GRIPS Discussion Paper 23-7

Exact Likelihood for Inverse Gamma Stochastic Volatility Models

By

Roberto Leon-Gonzalez
Blessings Majoni

July 2023

GRIPS
NATIONAL GRADUATE INSTITUTE FOR POLICY STUDIES

National Graduate Institute for Policy Studies
7-22-1 Roppongi, Minato-ku,
Tokyo, Japan 106-8677

Exact Likelihood for Inverse Gamma Stochastic Volatility Models[1]

Roberto Leon-Gonzalez[a,b], **Blessings Majoni**[a]

[a]National Graduate Institute for Policy Studies (GRIPS), Japan

[b] The Rimini Centre for Economic Analysis (RCEA)

July, 2023

[1]The authors thank participants of the 24th International Conference on Computational Statistics (Bologna) and the 16th International Conference on Computational and Financial Econometrics (London) for helpful comments and suggestions. They also thank the Japan Society for the Promotion of Science (JSPS) for financial support under grant 19K01588. Roberto Leon-Gonzalez is a senior fellow of The Rimini Centre for Economic Analysis.

Abstract

We obtain a novel analytic expression of the likelihood for a stationary inverse gamma Stochastic Volatility (SV) model. This allows us to obtain the Maximum Likelihood Estimator for this non linear non gaussian state space model. Further, we obtain both the filtering and smoothing distributions for the inverse volatilities as mixture of gammas and therefore we can provide the smoothed estimates of the volatility. We show that by integrating out the volatilities the model that we obtain has the resemblance of a GARCH in the sense that the formulas are similar, which simplifies computations significantly. The model allows for fat tails in the observed data. We provide empirical applications using exchange rates data for 7 currencies and quarterly inflation data for four countries. We find that the empirical fit of our proposed model is overall better than alternative models for 4 countries currency data and for 2 countries inflation data.

Keywords: Hypergeometric Function, Particle Filter, Parallel Computing, Euler Acceleration.

JEL: C32, C58

MSC: 62M10, 62F99, 62J99, 60J05

1 Introduction

For most non-linear or non-Gaussian state space models it is difficult to obtain the likelihood function in closed form. This prevents the use of Maximum Likelihood Estimation (MLE). As a result most studies use Bayesian estimation with Markov Chain Monte Carlo (MCMC) methods. Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models are simpler to estimate than Stochastic Volatility (SV) models, because the likelihood for a GARCH model can be easily calculated in closed form (e.g. Engle (1982), Bollerslev (1987)). However, SV models have often been found to outperform GARCH models in empirical studies for both macroeconomic and financial data (e.g. Chan & Grant (2016) and Kim et al. (1998)). In addition, unlike GARCH models, SV models provide not only filtered estimates but also smoothed estimates of the volatility.

Although in linear Gaussian state space models the likelihood is available in closed form and can easily be calculated with the Kalman Filter algorithm (e.g. Durbin and Koopman (2012)), few studies have attempted to obtain a closed form expression for the likelihood in nonlinear non-Gaussian state space models. Shephard (1994) obtains a closed form expression for the likelihood of a non-stationary SV model known as Local Scale Model, showing the similarities to GARCH models. Uhlig (1997) builds on and generalizes Shephard (1994) to the multivariate case. They obtain an analytic expression for the likelihood and posterior density of a SV non-stationary restricted singular Wishart model. Creal (2017) obtains an analytic expression for the likelihood in a SV gamma model and shows that analytic expressions for the likelihood could also be obtained for a family of non linear non Gaussian state space models. The gamma SV model in Creal (2017) implies a variance-gamma distribution for the data and this distribution has thin tails (Madan & Seneta, 1990). In contrast, inverse gamma SV models imply a student-t distribution, thus, they can account for the fat tails that are observed in most macroeconomic and financial data (Leon-Gonzalez, 2019).

The purpose of this study is to obtain an analytic expression of the likelihood for the inverse gamma SV model. This exact likelihood solution will allow the estimation of the parameters and unobserved states for this non linear and non gaussian state space model by MLE. Without the likelihood expression, estimation of non linear non gaussian state space models generally involves bayesian methods such as Markov Chain Monte Carlo. We show that by marginalising out the volatilities, the model that we obtain has the resemblance of a GARCH in the sense that the formulas that we get are similar, which simplifies computations significantly. Moreover, the likelihood function proposed in this paper can be calculated efficiently using a simple recursion. The calculations can be accelerated by doing computations in parallel, as well as by applying Euler or other acceleration techniques to the Gauss hypergeometric functions in the likelihood. In addition to obtaining the exact likelihood, we obtain analytically the expressions for the smoothed and filtered estimates of the volatilities. We provide the computer code to calculate the likelihood as a user-friendly R package.

Section 2 reviews the literature on previous attempts to obtain analytically the likelihood expressions for non linear non gaussian state space models. Section 3 describes our model and derives the analytic expression of the likelihood. In addition the section provides the analytic expressions for the filtering and smoothing distributions of the volatilities. Section 4 evaluates the empirical performance and computational efficiency of the proposed novel algorithm with a comparison to other methods. We provide empirical applications using exchange rates data for 7 currencies to the US dollar and quarterly inflation data for four countries. Section 5 concludes.

2 Literature Review

2.1 Stochastic Volatility Models with an Exact Likelihood

There are very few non linear non gaussian state space models for which the likelihood can be obtained exactly. In what follows we review some of the SV models for which an analytic expression of the likelihood has been obtained.

To obtain the maximum likelihood estimates for a generalised non stationary local scale model, Shephard (1994) uses the conjugacy between the gamma and the beta distribution. Using our notation, their model for a univariate observed variable y_t can be expressed as:

$$y_t = x_t\beta + h_t^{-\frac{1}{2}}e_t, \qquad e_t \sim N(0,1)$$

where x_t is a vector of predetermined variables which could include lags of y_t, and the inverse of h_t is the time varying volatility. The law of motion for the volatilities is:

$$h_{t+1} = h_t\frac{\nu_t}{\lambda} \qquad \nu_t \sim Beta(\alpha_1, \alpha_2) \tag{2.1}$$

with $\alpha_2 = \frac{1}{2}$. The initial distribution is a gamma with parameters ν and S_1 such that h_1

has the following density function:

$$f(h_1|S_1) = h_1^{\frac{\nu}{2}-1} \exp\left(-\frac{h_1}{2S_1}\right)\frac{1}{\Gamma(\nu/2)(2S_1)^{\frac{\nu}{2}}} \tag{2.2}$$

where for mathematical convenience the initial density is restricted such that $\alpha_1 = \frac{\nu}{2}$. The parameters to be estimated are β, ν, λ and S_1. Note that, in contrast with the other models in this paper, the volatility follows a non-stationary process. As shown in subsection 6.6 of the Appendix, defining $Z = h_1 - \lambda h_2$ for $\in (0, \infty)$, the likelihood for this model can be obtained by integrating over the state variable Z. Given that the process for the stochastic volatility is multiplicative, the likelihood is as follows:

$$\pi(y_t|y_{1:t-1}) = \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)}\lambda^{\alpha_1}\left(\frac{S_{t+1}}{S_t}\right)^{\alpha_1}\frac{1}{\sqrt{2\pi}}\left(2\left((y_t - x_t\beta)^2 + \frac{1}{S_t}\right)^{-1}\right)^{\alpha_2} \tag{2.3}$$

where $S_t = \left((y_{t-1} - x_{t-1}\beta)^2 + \frac{1}{S_{t-1}}\right)^{-1}\frac{1}{\lambda}$ and $y_{1:t-1} = (y_1, y_2, ..., y_{t-1})$. To facilitate the reading here and in the following we do not write explicitly x_t as a conditioning argument.

The framework in Shephard (1994) provides a formal justification to Bayesian methods of variance discounting used in earlier literature (West & Harrison (2006), p.p. 360-361).

Creal (2017) shows that closed form solutions for the likelihood can be obtained for a family of non linear state space models with observation densities $p(y_t|h_t, x_t; \theta)$, in which the continuous valued time varying state variable h_t can be analytically integrated out conditionally on a discrete auxiliary variable z_t. x_t in these models are the predetermined regressors and θ is a parameter vector. The models in this class are defined as follows:

$$\begin{aligned}
y_t &\backsim p(y_t|h_t, x_t; \theta)\\
h_t &\backsim Gamma(\nu + z_t, c)\\
z_t &\backsim Poisson\left(\frac{\phi h_{t-1}}{c}\right)
\end{aligned}$$

where c is a scale parameter and ϕ determines the persistence of the state variable. For example Creal (2017) provides the following two alternative sufficient conditions for being able to integrate analytically these densities conditional on z_t:

$$\begin{aligned}
p(h_t|\alpha_1, \alpha_2, \alpha_3) &\propto h_t^{\alpha_1}\exp(\alpha_2 h_t + \alpha_3 h_t^{-1})\\
p(h_t|\alpha_1, \alpha_2, \alpha_3) &\propto h_t^{\alpha_1}(1 + h_t)^{\alpha_2}\exp(\alpha_3 h_t)
\end{aligned}$$

where $\alpha_{1:3}$ are functions of only the observations and parameters of the model. Thus, the contribution to the likelihood of one observation conditional on z_t can be obtained by integrating out the continuous state variables h_t analytically. The model that is obtained after integration simplifies to a Markov Switching model over the support of the non-negative discrete state variables z_t. The likelihood for these Markov Switching models can therefore be obtained recursively. Creal (2017) gives the detailed recursive formulas to obtain the likelihood for some specific models within this family, such as the gamma stochastic volatility models, stochastic duration models, stochastic count models and cox processes.

The gamma SV model by Creal (2017) can be expressed as follows:

$$y_t = \mu + x_t\beta + \gamma h_t + \sqrt{h_t}e_t, \qquad\qquad e_t \sim N(0,1)$$

where γ determines the skewness. When $\gamma = 0$ the model implies a variance-gamma distribution for the observed variable, which has thin tails (Madan & Seneta, 1990). The initial stationary distribution is $h_1 \backsim \text{Gamma}\big(\nu, \frac{c}{1-\phi}\big)$ and the unconditional mean is $E(h_1) = \frac{\nu c}{1-\phi}$.

More recently Sundararajan & Barreto-Souza (2023) propose a composite likelihood approach for the same model that we analyze in this paper, and which was estimated with Bayesian methods earlier by Leon-Gonzalez (2019). While they do not obtain the MLE as we do, their approach uses an expectation maximization algorithm to find the maximum of the composite likelihood, albeit with some restrictions.

3 Model Specification, Likelihood and Volatility Estimates

The model that we analyze is the same as in Leon-Gonzalez (2019) and assumes that the distribution of the one dimensional y_t conditional on an observed predetermined vector of regressors x_t can be described as follows:

$$y_t = \mu + x_t\beta + e_t, \qquad\qquad e_t|k_t \sim N\left(0, \frac{1}{k_t B^2}\right) \tag{3.1}$$

where β is a conformable vector of coefficients, μ and B^2 are scalar parameters and e_t is independent of x_t. The state variable k_t follows an autoregressive Gamma process (Gouriéroux & Jasiak, 2006) which can be described by writing $k_t = z_t'z_t$, where z_t is a $n \times 1$ vector that has the following Gaussian AR(1) representation:

$$z_t = \rho z_{t-1} + \varepsilon_t \qquad\qquad \varepsilon_t \sim N(0, \theta^2 I_n) \tag{3.2}$$

where ρ is a scalar that controls the persistence of the volatility, with $|\rho| < 1$. The stationary and initial distribution of the time varying inverse volatility k_t is a gamma with n degrees of freedom, such that $k_1 \backsim \text{Gamma}\big(n/2, \frac{2\theta^2}{1-\rho^2}\big)$. Therefore we have that $E(\frac{1}{k_t B^2}) = E(Var(e_t|k_t)) = \frac{1}{B^2}\frac{1-\rho^2}{n-2}$, provided that $n > 2$, where as a normalization we assume $\theta^2 = 1$ because we have B^2 in (3.1). For $0 < n \leq 2$ the model is well-defined but the volatility does not have a finite mean. The conditional distribution of $k_t|k_{t-1}$ is a non central chi-squared times a parameter constant that can be written as a mixture of gammas. The noncentral chi squared is well defined for non-integer values of n, so we will treat the unknown parameter n as a continuous parameter.

