T-1}M_{T-2}\end{pmatrix}\\ \nu_{jT-3} &= \nu_{T|T}\\ \Delta_{jT-3} &= \begin{pmatrix}\Delta_{T-3|T-3} & 0 & 0\\ 0 & \Delta_\eta & 0\\ 0 & 0 & \tilde{\Delta}_{T-2}\end{pmatrix}.\end{aligned}$$

The marginal distribution of $x_{T-3}|D_T$ is

$$\mu_{T-3|T} = \mu_{T-3|T-3} + J_{T-3}(\mu_{T-2|T} - \mu_{T-2|T-3})$$
$$\Sigma_{T-3|T} = \Sigma_{T-3|T-3} + J_{T-3}(\Sigma_{T-2|T} - \Sigma_{T-2|T-3})J'_{T-3}$$
$$\Gamma_{T-3|T} = \begin{pmatrix} \Gamma_{T-3|T-3} \\ N_{T-3} \\ N_{T-2}M_{T-3} \\ N_{T-1}M_{T-2}M_{T-3} \end{pmatrix}$$
$$\nu_{T-3|T} = \nu_{T|T}$$
$$\Delta_{T-3|T} = \begin{pmatrix} \Delta_{T-3|T-3} & 0 \\ 0 & \tilde{\Delta}_{T-3} \end{pmatrix}.$$

with

$$\tilde{\Delta}_{T-3} = \begin{pmatrix} \Delta_\eta & 0 \\ 0 & \tilde{\Delta}_{T-2} \end{pmatrix} + \begin{pmatrix} \Gamma_\eta \\ N_{T-2} \\ N_{T-1}M_{T-2} \end{pmatrix} L_{T-3} \begin{pmatrix} \Gamma_\eta \\ N_{T-2} \\ N_{T-1}M_{T-2} \end{pmatrix}'$$

(note that the dimension of $\tilde{\Delta}_{T-2}$ is just large enough to make the sum fit).

6. Smoothing formulas for period T-4

Guess for CSN parameters of $x_{T-4}|D_T$:

$$\mu_{T-4|T} = \mu_{T-4|T-4} + J_{T-4}(\mu_{T-3|T} - \mu_{T-3|T-4})$$
$$\Sigma_{T-4|T} = \Sigma_{T-4|T-4} + J_{T-4}(\Sigma_{T-3|T} - \Sigma_{T-3|T-4})J'_{T-4}$$
$$\Gamma_{T-4|T} = \begin{pmatrix} \Gamma_{T-4|T-4} \\ N_{T-4} \\ N_{T-3}M_{T-4} \\ N_{T-2}M_{T-3}M_{T-4} \\ N_{T-1}M_{T-2}M_{T-3}M_{T-4} \end{pmatrix}$$
$$\nu_{T-4|T} = \nu_{T|T}$$
$$\Delta_{T-4|T} = \begin{pmatrix} \Delta_{T-4|T-4} & 0 \\ 0 & \tilde{\Delta}_{T-4} \end{pmatrix}.$$

with

$$\tilde{\Delta}_{T-4} = \begin{pmatrix} \Delta_\eta & 0 \\ 0 & \tilde{\Delta}_{T-3} \end{pmatrix} + \begin{pmatrix} \Gamma_\eta \\ N_{T-3} \\ N_{T-2}M_{T-3} \\ N_{T-1}M_{T-2}M_{T-3} \end{pmatrix} L_{T-4} \begin{pmatrix} \Gamma_\eta \\ N_{T-3} \\ N_{T-2}M_{T-3} \\ N_{T-1}M_{T-2}M_{T-3} \end{pmatrix}'$$