Then, given the properties of a gamma, the conditional mean of the inverse volatility k_t given previous history of k_t is a weighted average of the unconditional mean of k_t and its previous value k_{t-1}.

$$E(k_t|k_{t-1}) = \rho^2 k_{t-1} + (1-\rho^2)E(k_t)$$

where $\rho^2 k_{t-1}$ represents the non centrality parameter. k_t is correlated with its previous value and this generates the persistence in the squared residuals, a characteristic feature of time-varying variance models.

The inverse gamma specification implies a student-t distribution with n degrees of freedom for y_t thus enabling us to model heavy tailed distributions. In contrast, the gamma SV model (Creal, 2017) implies a variance-gamma distribution, which has thin tails (Madan & Seneta, 1990). The local scale model of Shephard (1994) is non-stationary, unlike ours which is stationary. In addition, the local scale model requires a restriction on the initial distribution for conjugacy (i.e. $\nu = 2\alpha_1$).

Integrating out analytically the volatilities in our model not only allows us to get a closed form expression for the likelihood, but also to see the similarity of our model to GARCH models. In particular we can see that the variance at each point in time given previous data is a (nonlinear) function of previous residuals. Using the filtering distributions in Section 3.2, we obtain the following:

- $y_1|k_1 \sim N(\mu + x_1\beta, (B^2 k_1)^{-1})$, where k_1 is a gamma. Therefore the first observation is a student-t with n degrees of freedom.

- Similarly for the second observation $y_2|y_1, k_2 \sim N(\mu + x_2\beta, (B^2k_2)^{-1})$, where $k_2|y_1$ is a mixture of gammas. $E(k_2|y_1)$ is a nonlinear function of past residuals.

- For any t, $y_t|y_{t-1}, ..., y_1, k_t \sim N(\mu + x_t\beta, (B^2k_t)^{-1})$, where $k_t|y_{t-1}, ..., y_1$ is a mixture of gammas, whose expected value is a nonlinear function of all past residuals.

Thus, integrating out the volatilities gives a structure similar to GARCH models, but with a different functional form and distribution.

3.1 The Likelihood

The following proposition, whose proof is in Appendix 6.2, gives the likelihood for the model described in equations (3.1)-(3.2).

Proposition 3.1. *Let $e_t = y_t - \mu - x_t\beta$ for $t = 1, ..., T$. The likelihood for the first observation is:*

$$L(y_1) = (2\pi)^{-\frac{1}{2}}\sqrt{B^2}2^{\frac{1}{2}}\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})}\left|B^2e_1^2 + V_1^{-1}\right|^{-\frac{n+1}{2}}V_1^{-\frac{n}{2}}$$

for the second is:

$$L(y_2|y_1) = (2\pi)^{-\frac{1}{2}}\sqrt{B^2}\frac{2^{\frac{n+1}{2}}}{2^{\frac{n}{2}}}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\frac{(B^2e_2^2+1)^{-\frac{n+1}{2}}}{(1-\delta_2)^{-\frac{n+1}{2}}}\hat{C}_2$$

for the third is:

$$L(y_3|y_2, y_1) = (2\pi)^{-\frac{1}{2}}\sqrt{B^2}\frac{1}{c_3}\sum_{h_2=0}^{\infty}\widetilde{C}_{2,h_2}\frac{\Gamma\left(\frac{n+1+2h_2}{2}\right)}{(B^2e_3^2+1)^{\frac{n+1}{2}}}(2S_3)^{\frac{n+1+2h_2}{2}}\frac{2^{\frac{n+1}{2}}}{2^{\frac{n}{2}}}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\hat{C}_3$$

for the fourth is:

$$L(y_4|y_3, y_2, y_1) = (2\pi)^{-\frac{1}{2}}\sqrt{B^2}\frac{1}{c_4}\sum_{h_3=0}^{\infty}\widetilde{C}_{3,h_3}\frac{\Gamma\left(\frac{n+1+2h_3}{2}\right)}{(B^2e_4^2+1)^{\frac{n+1}{2}}}(2S_4)^{\frac{n+1+2h_3}{2}}\frac{2^{\frac{n+1}{2}}}{2^{\frac{n}{2}}}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\hat{C}_4$$

and for any $t \geq 3$ is

$$L(y_t|y_{1:t-1}) = (2\pi)^{-\frac{1}{2}}\sqrt{B^2}\frac{1}{c_t}\sum_{h_{t-1}=0}^{\infty}\widetilde{C}_{t-1,h_{t-1}}\frac{\Gamma\left(\frac{n+1+2h_{t-1}}{2}\right)}{(B^2e_t^2+1)^{\frac{n+1}{2}}}(2S_t)^{\frac{n+1+2h_{t-1}}{2}}\frac{2^{\frac{n+1}{2}}}{2^{\frac{n}{2}}}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\hat{C}_t$$

where:

$$
\begin{aligned}
&V_1 = (1-\rho^2)^{-1}\\
&\widetilde{V}_2^{-1} = V_1^{-1} + B^2 e_1^2\\
&\delta_2 = \rho^2(\widetilde{V}_2^{-1}+\rho^2)^{-1}\\
&Z_2 = (1+B^2e_2^2)^{-1}\delta_2\\
&\widetilde{C}_{2,h_2} = \frac{[(n+1)/2]_{h_2}}{[n/2]_{h_2}}\left(\frac{1}{2}\rho^2(\widetilde{V}_2^{-1}+\rho^2)^{-1}\right)^{h_2}\frac{1}{h_2!}\\
&\widetilde{C}_{3,h_3} = \sum_{h_2=0}^{\infty}\widetilde{C}_{2,h_2}\Gamma\left(\frac{n+1+2h_2}{2}\right)\frac{[(n+1)/2+h_2]_{h_3}}{[n/2]_{h_3}}\left(\frac{1}{2}\rho^2S_3\right)^{h_3}\frac{1}{h_3!}(2S_3)^{\frac{n+1+2h_2}{2}}\\
&c_3 = {}_2F_1\left(\frac{n+1}{2},\frac{n+1}{2};\frac{n}{2};\delta_3\right)\Gamma\left(\frac{n+1}{2}\right)(1-\rho^2S_3)^{-\frac{n+1}{2}}(2S_3)^{\frac{n+1}{2}}\\
&\hat{C}_t = {}_2F_1\left(\frac{n+1+2h_{t-1}}{2},\frac{n+1}{2};\frac{n}{2};Z_t\right)\quad \textit{for } t \geq 2 \textit{ and where } h_1 = 0
\end{aligned}
$$

for $T \geq t \geq 3$

$$
\begin{aligned}
&S_t = (1+B^2e_{t-1}^2+\rho^2)^{-1}\\
&\widetilde{V}_t^{-1} = 1 + B^2e_{t-1}^2\\
&Z_t = (B^2e_t^2+1)^{-1}S_t\rho^2\\
&\delta_t = \left((1-\rho^2S_t)^{-1}S_t\rho^2(\widetilde{V}_{t-1}^{-1}+\rho^2)^{-1}\right)
\end{aligned}
$$

and for $T+1 \geq t \geq 4$

$$
c_t = \sum_{h_{t-1}=0}^{\infty}\widetilde{C}_{t-1,h_{t-1}}(1-\rho^2S_t)^{-\frac{n+1+2h_{t-1}}{2}}\Gamma\left(\frac{n+1+2h_{t-1}}{2}\right)(2S_t)^{\frac{n+1+2h_{t-1}}{2}}
$$

$$
\widetilde{C}_{t-1,h_{t-1}} = \sum_{h_{t-2}=0}^{\infty}\widetilde{C}_{t-2,h_{t-2}}\Gamma\left(\frac{n+1+2h_{t-2}}{2}\right)\frac{[(n+1)/2+h_{t-2}]_{h_{t-1}}}{[n/2]_{h_{t-1}}}\left(\frac{1}{2}\rho^2S_{t-1}\right)^{h_{t-1}}\frac{(2S_{t-1})^{\frac{n+1+2h_{t-2}}{2}}}{h_{t-1}!}
$$

and $S_{T+1} = (1+B^2e_T^2)^{-1}$

The rising factorial is denoted as $[x]_h$ and ${}_2F_1$ denotes a hypergeometric function (e.g. Muirhead (2005, p. 20)). There are a number of transformations to the ${}_2F_1$ hypergeometric functions above to accelerate their convergence. Abramowitz et al. (1988, p.559) defines several transformations such as the Euler transformation:

$$ {}_2F_1(a,b;c;z) = (1-z)^{c-a-b}{}_2F_1(c-a,c-b;c;z) $$

or a linear combination approach:

$$ \begin{aligned} {}_2F_1(a,b;c;z) = &\frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}{}_2F_1(a,b;a+b-c+1;1-z) \\ &+(1-z)^{c-a-b}\frac{\Gamma(c)\Gamma(a+b-c)}{\Gamma(a)\Gamma(b)}{}_2F_1(c-a,c-b;c-a-b+1;1-z) \\ &\text{for } (|arg(1-z)| < \pi) \end{aligned} $$

The expression for $\hat{C}_t$ above transformed using the Euler transformation becomes:

$$ \hat{C}_t = (1-Z_t)^{-\frac{n+2+2h_{t-1}}{2}}{}_2F_1\bigg(-\frac{1+2h_{t-1}}{2},-\frac{1}{2};\frac{n}{2};Z_t\bigg) \text{ for t} \geq 2 \text{ and where } h_1 = 0 $$

In our coding we used the Euler acceleration only for $\hat{C}_2$ and c_3, because for larger values of t the acceleration did not converge when h was large. Regarding the linear combination approach, although we did not implement it in our code for the R package, the acceleration converges. We accelerated the calculations by implementing parallel computing in the code. This is possible because many of the coefficients in the series are the same for every t, therefore they only need to be computed once, which can be done in parallel. We also calculate all the $\hat{C}_t$ in parallel. As shown in Section 4, this drastically reduces computation time. The derivatives of the log-likelihood can be obtained as a byproduct of the likelihood calculation.

After integrating out the volatilities, this likelihood can be calculated recursively starting with y_1, which is the first observation, to y_T. This likelihood is easy to compute and it always converges since $|Z_t| < 1$ for all values of t. We truncate the number of terms to calculate the hypergeometric functions to around 350 to ensure convergence, and the sums are truncated at about $h = 350$. These truncation values seemed to be sufficient as explained in Table 1 in our application using inflation data.

3.2 Joint Smoothing and Filtering Distributions

In this subsection, we provide the analytical expressions for both the joint smoothing and filtering distributions for the volatilities. Propositions 3.2, 3.3 and 3.4, proved in the Appendix, provide the smoothing distributions in alternative forms. Proposition 3.2 and 3.3 give the conditional distributions $\pi(k_t|k_{(t+1):T}, y_{1:T})$, and $\pi(k_t|k_{1:(t-1)}, y_{1:T})$, respectively, while Proposition 3.4 gives the marginals $\pi(k_t|y_{1:T})$. The filtering distributions are stated after Proposition 3.4.

Proposition 3.2. *The joint posterior distribution $\pi(k_{1:T}|y_{1:T})$ can be obtained from the following conditional densities each of which is a mixture of gammas:*

$$\pi(k_t|k_{(t+1):T}, y_{1:T}) \propto |k_t|^{\frac{n+1-2}{2}} \exp\left(-\frac{1}{2} S_{t+1}^{-1} k_t \right) \sum_{h=0}^{\infty} \left(C_{t,h} |k_t|^h \right),\ t = 1, ..., T$$

where

$$C_{1,h} = \frac{1}{h!} \frac{1}{[n/2]_h} \left(\frac{1}{4} \rho^2 k_2 \right)^h$$
$$S_2 = (1 + B^2 e_1^2)^{-1}$$
$$S_{T+1} = (1 + B^2 e_T^2)^{-1}$$

for $3 \leq t \leq T$

$$S_t = (1 + B^2 e_{t-1}^2 + \rho^2)^{-1}$$

and for $2 \leq t < T$*:*

$$C_{t,h} = \sum_{h_t=0}^{h} \widetilde{C}_{t,h-h_t} \frac{1}{[n/2]_{h_t}} \left(\frac{1}{4} \rho^2 \right)^{h_t} \frac{k_{t+1}^{h_t}}{h_t!}$$

while for $t = T$*,* $C_{t,h} = \widetilde{C}_{t,h}$*, and where* $\widetilde{C}_{t,h}$ *has been defined in Proposition 3.1.*

Proposition 3.3. *The density* $\pi(k_t|k_{1:(t-1)}, y_{1:T})$ *is a mixture of gamma distributions and its kernel is proportional to:*

$$\pi(k_{\text{T-s}}|k_{1:(\text{T-s-1})}, y_{1:T}) \propto |k_{\text{T-s}}|^{\frac{n+1-2}{2}} \exp\left(-\frac{1}{2} S_{\text{T-s}+1}^{-1} k_{\text{T-s}} \right) \left(\sum_{h=0}^{\infty} a_{\text{T-s},h} k_{\text{T-s}}^h \right) \qquad s = 0, \ldots, T-1$$

where

$$a_{\text{T-s},h} = \sum_{h_{\text{T-s}}=0}^{h} \widetilde{a}_{\text{T-s},h-h_{\text{T-s}}} \frac{1}{(h_{\text{T-s}})!} \frac{1}{[n/2]_{h_{\text{T-s}}}} \left(\frac{1}{4} \rho^2 \right)^{h_{\text{T-s}}} k_{\text{T-S}-1}^{h_{\text{T-s}}}, \qquad s = 1, \ldots, T-2$$

and

$$\widetilde{a}_{\text{T-s},h_{\text{T-s+1}}} = \sum_{h_{\text{T-s+2}}=0}^{\infty} \widetilde{a}_{\text{T-s+1},h_{\text{T-s+2}}} \Gamma\left(\frac{n+1}{2} + h_{\text{T-s+2}}\right) \frac{[(n+1)/2 + h_{\text{T-s+2}}]_{h_{\text{T-s+1}}}}{[n/2]_{h_{\text{T-s+1}}}}$$
$$\times \frac{\left(\frac{1}{2}\rho^2 S_{\text{T-s+2}}\right)^{h_{\text{T-s+1}}}}{(h_{\text{T-s+1}})!} (2S_{\text{T-s+2}})^{\frac{n+1+2h_{\text{T-s+2}}}{2}} \qquad s = 2, \ldots, T-1$$

with,

$$a_{T,h} = \frac{1}{h!} \frac{1}{[n/2]_h} \left(\frac{1}{4}\rho^2 k_{T-1}\right)^h$$

$$\widetilde{a}_{T-1,h_T} = \frac{[(n+1)/2]_{h_T}}{[n/2]_{h_T}} \frac{\left(\frac{1}{2}\rho^2 S_{T+1}\right)^{h_T}}{(h_T)!}$$

For the case when $s = T-1$, *we have* $a_{\text{T-s},h} = a_{1,h} = \widetilde{a}_{1,h}$.

We can integrate $\pi(k_{1:T})\pi(y_{1:T}|k_{1:T})$ with respect to $k_{1:(t-1)}$ and with respect to $k_{(t+1):T}$ to obtain the following proposition which gives the marginal density $\pi(k_t|y_{1:T})$ for $t = 2, \ldots, T-1$. Note that for $t = T$ or $t = 1$ the marginal densities are given by Propositions 3.2 and 3.3, respectively.

Proposition 3.4. *The density of* $\pi(k_t|y_{1:T})$ *is that of a mixture of gammas and its kernel is given by:*

$$\pi(k_t|y_{1:T}) \propto |k_t|^{\frac{n+1-2}{2}} \exp\left(-\frac{1}{2}S_{t+1}^{-1}k_t\right) \sum_{h=0}^{\infty} \widetilde{D}_{t,h}|k_t|^h$$

for $t = 2, \ldots, T-1$, *where for* $2 \leq t < T-1$:

$$\widetilde{D}_{t,h} = \sum_{h_t=0}^{h} \sum_{h_{t+1}=0}^{\infty} \widetilde{C}_{t,h-h_t} \frac{1}{[n/2]_{h_t}} \left(\frac{1}{4}\rho^2\right)^{h_t} \frac{1}{h_t!} \frac{\Gamma\left(\frac{n+1}{2} + h_t + h_{t+1}\right)}{(S_{t+2}^{-1}/2)^{\frac{n+1}{2}+h_t+h_{t+1}}} \widetilde{a}_{t+1,h_{t+1}}$$

and for $t = T-1$:

$$\widetilde{D}_{T-1,h} = \sum_{h_{T-1}=0}^{h} \widetilde{C}_{T-1,h-h_{T-1}} \frac{1}{[n/2]_{h_{T-1}}} \left(\frac{1}{4}\rho^2\right)^{h_{T-1}} \frac{1}{(h_{T-1})!} \frac{\Gamma\left(\frac{n+1}{2} + h_{T-1}\right)}{(S_{T+1}^{-1}/2)^{\frac{n+1}{2}+h_{T-1}}}$$

where $\widetilde{a}_{t+1,h}$ *was defined in Proposition 3.3 and* $\widetilde{C}_{t,h_t}$ *was defined in Proposition 3.1.*

Regarding the filtering distributions, they were obtained in the proof of Proposition 3.1.

They are a mixture of gammas and the kernel is given by:

$$\pi(k_t|y_{t-1}, y_{t-2}, ..., y_1) \propto |k_t|^{\frac{n-2}{2}} \exp\left(-\frac{1}{2}k_t\right) \sum_{h=0}^{\infty} \left(\widetilde{C}_{t,h}|k_t|^h\right),\ t = 1, ..., T$$

where the recursive constants are defined in Proposition 3.1.

4 Empirical Applications

4.1 Macroeconomic Data

To illustrate the efficiency and usefulness of our proposed novel addition to the SV literature, we provide macroeconomic applications using inflation data for the UK, Japan, US and Brazil. The data series were all sourced from the Federal Reserve Bank of St Louis Fred database as the Consumer Price Index (CPI) data and inflation was constructed using the following formula:

$$\text{Inflation} = \frac{CPI_t - CPI_{t-1}}{CPI_{t-1}} \times 100$$

The number of observations for each series were determined by availability of data. UK data thus covers the period 1960Q2 to 2022Q1 and Japan data is obtained for the period 1960Q4 to 2022Q1. The US inflation data covers the period 1960Q1 to 2021Q4. Due to unavailability of data for earlier years for Brazil we have observations for the period 1981Q1 to 2021Q4. y_t is the level of inflation and x_t contains a constant and 4 lags of y_t. Therefore, for each series we have 244, 242, 244 and 160 observations, respectively, after constructing the lags.

Figure 1 illustrates the quarterly inflation series for the four countries in levels. The trend for the evolution of inflation for the US, UK and Japan in the early 1970's and 1980's have slight similarities. However, in later years across all series, inflation evolves differently.

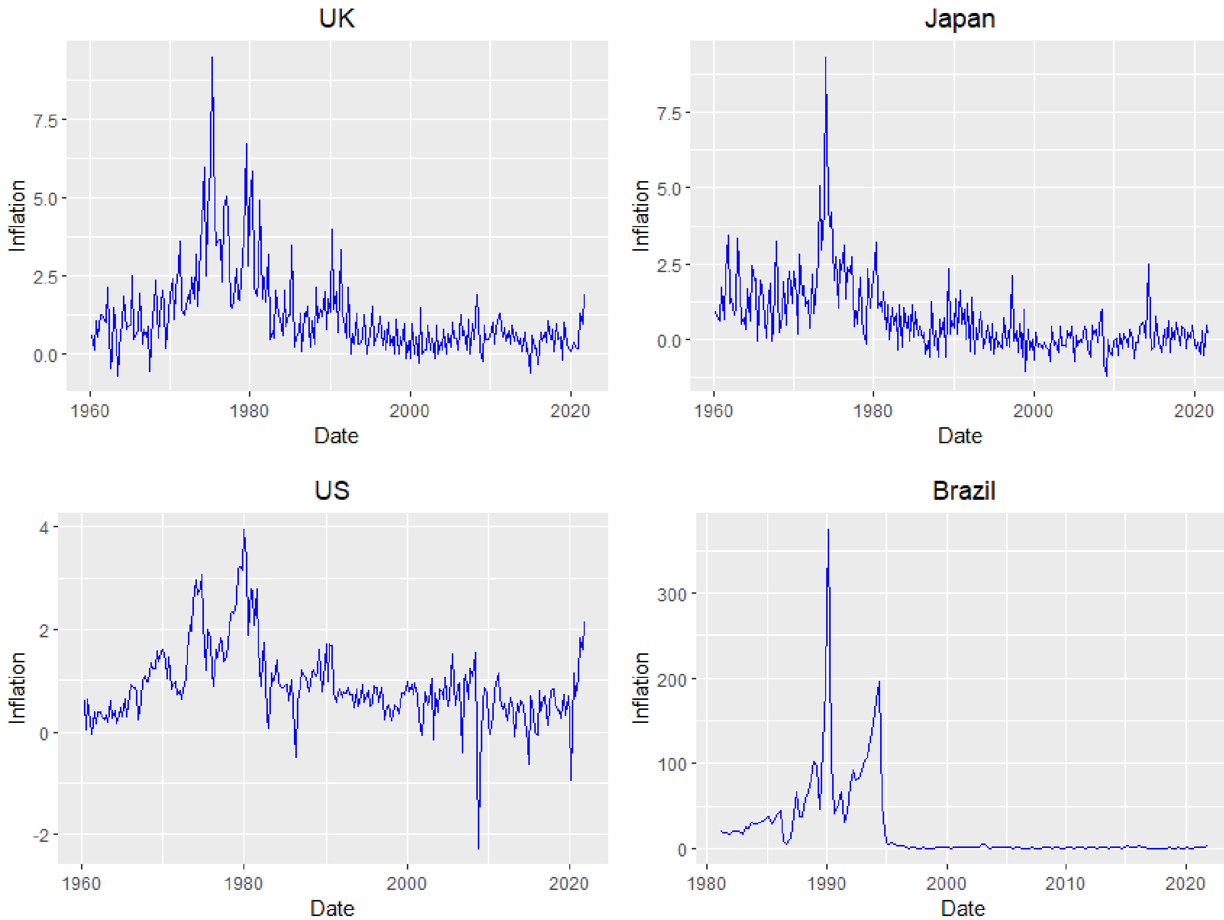

Figure 1: Inflation Rates. The x-axis plots the dates that correspond to the end of each year for the quarterly observations. The y-axis plots the Inflation Rates

Figure 2 shows the Ordinary Least Squares (OLS) residuals for the four countries over the sample period, after regressing the level of inflation on its 4 lags and an intercept. The spikes in volatility observed for Brazil inflation show that the series accumulates periods of consistent high volatility continuously. Overall for all countries, volatility patterns exhibit some extreme values suggesting that models that assume heavier tailed distributions might fit better and improve forecasting.