7. Smoothing formulas for any time period

The CSN parameters for $x_t|D_T$ are

$$\begin{aligned}
\mu_{t|T} &= \mu_{t|t} + J_t(\mu_{t+1|T} - \mu_{t+1|t}) \\
\Sigma_{t|T} &= \Sigma_{t|t} + J_t(\Sigma_{t+1|T} - \Sigma_{t+1|t})J_t' \\
\Gamma_{t|T} &= \begin{pmatrix} \Gamma_{t|t} \\ N_t \\ N_{t+1}M_t \\ N_{t+2}M_{t+1}M_t \\ \vdots \\ N_{T-1} \cdot \ldots \cdot M_{t+2}M_{t+1}M_t \end{pmatrix} \\
\nu_{t|T} &= \nu_{T|T} \\
\Delta_{t|T} &= \begin{pmatrix} \Delta_{t|t} & 0 \\ 0 & \tilde{\Delta}_t \end{pmatrix}.
\end{aligned}$$

with

$$\tilde{\Delta}_t = \begin{pmatrix} \Delta_\eta & 0 \\ 0 & \tilde{\Delta}_{t+1} \end{pmatrix} + \begin{pmatrix} \Gamma_\eta \\ N_{t+1} \\ N_{t+2}M_{t+1} \\ N_{T-1} \cdot \ldots \cdot M_{t+2}M_{t+1} \end{pmatrix} L_t \begin{pmatrix} \Gamma_\eta \\ N_{t+1} \\ N_{t+2}M_{t+1} \\ N_{T-1} \cdot \ldots \cdot M_{t+2}M_{t+1} \end{pmatrix}'$$

with

$$\begin{aligned}
J_t &= \Sigma_{t|t}G'\Sigma_{t+1|t}^{-1} \\
M_t &= \Sigma_{t+1|T}J_t'\Sigma_{t|T}^{-1} \\
N_t &= -\Gamma_\eta G + \Gamma_\eta M_t \\
L_t &= \Sigma_{t+1|T} - M_t\Sigma_{t|T}M_t'
\end{aligned}$$

8. More compact formulas for any time period

We can write the formulas for general t more neatly via the following steps:

- Replace J_t with $\Sigma_{t|t} G' \Sigma_{t+1|t}^{-1}$.
- Replace L_t with $\Sigma_{t+1|T} - M_t \Sigma_{t|T} M_t'$
- Define $O_t \equiv \begin{bmatrix} N_t \\ O_{t+1} M_t \end{bmatrix}$ with $O_{T-1} \equiv N_{T-1}$ (Note that O_T is not defined)

After implementing the above steps, it is easy to see that we get:

$$\begin{aligned}
\mu_{t|T} &= \mu_{t|t} + \Sigma_{t|t} G' \Sigma_{t+1|t}^{-1} (\mu_{t+1|T} - \mu_{t+1|t}) \\
\Sigma_{t|T} &= \Sigma_{t|t} + \Sigma_{t|t} G' \Sigma_{t+1|t}^{-1} (\Sigma_{t+1|T} - \Sigma_{t+1|t}) \Sigma_{t+1|t}^{-1} G \Sigma_{t|t} \\
\Gamma_{t|T} &= \begin{pmatrix} \Gamma_{t|t} \\ O_t \end{pmatrix} \\
\nu_{t|T} &= \nu_{T|T} \\
\Delta_{t|T} &= \begin{pmatrix} \Delta_{t|t} & 0 \\ 0 & \tilde{\Delta}_t \end{pmatrix}.
\end{aligned}$$

with

$$\tilde{\Delta}_t = \begin{pmatrix} \Delta_\eta & 0 \\ 0 & \tilde{\Delta}_{t+1} \end{pmatrix} + \begin{pmatrix} \Gamma_\eta \\ O_{t+1} \end{pmatrix} (\Sigma_{t+1|T} - M_t \Sigma_{t|T} M_t') \begin{pmatrix} \Gamma_\eta \\ O_{t+1} \end{pmatrix}'$$

References

Genton, M. G. (2004). *Skew-Elliptical Distributions and Their Applications - A Journey Beyond Normality*. S.l.: CRC PRESS.

Rezaie, J., & Eidsvik, J. (2014). Kalman filter variants in the closed skew normal setting. *Computational Statistics & Data Analysis*, *75*, 1–14. doi:10.1016/j.csda.2014.01.014.