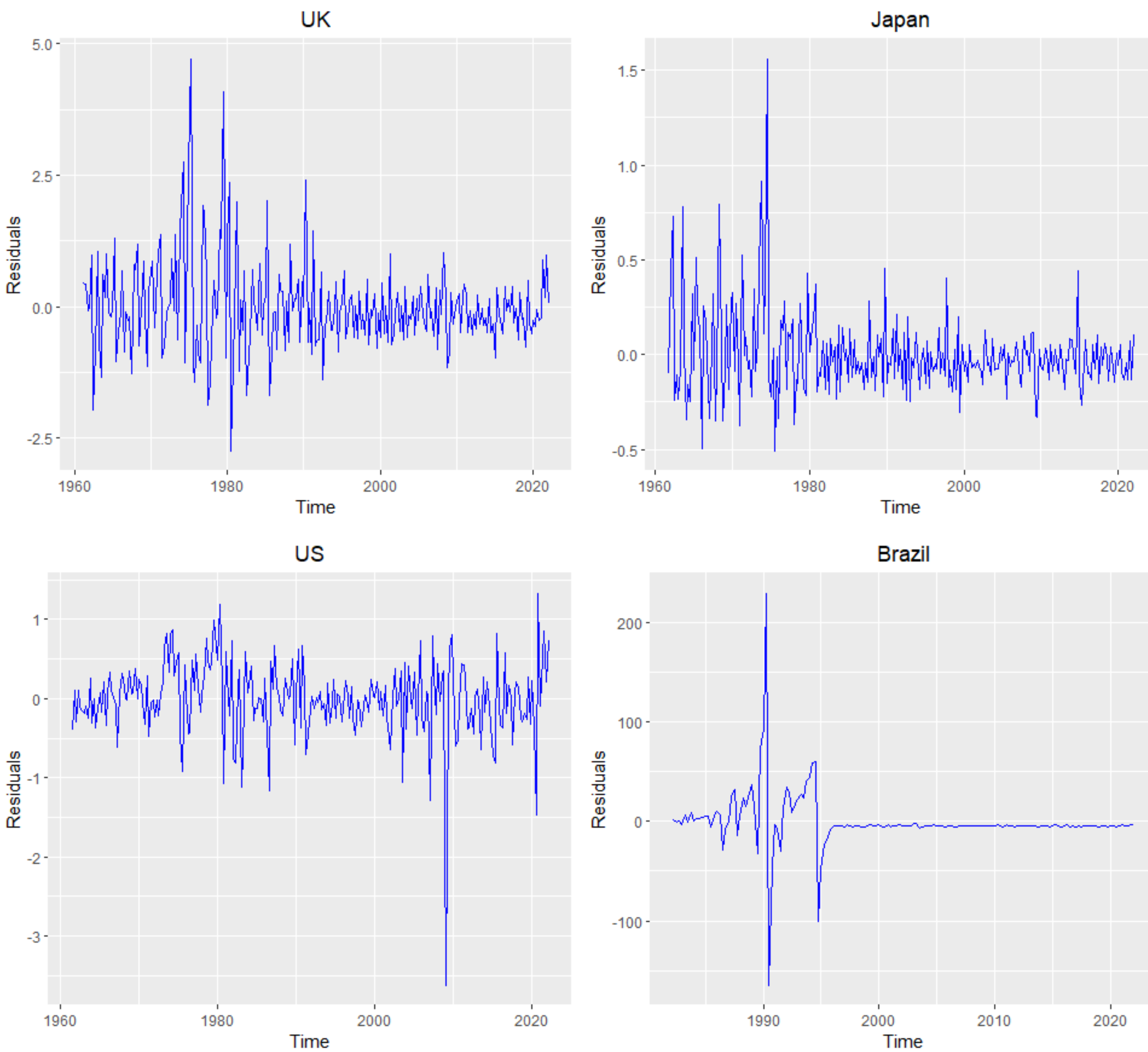

Figure 2: Residuals Plots. The x-axis plots the time period. The y-axis plots the OLS Residuals

In the maximization algorithm, the initial values for the slope coefficients are equal to the OLS estimates, and for the rest of the parameters we choose values such that the mean volatility implied by the model equals that of the data. We truncate the calculation of hypergeometric functions at 300 terms and we truncate h_t in the likelihood at $h_t = 300$ to ensure convergence.

4.1.1 Smoothed Estimates of the Volatilities

Using the smoothing distributions we are able to obtain an estimate of the variance of e_t at each point of time given all available data: $E(var(e_t|k_t)) = E(var(y_t|x_t, k_t))$, where the expectation is with respect to the smoothing distribution of k_t (i.e. $\pi(k_t|y_{1:T})$). This is in contrast to the commonly used GARCH MLE estimates, which can only provide the filtered estimates of the variance: $var(e_t|y_{1:(t-1)})$. Figure 3 compares the MLE smoothed estimates of the variance at each point in time for each country, to the moving average of the squared OLS residuals obtained from 5 continuous squared residuals.

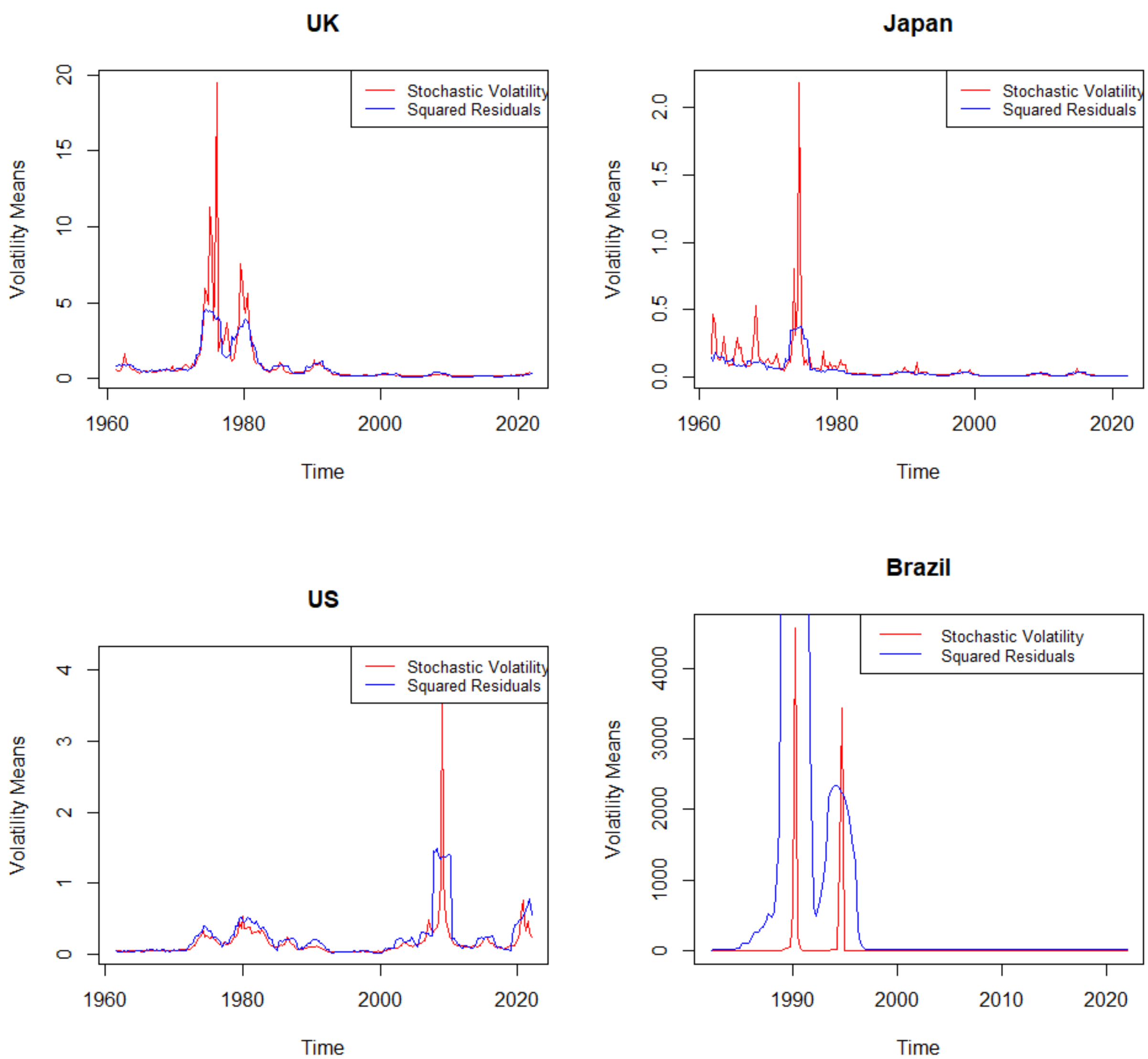

Figure 3: Smoothed Estimates of the Volatilities. The red lines show the smoothed estimates of the volatilities compared to the moving average of OLS squared residuals displayed in blue

The periods with high residuals coincide with periods of high estimated stochastic volatility each point in time for all the four countries. In particular for the US and UK the estimates reflect the expectations for high volatility trends observed during periods such as the Great recession and smaller peaks in volatility representing the covid recessions.

4.1.2 Accuracy Check

We compare our novel algorithm to the Particle Filter to check the accuracy of our computations. Particle filters are commonly used in practice for calculating the likelihood function. Literature has it that they provide an unbiased estimate of the likelihood (see e.g Moral (2004), proposition 7.4.1). We use the UK inflation data for this exercise. Parameter values for both algorithms are set at the maximum likelihood estimates. To evaluate each value of the likelihood we use the average of 110 independent replications of the particle filter proposed in Chan et al. (2020). We set the number of particles to twice the sample size T, that is each particle filter has $T * 2$ particles. We obtained 100 values for the log-likelihood using this method and plot them in Figure 4 together with the value provided by our algorithm.

The exact log likelihood estimate for the UK inflation data is -229.87. The figure shows that the particle filter value for the log-likelihood goes above and below our exact value. Therefore our solution seems accurate.

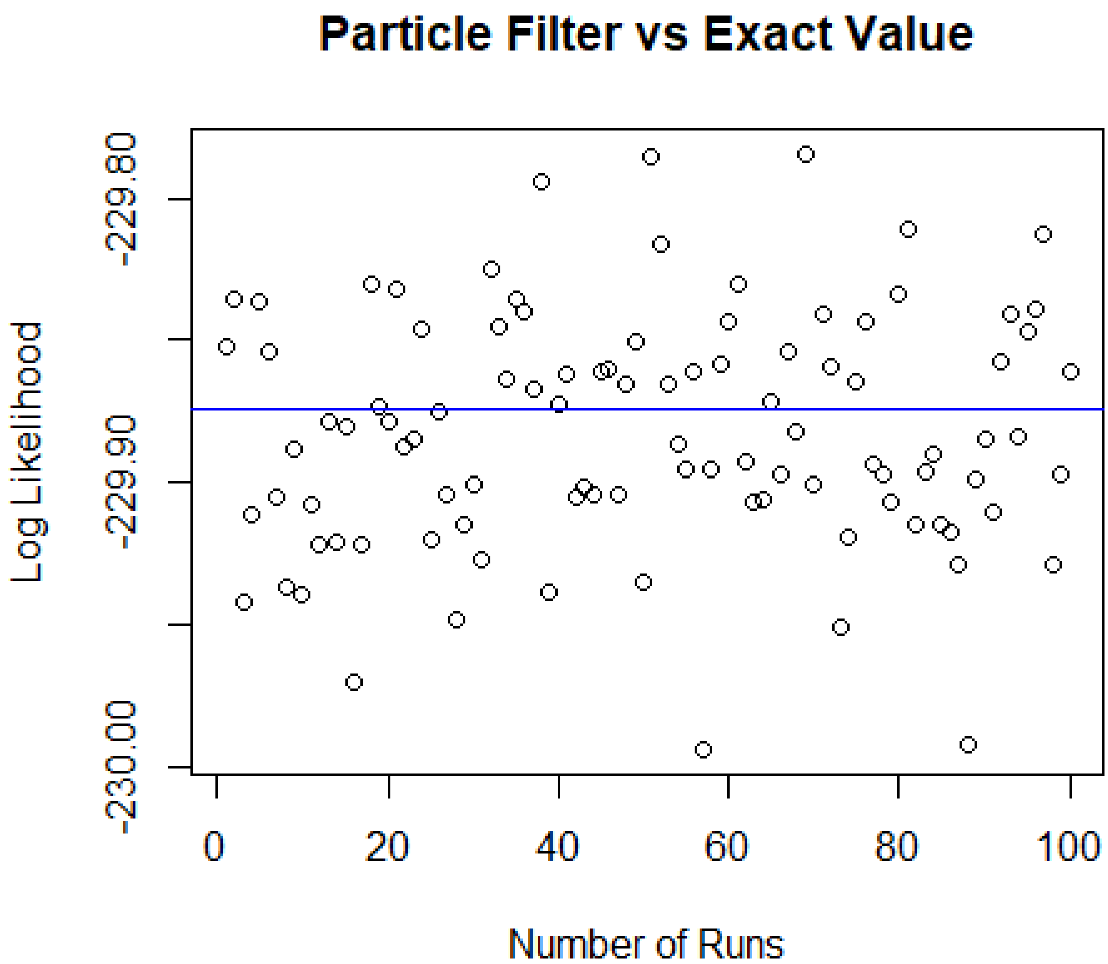

Figure 4: Particle Filter Estimates. The horizontal blue line represents the exact value obtained using our novel algorithm. Small circles show the 100 log-likelihood estimates, each of which was obtained by averaging 110 runs of the particle filter

4.1.3 Computational Efficiency

In order to calculate the likelihood, we need to truncate the number of terms that are added for the hypergeometric functions ($niter$), and also we need to truncate h. For simplicity we use the same truncation points for both. Table 1 shows the values of the log likelihood obtained for several truncation values, using the MLE estimates for the parameter values and the four datasets. The value of the log-likelihood remains stable at truncation points of 150 (Japan), 200 (US), 300 (UK) and 350 (Brazil).

Using a truncation point of 350, the computation time for one evaluation of the likelihood in seconds for the UK inflation dataset ($T = 244$) is 0.24, 0.39, 0.72 and 2.60 when using 18, 8, 4, or just one computing thread, respectively. For the UK exchange rate dataset ($T = 999$) that we use in Section 4.2 a truncation point of 350 was also adequate, and the computation times for the same increase to 0.82, 1.42, 2.72, 10.07, respectively. The coding was done in C++, linked to the R software and executed in a Ryzen threadripper 3970x processor.

Table 1: Likelihood at different truncation parameter values

	UK	**Japan**	**US**	**Brazil**
$niter = h = 100$	-234.59	102.58	-124.61	-392.51
$niter = h = 150$	-230.48	102.67	-124.58	-387.29
$niter = h = 200$	-229.91	102.67	-124.57	-385.91
$niter = h = 300$	-229.87	102.67	-124.57	-385.63
$niter = h = 350$	-229.87	102.67	-124.57	-385.62
$niter = h = 400$	-229.87	102.67	-124.57	-385.62

4.1.4 Parameter Estimates and Model Comparison

Maximum likelihood parameter estimates are reported in Table 2 for our model using quarterly inflation data for the UK, Japan, US and Brazil and their standard errors in parenthesis. β_0 is the coefficient of the intercept while $\beta_{1:4}$ are the coefficients of the lags. Throughout the maximum likelihood estimation, we imposed the constraint $0 < \rho < 1$ on the persistence of volatility.

Table 2: Inverse Gamma SV Model Maximum Likelihood Estimates

Parameter	**UK**	**Japan**	**US**	**Brazil**
B^2	0.0653	2.2868	0.2845	0.0127
	(0.0354)	(1.2679)	(0.1670)	(0.0064)
ρ	0.9849	0.9734	0.9577	0.9964
	(0.0091)	(0.0159)	(0.0252)	(0.0048)
n	2.2527	2.0529	3.2136	0.7010
	(0.6534)	(0.4724)	(0.8377)	(0.1374)
β_0	0.1148	0.0053	0.1053	-0.1030
	(0.0492)	(0.0078)	(0.0418)	(0.0810)
β_1	0.1256	0.0222	0.5772	1.0604
	(0.0529)	(0.0557)	(0.0701)	(0.0607)
β_2	0.1627	0.2592	0.0500	-0.4053
	(0.0479)	(0.0537)	(0.0731)	(0.0499)
β_3	-0.1005	0.0247	0.3304	0.4889
	(0.0483)	(0.0517)	(0.0719)	(0.0924)
β_4	0.6140	0.4291	-0.0747	-0.0652
	(0.0485)	(0.0530)	(0.0638)	(0.0315)

The coefficients of the lags are mostly significant, and the estimates of ρ indicate high persistence of the volatility in all countries. In all cases except Brazil, the estimated values of n are bigger than 2, implying a finite value for the expected value of volatility. For Brazil we have $n = 0.7$, implying that y_t has very fat tails, similar to those of a Cauchy distribution.

We compare the empirical performance of the following 7 models:

M_1: Homoscedastic

M_2: Local scale model (Shephard, (1994))

M_3: Univariate GARCH(1,1) with normal errors

M_4: Univariate GARCH(1,1) with student t errors

M_5: Log Normal stochastic volatility (e.g. Kim et al. (1998))

M_6: Gamma stochastic volatility

M_7: Inverse Gamma stochastic volatility

Except M_5 all models are estimated by MLE. The model M_5 is estimated using Bayesian methods with the R package *stochvol* (Kastner (2016)), using the default non-informative priors implemented in the package. For this model the value of the log-likelihood at the posterior mean of parameters is evaluated by averaging 50 independent replications of a bootstrap particle filter, with each particle filter having a number of particles equal to 60 times the sample size. The numerical standard error of the log-likelihood estimate was smaller than 0.02 in all cases. Both the Gaussian and Student t GARCH are specified as $GARCH(1,1)$, thus they have 8 parameters and 9 parameters respectively given that we have 4 lags and an intercept. The stochastic volatility models have 8 parameters except for the gamma SV model which has an additional parameter for the skewness of volatility.

Table 3 reports the log likelihood values at the maximum likelihood estimates and Table 4 reports the values of the Bayesian Information Criterion (BIC, Schwarz 1978). As expected the homoscedastic model is the worst of all models for all countries. In terms of the log-likelihood the inverse gamma model is the best for the US, and the gamma SV model is the best for the UK and Japan. For Brazil the GARCH(1,1) with student-t errors has the best value of the log-likelihood, but when penalizing for the number of parameters using the BIC (the smaller the better) the inverse gamma SV model is the best. In summary, using the BIC the gamma SV model is the best for the UK and Japan, and the inverse gamma SV model is the best for the US and Brazil. In the case of the UK and Japan the asymmetry parameter of the Gamma SV model was estimated to be large, which might be the reason for the better performance of this model. In the case of Brazil and the US the residuals appear to have more abrupt changes, which might be the reason for the better performance of the inverse Gamma SV model.

Table 3: Inflation Rates Model Comparisons: Log Likelihood

Model	**UK**	**Japan**	**US**	**Brazil**
M_1	-306.74	18.39	-165.42	-763.33
M_2	-230.04	100.17	-129.90	-395.30
M_3	-233.01	90.87	-147.72	-387.76
M_4	-227.74	107.06	-133.34	-383.97
M_5	-229.08	101.96	-126.74	-389.63
M_6	-220.88	112.09	-129.33	-475.07
M_7	-229.87	102.67	-124.57	-385.62

Table 4: Inflation Rates Model Comparisons: BIC

Model	**Parameters**	**UK**	**Japan**	**US**	**Brazil**
T		244	242	244	160
M_1	6	646.46	-3.85	363.83	1557.12
M_2	8	504.05	-156.42	303.77	831.21
M_3	8	509.99	-137.83	339.41	816.11
M_4	9	504.95	-164.73	316.15	813.61
M_5	8	502.13	-160.00	297.47	819.85
M_6	9	491.24	-174.79	308.14	995.81
M_7	8	503.72	-161.43	293.12	811.84

4.2 Exchange Rates Data Application

We use 1000 daily exchange rate observations for 7 currencies (GBP, EUR, JPY, CND, AUD, BRL, ZAR) to the USD. The data for the first 6 currencies were obtained from the Board of Governors of the Federal Reserve and covers the period beginning 5 March 2019 and ending 3 March 2023. ZAR was obtained from the South African Reserve Bank for the period 7 May 2019 to 3 March 2023. In this analysis y_t is the first differences of the log exchange rate. All models include an intercept but we include no regressors (i.e. x_t is empty).

Figure 5 shows the normalised exchange rates for the 7 countries. We calculate the percentage of times that the absolute value of the normalised exchange rate goes beyond 1.96 standard deviations. The JPY, BRL, GBP, CAD, EUR, and AUD have thicker tails than a normal distribution with 5.8%, 5.7%, 5.9%, 5.2%, 6.5% and 6.1% proportions respectively. The ZAR has slightly thinner tails to the normal with 4.8% of the proportion going beyond 1.96 standard deviations.

In addition we obtain the proportion where the absolute value of the normalised exchange rate goes beyond 3.0902 standard deviations, which is 0.2% for a normal distribution. The ZAR has the lowest proportion, with 0.4%, but still larger than the normal. The JPY, BRL, GBP, CAD, EUR, and AUD distribution proportions are 1.8%, 1.0%, 1.6%,1.3%, 1.2%, 0.9%, respectively, all of them much greater than the normal.

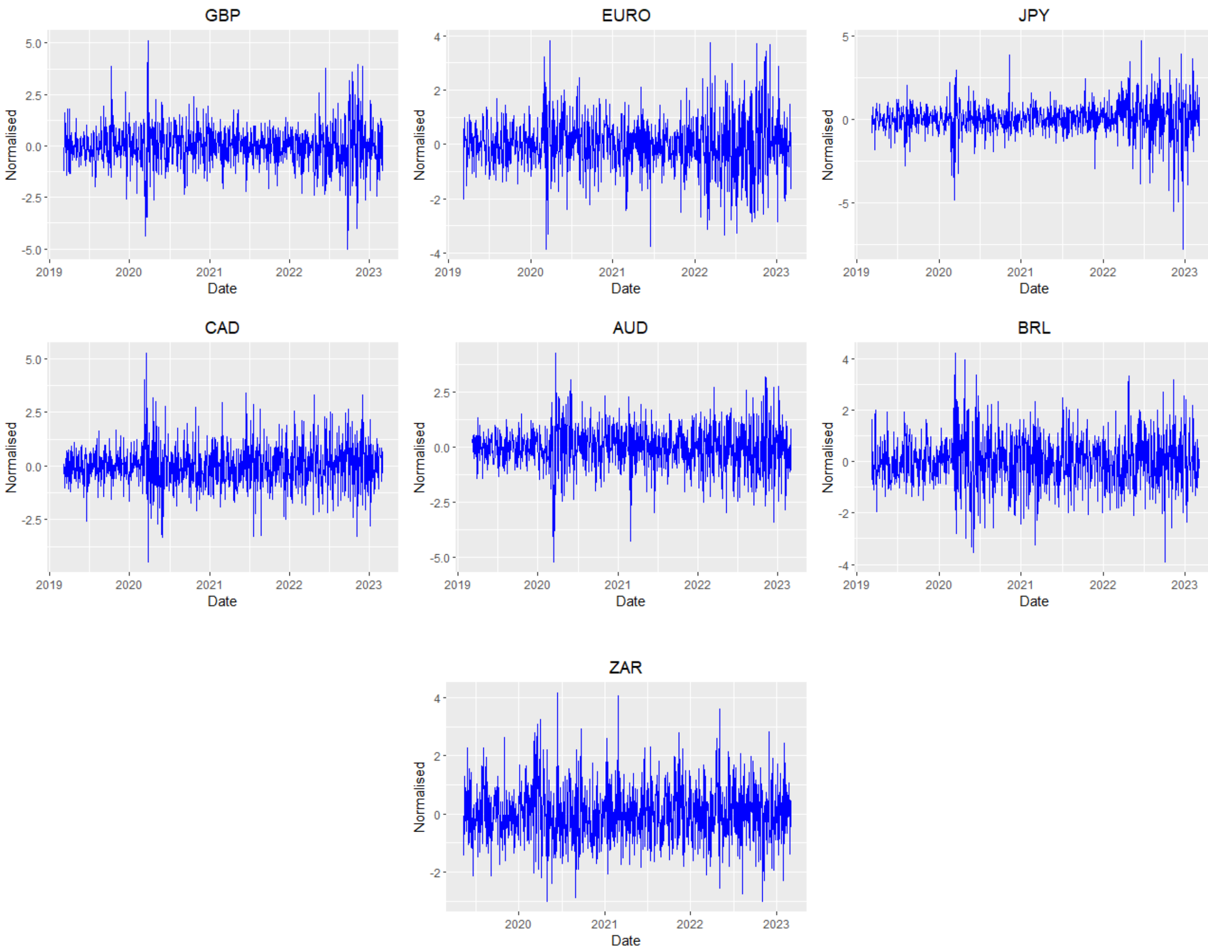

Figure 5: Normalised Exchange Rates. y_t was normalised by subtracting its mean and dividing by the standard deviation. The x-axis plots the dates that correspond to the end of each year for the daily observations. The y-axis plots the normalised y_t

Table 5 shows the log likelihood values and Table 6 the BIC values (the smaller the better) across all the 7 models listed above. The best model for the ZAR, which has the thinnest tails, is the Gamma SV model, both in terms of the likelihood and the BIC. For all the other currencies the GARCH(1,1) with student-t errors has the highest log-likelihood values. However, when taking into account the number of parameters using the BIC, this model is the best only for the EUR and JP. The inverse Gamma SV model is the best for all the other currencies, GBP, CAD, AUD, BRL, with the log normal SV model being equally good for the GBP and BRL.

Table 5: Exchange Rates Model Comparisons: Log likelihood

Model	**GBP**	**EUR**	**JPY**	**CAD**	**AUD**	**BRL**	**ZAR**
M_1	3659.66	3962.76	3770.79	3976.17	3551.11	3123.64	3236.27
M_2	3754.46	4053.28	3962.96	4034.74	3637.23	3167.60	3244.68
M_3	3747.26	4044.27	3927.91	4027.51	3632.31	3165.52	3249.98
M_4	3765.21	4059.93	3987.26	4041.50	3641.47	3171.54	3253.80
M_5	3762.50	4055.48	3971.76	4036.88	3638.15	3168.97	3251.93
M_6	3759.35	4053.41	3967.96	4034.91	3633.98	3168.72	3257.63
M_7	3762.81	4055.50	3973.79	4038.36	3640.23	3168.94	3252.42

Table 6: Exchange Rates Model Comparisons: BIC

Model	**Parameters**	**GBP**	**EUR**	**JPY**	**CAD**	**AUD**	**BRL**	**ZAR**
T		999	999	999	999	999	999	999
M_1	2	-7305.51	-7911.71	-7527.77	-7938.52	-7088.41	-6233.46	-6458.73
M_2	4	-7481.30	-8078.92	-7898.28	-8041.86	-7246.83	-6307.58	-6461.73
M_3	4	-7466.89	-8060.91	-7828.19	-8027.39	-7236.99	-6303.42	-6472.33
M_4	5	-7495.89	-8085.33	-7939.98	-8048.47	-7248.40	-6308.55	-6473.07
M_5	4	-7497.37	-8083.33	-7915.89	-8046.13	-7248.67	-6310.31	-6476.23
M_6	5	-7484.16	-8072.29	-7901.38	-8035.29	-7233.42	-6302.92	-6480.73
M_7	4	-7497.99	-8083.37	-7919.96	-8049.09	-7252.83	-6310.25	-6477.22

5 Conclusions

This paper obtained an analytic expression for the likelihood of an inverse gamma SV model. As a result it is possible to obtain the Maximum Likelihood estimator. The exact value of the likelihood is also useful for Bayesian estimation and model comparison. Within the literature of nonlinear or non Gaussian state space models this novel approach is one of the very few methods that allow MLE because we are able to obtain the likelihood exactly. We provide the explicit formulas for this likelihood as well as the code to calculate it. Furthermore, we obtained the filtering and smoothing distributions for the inverse volatilities as mixture of gammas, allowing exact sampling from these distributions. Inverse gamma SV models can account for fat tails, which are observed in most macroeconomic and financial data. The approach that we use to obtain the likelihood expression is a result of integrating out the volatilities in the model. This approach is computationally efficient, simple and accurate.

The empirical fit of the inverse gamma SV model is better than other alternative models in the literature with inflation data for two countries and for 4 exchange rates series as shown in the empirical exercises.

References

Abramowitz, M., Stegun, I. A., & Romer, R. H. (1988). *Handbook of mathematical functions with formulas, graphs, and mathematical tables.* American Association of Physics Teachers.

Bauwens, L., Lubrano, M., & Richard, J.-F. (2000). *Bayesian inference in dynamic econometric models.* OUP Oxford.

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *The Review of Economics and Statistics*, *69*(3), 542–547.

Chan, J. C., Doucet, A., León-González, R., & Strachan, R. W. (2020). Multivariate stochastic volatility with co-heteroscedasticity. *GRIPS Discussion Papers*, *20-09*.

Chan, J. C., & Grant, A. L. (2016). Modeling energy price dynamics: GARCH versus stochastic volatility. *Energy Economics*, *54*, 182–189.

Creal, D. D. (2017). A class of non-gaussian state space models with exact likelihood inference. *Journal of Business & Economic Statistics*, *35*(4), 585–597.

Durbin, J., & Koopman, S. J. (2012). *Time series analysis by state space methods.* OUP Oxford.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the econometric society*, *50*(4), 987–1007.

Gouriéroux, C., & Jasiak, J. (2006). Autoregressive gamma processes. *Journal of Forecasting*, *25*(2), 129–152.

Kim, S., Shephard, N., & Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. *The Review of Economic Studies*, *65*(3), 361–393.

Leon-Gonzalez, R. (2019). Efficient bayesian inference in generalized inverse gamma processes for stochastic volatility. *Econometric Reviews*, *38*(8), 899-920.

Madan, D. B., & Seneta, E. (1990). The variance gamma model for share market returns. *Journal of Business*, *63*(4), 511–524.

Moral, P. (2004). *Feynman-kac formulae: genealogical and interacting particle systems with applications.* Springer.

Muirhead, R. J. (2005). *Aspects of multivariate statistical theory.* John Wiley & Sons.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464.

Shephard, N. (1994). Local scale models: State space alternative to integrated GARCH processes. *Journal of Econometrics*, *60*(1-2), 181–202.

Sundararajan, R. R., & Barreto-Souza, W. (2023). Student-t stochastic volatility model with composite likelihood EM-algorithm. *Journal of Time Series Analysis*, *44*, 125-147.

Uhlig, H. (1997). Bayesian vector autoregressions with stochastic volatility. *Econometrica: Journal of the Econometric Society*, *65*(1), 59–73.

West, M., & Harrison, J. (2006). *Bayesian forecasting and dynamic models.* Springer Science & Business Media.

6 Appendix

Proofs for the Lemmas and Proposition results in the paper are stated here.

6.1 Proof of Lemma

To derive the likelihood we will make use of the following lemma, which is a slightly modified version of Theorem 7.3.4 in Muirhead (2005).

Lemma 6.1. *For integers $p \leq q$*

$$\int |K|^{\frac{n+1-2}{2}} \exp\left(-\frac{1}{2}AK\right) {}_pF_q\left(a_1,...,a_p;b_1,...,b_q;\frac{1}{4}BK\right)dK = \Gamma\left(\frac{n+1}{2}\right)\left|\frac{1}{2}A\right|^{-\frac{n+1}{2}} \times \\ {}_{p+1}F_q\left(a_1,...,a_p,\frac{n+1}{2};b_1,...,b_q;\frac{1}{2}BA^{-1}\right)$$

where $(n+1)/2 > 0$ and ${}_pF_q(.)$ is a hypergeometric function of scalar argument, provided that in the case $p = q$ we have that $|0.5BA^{-1}| < 1$.

Proof. We apply Theorem 7.3.4 in Muirhead (2005) after making a change of variables. Let $X = \frac{1}{4}BK$ such that $K = 4XB^{-1}$. Thus we have:

$${}_pF_q\left(a_1,...,a_p;b_1,....,b_q;\frac{1}{4}BK\right) = {}_pF_q\left(a_1,...,a_p;b_1,...,b_q;X\right)$$

Therefore the integral becomes as follows:

$$\int |X|^{\frac{n+1-2}{2}}|4B^{-1}|^{\frac{n+1-2}{2}} \exp\Big(-\frac{1}{2}4AXB^{-1}\Big)_pF_q\Big(a_1,...,a_p;b_1,...,b_q;X\Big)dK$$

We use the Jacobian $dK = |4B^{-1}|dX$ to integrate with respect to X:

$$\int |X|^{\frac{n+1-2}{2}} \exp(-2XB^{-1}A)_pF_q\Big(a_1,...,a_p;b_1,...,b_q;X\Big)dX|4B^{-1}|^{\frac{n+1}{2}}$$

This integral is the same as in the theorem, therefore, when we integrate out X we get the following:

$$\int \frac{|X|^{\frac{n+1-2}{2}}}{|4B^{-1}|^{-\frac{n+1}{2}}} \exp(-X2B^{-1}A)_pF_q\Big(a_1,...,a_p;b_1,...,b_q;X\Big)dX = \Gamma\Big(\frac{n+1}{2}\Big)\Big|\frac{1}{2}A\Big|^{-\frac{n+1}{2}} \times \\ {}_{p+1}F_q\Big(a_1,...,a_p,\frac{n+1}{2};b_1,...,b_q;\frac{1}{2}BA^{-1}\Big)$$

□

6.2 Proof of Proposition 3.1

Proof. k_1 is a gamma, Bauwens et al. (2000) gives the prior density for k_1 as:

$$|k_1|^{\frac{n-2}{2}} \exp\Big(-\frac{1}{2}\big(k_1(1-\rho^2)\big)\Big)\frac{1}{c_0} \tag{6.1}$$

where $c_0 = \frac{\Gamma(\frac{n}{2})}{(\frac{1-\rho^2}{2})^{\frac{n}{2}}}$, is a constant and Γ is a gamma function. Let $V_1^{-1} = (1-\rho^2)$, thus, the likelihood for the first observation is as follows:

$$\begin{aligned} L(y_1) &= \int L(y_1 \mid k_1)\pi(k_1)dk_1 \\ &= \int (2\pi)^{-\frac{1}{2}}\sqrt{B^2}k_1^{\frac{1}{2}} \exp\Big(-\frac{1}{2}e_1^2B^2k_1\Big)k_1^{\frac{n-2}{2}} \exp\Big(-\frac{1}{2}(1-\rho^2)k_1\Big)\frac{1}{c_0}dk_1 \end{aligned} \tag{6.2}$$

The integral is with respect to k_1, so after rearranging and combining like terms we have;

$$L(y_1) = \int (2\pi)^{-\frac{1}{2}}\sqrt{B^2}k_1^{\frac{n+1-2}{2}} \exp\Big(-\frac{1}{2}(B^2e_1^2+V_1^{-1})k_1\Big)\frac{1}{c_0}dk_1$$

where $k_1^{\frac{n+1-2}{2}} \exp(-\frac{1}{2}(B^2e_1^2+V_1^{-1})k_1)$ is the kernel of a gamma with $n+1$ degrees of freedom. Let $\widetilde{V}_2 = (B^2e_1^2+V_1^{-1})^{-1}$, therefore, the density of $k_1|y_1$ is:

$$\pi(k_1|y_1) = k_1^{\frac{n+1-2}{2}} \exp\Big(-\frac{1}{2}k_1\widetilde{V}_2^{-1}\Big)\frac{1}{\bar{c}_0} \tag{6.3}$$

with $\bar{c_0} = \frac{\Gamma(\frac{n+1}{2})}{(\frac{\widetilde{V_2}^{-1}}{2})^{\frac{n+1}{2}}}$. Thus, we have the likelihood as:

$$L(y_1) = (2\pi)^{-\frac{1}{2}}\sqrt{B^2}\Gamma\bigg(\frac{n+1}{2}\bigg)\bigg|\frac{B^2e_1^2+V_1^{-1}}{2}\bigg|^{-\frac{n+1}{2}}\frac{1}{c_0}$$

Taking into account c_0 we can write the likelihood for $t=1$ as:

$$L(y_1) = (2\pi)^{-\frac{1}{2}}\sqrt{B^2}2^{\frac{1}{2}}\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})}\big|B^2e_1^2+V_1^{-1}\big|^{-\frac{n+1}{2}}V_1^{-\frac{n}{2}}$$

Define $k_{1:2} = (k_1, k_2)$, then we have the likelihood for the second observation as:

$$L(y_2|y_1) = \int L(y_2|k_{1:2}, y_1)\pi(k_{1:2}|y_1)dk_{1:2}$$

where $\pi(k_{1:2}|y_1) = \pi(k_1|y_1)\pi(k_2|k_1, y_1)$. The prior for k_t unconditionally is a gamma. However, $k_t|k_{t-1}$ is a non central chi-squared. Muirhead (2005, p. 442) gives this non central chi-squared density as follows:

$$\pi(k_t|k_{t-1}) = k_t^{\frac{n-2}{2}}\exp\bigg(-\frac{1}{2}k_t\bigg){}_0F_1\bigg(\frac{n}{2};\frac{1}{4}\rho^2k_{t-1}k_t\bigg)\exp\bigg(-\frac{1}{2}\rho^2k_{t-1}\bigg)\bigg(\Gamma\bigg(\frac{n}{2}\bigg)\bigg)^{-1}\frac{1}{c} \quad (6.4)$$

where ${}_0F_1$ is a hypergeometric function, ρ^2k_{t-1} is the non-centrality parameter and $c = 2^{\frac{n}{2}}$. We can then write the likelihood for the second observation given the first as :

$$L(y_2|y_1) = \int(2\pi)^{-\frac{1}{2}}\sqrt{B^2}k_2^{\frac{1}{2}}\exp\bigg(-\frac{1}{2}B^2e_2^2k_2\bigg)\pi(k_{1:2}|y_1)dk_{1:2} \quad (6.5)$$

We integrate first with respect to k_1. Define l_2 as representing all the elements in $\pi(k_2|k_1)$ as given by (6.4) that do not depend on k_1 as follows:

$$l_2 = \bigg(k_2^{\frac{n-2}{2}}\exp\bigg(-\frac{1}{2}k_2\bigg)\bigg)^{-1}\bigg(\frac{1}{\Gamma(\frac{n}{2})}\bigg)^{-1}\bigg(\frac{1}{c}\bigg)^{-1} \quad (6.6)$$

Given that $\pi(k_2|k_1, y_1) = \pi(k_2|k_1)$, and given (6.4) and (6.3), we can write $\pi(k_2|y_1)$ as follows:

$$\pi(k_2|y_1) = \int\pi(k_2|k_1, y_1)\pi(k_1|y_1)dk_1 =$$
$$\int\frac{1}{\bar{c_0}}k_1^{\frac{n+1-2}{2}}\exp\bigg(-\frac{1}{2}(\widetilde{V_2}^{-1}k_1)\bigg)\exp\bigg(-\frac{1}{2}(\rho^2k_1)\bigg){}_0F_1\bigg(\frac{n}{2};\frac{1}{4}\rho^2k_1k_2\bigg)\frac{1}{l_2}dk_1$$

where we have used the expression for $\pi(k_1|y_1)$ in (6.3). We can write the above integral more compactly as:

$$\int \pi(k_2|k_1,y_1)\pi(k_1|y_1)dk_1 = \int \frac{1}{\bar{c}_0} k_1^{\frac{n+1-2}{2}} \exp\bigg(-\frac{1}{2}(\widetilde{V_2}^{-1}+\rho^2)k_1\bigg)\,{}_0F_1\bigg(\frac{n}{2};\frac{1}{4}\rho^2 k_1 k_2\bigg)\frac{1}{l_2}dk_1$$

Applying Lemma 6.1 the solution to this integral is as follows:

$$\begin{aligned}\pi(k_2|y_1) = &\int \pi(k_2|k_1,y_1)\pi(k_1|y_1)dk_1 = \\ &\frac{1}{\bar{c}_0}\Gamma\bigg(\frac{n+1}{2}\bigg)\bigg|\frac{\widetilde{V_2}^{-1}+\rho^2}{2}\bigg|^{-\frac{n+1}{2}}\,{}_1F_1\bigg(\frac{n+1}{2};\frac{n}{2};\frac{1}{2}k_2\rho^2(\widetilde{V_2}^{-1}+\rho^2)^{-1}\bigg)\frac{1}{l_2}\end{aligned} \tag{6.7}$$

Given (6.6) and (6.7), the distribution of $k_2|y_1$ is a mixture of gammas as follows:

$$\pi(k_2|y_1) \propto k_2^{\frac{n-2}{2}} \exp\bigg(-\frac{1}{2}k_2\bigg)\,{}_1F_1\bigg(\frac{n+1}{2};\frac{n}{2};\frac{1}{2}k_2\rho^2(\widetilde{V_2}^{-1}+\rho^2)^{-1}\bigg) \tag{6.8}$$

The normalising constant for this density function can be obtained in closed form. Lemma 6.1 gives the solution to this integral, thus, we have:

$$\int k_2^{\frac{n-2}{2}} \exp\bigg(-\frac{1}{2}k_2\bigg)\,{}_1F_1\bigg(\frac{n+1}{2};\frac{n}{2};\frac{1}{2}k_2\delta_2\bigg)dk_2 = \Gamma\bigg(\frac{n}{2}\bigg)2^{\frac{n}{2}}\,{}_2F_1\bigg(\frac{n+1}{2},\frac{n}{2};\frac{n}{2};\delta_2\bigg) \tag{6.9}$$

where $\delta_2 = \rho^2(\widetilde{V_2}^{-1}+\rho^2)^{-1}$. This ${}_2F_1\big(\frac{n+1}{2},\frac{n}{2};\frac{n}{2};\delta_2\big)$ function has the same terms in the denominator and the numerator thus they cancel out and we have:

$${}_2F_1\bigg(\frac{n+1}{2},\frac{n}{2};\frac{n}{2};\delta_2\bigg) = {}_1F_0\bigg(\frac{n+1}{2};\delta_2\bigg) \tag{6.10}$$

Therefore, this function simplifies to a known solution for $|\delta_2| < 1$, see Muirhead (2005, p.261) .

$${}_1F_0\bigg(\frac{n+1}{2};\delta_2\bigg) = (1-\delta_2)^{-\frac{n+1}{2}} \tag{6.11}$$

Therefore the normalising constant becomes:

$$\Gamma\bigg(\frac{n}{2}\bigg)2^{\frac{n}{2}}\,{}_1F_0\bigg(\frac{n+1}{2};\delta_2\bigg) = \Gamma\bigg(\frac{n}{2}\bigg)2^{\frac{n}{2}}(1-\delta_2)^{-\frac{n+1}{2}}$$

Given this normalising constant, we have the density for $\pi(k_2|y_1)$ from 6.8 as follows:

$$\pi(k_2|y_1) = \frac{1}{c_1}k_2^{\frac{n-2}{2}} \exp\bigg(-\frac{1}{2}k_2\bigg)\,{}_1F_1\bigg(\frac{n+1}{2};\frac{n}{2};\frac{1}{2}k_2\rho^2(\widetilde{V_2}^{-1}+\rho^2)^{-1}\bigg)$$

where $c_1 = \Gamma\big(\frac{n}{2}\big)2^{\frac{n}{2}}(1-\delta_2)^{-\frac{n+1}{2}}$. Thus, the likelihood for the second observation is as follows:

$$L(y_2|y_1) = \int \pi(y_2|k_2, y_1)\pi(k_2|y_1)dk_2$$
$$= \int (2\pi)^{-\frac{1}{2}}\sqrt{B^2}k_2^{\frac{n+1-2}{2}} \exp\Big(-\frac{1}{2}(B^2e_2^2+1)k_2\Big)\frac{1}{c_1}{}_1F_1\Big(\frac{n+1}{2};\frac{n}{2};\frac{1}{2}k_2\rho^2(\widetilde{V}_2^{-1}+\rho^2)^{-1}\Big)dk_2$$

Using again Lemma 6.1 and taking into account c_1, the likelihood for the second observation is:

$$L(y_2|y_1) = (2\pi)^{-\frac{1}{2}}\sqrt{B^2}\frac{2^{\frac{n+1}{2}}}{2^{\frac{n}{2}}}\frac{\Gamma\big(\frac{n+1}{2}\big)}{\Gamma\big(\frac{n}{2}\big)}\frac{(B^2e_2^2+1)^{-\frac{n+1}{2}}}{(1-\delta_2)^{-\frac{n+1}{2}}}{}_2F_1\Big(\frac{n+1}{2},\frac{n+1}{2};\frac{n}{2};(B^2e_2^2+1)^{-1}\delta_2\Big)$$

Thus we get a Gauss hypergeometric function which can be evaluated easily. Let $Z_2 = (B^2e_2^2+1)^{-1}\delta_2$ and $\hat{C}_2 = {}_2F_1\big(\frac{n+1}{2},\frac{n+1}{2};\frac{n}{2};Z_2\big)$. This series converges because $|Z_2| < 1$ (Abramowitz et al., 1988). To accelerate the convergence of this series we apply the Euler transformation as in Abramowitz et al. (1988) and thus we get:

$${}_2F_1\Big(\frac{n+1}{2},\frac{n+1}{2};\frac{n}{2};Z_2\Big) = (1-Z_2)^{-\frac{n+2}{2}}{}_2F_1\Big(-\frac{1}{2},-\frac{1}{2};\frac{n}{2};Z_2\Big) \tag{6.12}$$

Thus $\hat{C}_2 = {}_2F_1\big(\frac{n+1}{2},\frac{n+1}{2};\frac{n}{2};Z_2\big) = (1-Z_2)^{-\frac{n+2}{2}}{}_2F_1\big(-\frac{1}{2},-\frac{1}{2};\frac{n}{2};Z_2\big)$, then we can write the $L(y_2|y_1)$ as follows:

$$L(y_2|y_1) = (2\pi)^{-\frac{1}{2}}\sqrt{B^2}\frac{2^{\frac{n+1}{2}}}{2^{\frac{n}{2}}}\frac{\Gamma\big(\frac{n+1}{2}\big)}{\Gamma\big(\frac{n}{2}\big)}\frac{(B^2e_2^2+1)^{-\frac{n+1}{2}}}{(1-\delta_2)^{-\frac{n+1}{2}}}\hat{C}_2$$

The density of k_t for the third observation is given by:

$$\pi(k_3|y_2, y_1) = \int \pi(k_3|k_2)\pi(k_2|y_2, y_1)dk_2$$

where $\pi(k_2|y_2, y_1) \propto \pi(k_2|y_1)L(y_2|k_2, y_1)$. The distribution for $\pi(k_2|y_1)$ in (6.8) can be written as follows:

$$\pi(k_2|y_1) \propto \sum_{h_2=0}^{\infty}\widetilde{C}_{2,h_2}k_2^{\frac{n+2h_2-2}{2}}\exp\Big(-\frac{1}{2}k_2\Big)$$

where $\widetilde{C}_{2,h_2} = \frac{[(n+1)/2]_{h_2}}{[n/2]_{h_2}}\big(\frac{1}{2}\rho^2(\widetilde{V}_2^{-1}+\rho^2)^{-1}\big)^{h_2}\frac{1}{h_2!}$. Thus we have:

$$\pi(k_2|y_2, y_1) \propto \sum_{h_2=0}^{\infty}\widetilde{C}_{2,h_2}k_2^{\frac{n+1+2h_2-2}{2}}\exp\Big(-\frac{1}{2}k_2(B^2e_2^2+1)\Big) \tag{6.13}$$

Given (6.4) and (6.13) we have:

$$\pi(k_3|y_2,y_1) \propto \int k_3^{\frac{n-2}{2}} \exp\Big(-\frac{1}{2}k_3\Big) {}_0F_1\Big(\frac{n}{2};\frac{1}{4}\rho^2 k_2 k_3\Big) \exp\Big(-\frac{1}{2}\rho^2 k_2\Big)$$
$$\times \sum_{h_2=0}^{\infty} \widetilde{C}_{2,h_2} k_2^{\frac{n+1+2h_2-2}{2}} \exp\Big(-\frac{1}{2}k_2(B^2e_2^2+1)\Big)\frac{1}{\Gamma\big(\frac{n}{2}\big)2^{\frac{n}{2}}}dk_2$$

which simplifies to:

$$\pi(k_3|y_2,y_1) \propto \int k_3^{\frac{n-2}{2}} \exp\Big(-\frac{1}{2}k_3\Big) {}_0F_1\Big(\frac{n}{2};\frac{1}{4}\rho^2 k_2 k_3\Big) \exp\Big(-\frac{1}{2}(B^2e_2^2+1+\rho^2)k_2\Big)$$
$$\sum_{h_2=0}^{\infty} \widetilde{C}_{2,h_2} k_2^{\frac{n+1+2h_2-2}{2}} \frac{1}{\Gamma\big(\frac{n}{2}\big)2^{\frac{n}{2}}}dk_2$$

Using Lemma 6.1 the density of $k_3|y_2,y_1$ is thus:

$$\begin{aligned}\pi(k_3|y_2,y_1) = \frac{1}{c_3}k_3^{\frac{n-2}{2}} \exp\Big(-\frac{1}{2}k_3\Big) \sum_{h_2=0}^{\infty} \widetilde{C}_{2,h_2}\Gamma\Big(\frac{n+1+2h_2}{2}\Big)\\ {}_1F_1\Big(\frac{n+1+2h_2}{2};\frac{n}{2};\frac{1}{2}k_3\rho^2 S_3\Big)(2S_3)^{\frac{n+1+2h_2}{2}}\frac{1}{\Gamma\big(\frac{n}{2}\big)2^{\frac{n}{2}}}\end{aligned} \tag{6.14}$$

where $S_3 = (B^2e_2^2+1+\rho^2)^{-1}$ and c_3 is the normalising constant as in (6.9) as follows:

$$c_3 = \sum_{h_2=0}^{\infty} \widetilde{C}_{2,h_2}\Gamma\Big(\frac{n+1+2h_2}{2}\Big)(2S_3)^{\frac{n+1+2h_2}{2}} {}_2F_1\Big(\frac{n+1+2h_2}{2},\frac{n}{2};\frac{n}{2};\rho^2 S_3\Big)$$

Similar to (6.10) and (6.11), the hypergeometric function simplifies to get:

$$c_3 = \sum_{h_2=0}^{\infty} \widetilde{C}_{2,h_2}\Gamma\Big(\frac{n+1+2h_2}{2}\Big)(2S_3)^{\frac{n+1+2h_2}{2}}(1-\rho^2S_3)^{-\frac{n+1+2h_2}{2}}$$

Collecting terms dependent on h_2 we can write c_3 as

$$c_3 = \Bigg(\sum_{h_2=0}^{\infty} \frac{[(n+1)/2]_{h_2}}{[n/2]_{h_2}}[(n+1)/2]_{h_2}\frac{\delta_3^{h_2}}{h_2!}\Bigg)\Gamma\Big(\frac{n+1}{2}\Big)(1-\rho^2S_3)^{-\frac{n+1}{2}}(2S_3)^{\frac{n+1}{2}}$$

where $\delta_3 = \big((1-\rho^2S_3)^{-1}S_3\rho^2(\widetilde{V}_2^{-1}+\rho^2)^{-1}\big)$. This can be written as:

$$c_3 = {}_2F_1\Big(\frac{n+1}{2},\frac{n+1}{2};\frac{n}{2};\delta_3\Big)\Gamma\Big(\frac{n+1}{2}\Big)(1-\rho^2S_3)^{-\frac{n+1}{2}}(2S_3)^{\frac{n+1}{2}}$$

Using Euler's acceleration in (6.12) we can transform c_3 as:

$$c_3 = (1-\delta_3)^{-\frac{n+2}{2}} {}_2F_1\left(-\frac{1}{2}, -\frac{1}{2}; \frac{n}{2}; \delta_3\right)\Gamma\left(\frac{n+1}{2}\right)(1-\rho^2 S_3)^{-\frac{n+1}{2}}(2S_3)^{\frac{n+1}{2}}$$

Therefore the likelihood for $t = 3$ is as follows:

$$L(y_3|y_2, y_1) = \int \pi(y_3|k_3, y_2, y_1)\pi(k_3|y_2, y_1)dk_3$$

Thus we have from (6.14)

$$L(y_3|y_2, y_1) = \int (2\pi)^{-\frac{1}{2}}\sqrt{B^2}\frac{1}{c_3}k_3^{\frac{n+1-2}{2}} \exp\left(-\frac{1}{2}k_3(B^2e_3^2+1)\right)\sum_{h_2=0}^{\infty}\widetilde{C}_{2,h_2}\Gamma\left(\frac{n+1+2h_2}{2}\right)$$
$${}_1F_1\left(\frac{n+1+2h_2}{2}; \frac{n}{2}; \frac{1}{2}k_3\rho^2 S_3\right)(2S_3)^{\frac{n+1+2h_2}{2}}\frac{1}{\Gamma\left(\frac{n}{2}\right)2^{\frac{n}{2}}}dk_3$$

and using Lemma 6.1 we get:

$$L(y_3|y_2, y_1) = (2\pi)^{-\frac{1}{2}}\sqrt{B^2}\frac{1}{c_3}\sum_{h_2=0}^{\infty}\widetilde{C}_{2,h_2}\Gamma\left(\frac{n+1+2h_2}{2}\right)(2S_3)^{\frac{n+1+2h_2}{2}}\Gamma\left(\frac{n+1}{2}\right)2^{\frac{n+1}{2}}$$
$$(B^2e_3^2+1)^{-\frac{n+1}{2}}{}_2F_1\left(\frac{n+1+2h_2}{2}, \frac{n+1}{2}; \frac{n}{2}; (B^2e_3^2+1)^{-1}\rho^2 S_3\right)\frac{1}{\Gamma\left(\frac{n}{2}\right)2^{\frac{n}{2}}}$$

Letting $Z_3 = (B^2e_3^2+1)^{-1}\rho^2 S_3$, we can define $\hat{C}_3 = {}_2F_1\left(\frac{n+1+2h_2}{2}, \frac{n+1}{2}; \frac{n}{2}; Z_3\right)$. Thus, we have:

$$L(y_3|y_2, y_1) = (2\pi)^{-\frac{1}{2}}\sqrt{B^2}\frac{1}{c_3}\sum_{h_2=0}^{\infty}\widetilde{C}_{2,h_2}\frac{\Gamma\left(\frac{n+1+2h_2}{2}\right)}{(B^2e_3^2+1)^{\frac{n+1}{2}}}(2S_3)^{\frac{n+1+2h_2}{2}}\frac{2^{\frac{n+1}{2}}}{2^{\frac{n}{2}}}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\hat{C}_3$$

The filtering density of k_t for $t = 4$ is given by:

$$\pi(k_4|y_3, y_2, y_1) = \int \pi(k_4|k_3, y_1, y_2, y_3)\pi(k_3|y_3, y_2, y_1)dk_3 \qquad (6.15)$$

where $\pi(k_3|y_3, y_2, y_1) \propto \pi(k_3|y_2, y_1)L(y_3|k_3, y_2, y_1)$. Let:

$$\widetilde{C}_{3,h_3} = \sum_{h_2=0}^{\infty}\widetilde{C}_{2,h_2}\Gamma\left(\frac{n+1+2h_2}{2}\right)\frac{[(n+1)/2+h_2]_{h_3}}{[n/2]_{h_3}}\left(\frac{1}{2}\rho^2 S_3\right)^{h_3}\frac{1}{h_3!}(2S_3)^{\frac{n+1+2h_2}{2}} \qquad (6.16)$$

Then from (6.14) we have that the filtering distribution $k_3|y_2, y_1$ is a mixture of gammas as follows:

$$\pi(k_3|y_2, y_1) \propto \sum_{h_3=0}^{\infty} \widetilde{C}_{3,h_3} k_3^{\frac{n+2h_3-2}{2}} \exp\Big(-\frac{1}{2}k_3\Big)$$

As before, when we include the third observation, the distribution of $k_3|y_3, y_2, y_1$ is also a mixture of gammas and can be written as follows:

$$\pi(k_3|y_3, y_2, y_1) \propto \sum_{h_3=0}^{\infty} \widetilde{C}_{3,h_3} k_3^{\frac{n+1+2h_3-2}{2}} \exp\Big(-\frac{1}{2}k_3(B^2e_3^2+1)\Big)$$

Let $\widetilde{V_4}^{-1} = (B^2e_3^2 + 1)$. Then, using (6.15) and (6.4), we have the distribution of $k_4|y_3, y_2, y_1$ as follows:

$$\begin{aligned}\pi(k_4|y_3, y_2, y_1) \propto \int k_4^{\frac{n-2}{2}} \exp\Big(-\frac{1}{2}k_4\Big) {}_0F_1\Big(\frac{n}{2}; \frac{1}{4}\rho^2 k_3 k_4\Big) \exp\Big(-\frac{1}{2}\rho^2 k_3\Big) \frac{1}{\Gamma(\frac{n}{2})2^{\frac{n}{2}}} \\ \times \sum_{h_3=0}^{\infty} \widetilde{C}_{3,h_3} k_3^{\frac{n+1+2h_3-2}{2}} \exp\Big(-\frac{1}{2}k_3\widetilde{V_4}^{-1}\Big) dk_3\end{aligned} \tag{6.17}$$

Taking this integral with respect to k_3 we get:

$$\begin{aligned}\pi(k_4|y_3, y_2, y_1) \propto k_4^{\frac{n-2}{2}} \exp\Big(-\frac{1}{2}k_4\Big) \sum_{h_3=0}^{\infty} \widetilde{C}_{3,h_3} {}_1F_1\Big(\frac{n+1+2h_3}{2}; \frac{n}{2}; \frac{1}{2}\rho^2 k_4 (\widetilde{V_4}^{-1} + \rho^2)^{-1}\Big) \\ \Gamma\Big(\frac{n+1+2h_3}{2}\Big)(2S_4)^{\frac{n+1+2h_3}{2}} \frac{1}{\Gamma(\frac{n}{2})2^{\frac{n}{2}}}\end{aligned}$$

where $S_4 = (\widetilde{V_4}^{-1} + \rho^2)^{-1} = (B^2e_3^2 + 1 + \rho^2)^{-1}$. Let c_4 be the normalising constant, that is:

$$\begin{aligned}c_4 = \int k_4^{\frac{n-2}{2}} \exp\Big(-\frac{1}{2}k_4\Big) \sum_{h_3=0}^{\infty} \widetilde{C}_{3,h_3} {}_1F_1\Big(\frac{n+1+2h_3}{2}; \frac{n}{2}; \frac{1}{2}\rho^2 k_4 (\widetilde{V_4}^{-1} + \rho^2)^{-1}\Big) \\ \Gamma\Big(\frac{n+1+2h_3}{2}\Big)(2S_4)^{\frac{n+1+2h_3}{2}} \frac{1}{\Gamma(\frac{n}{2})2^{\frac{n}{2}}} dk_4\end{aligned}$$

Thus we get:

$$c_4 = \sum_{h_3=0}^{\infty} \widetilde{C}_{3,h_3} {}_2F_1\Big(\frac{n+1+2h_3}{2}, \frac{n}{2}; \frac{n}{2}; \rho^2 S_4\Big) \Gamma\Big(\frac{n+1+2h_3}{2}\Big)(2S_4)^{\frac{n+1+2h_3}{2}}$$

Using (6.10) and (6.11), this simplifies to:

$$c_4 = \sum_{h_3=0}^{\infty} \widetilde{C}_{3,h_3} (1-\rho^2 S_4)^{-\frac{n+1+2h_3}{2}} \Gamma\bigg(\frac{n+1+2h_3}{2}\bigg)(2S_4)^{\frac{n+1+2h_3}{2}}$$

Thus,

$$\pi(k_4|y_3,y_2,y_1) = \frac{1}{c_4} k_4^{\frac{n-2}{2}} \exp\bigg(-\frac{1}{2}k_4\bigg) \sum_{h_3=0}^{\infty} \widetilde{C}_{3,h_3}{}_1F_1\bigg(\frac{n+1+2h_3}{2};\frac{n}{2};\frac{1}{2}\rho^2 k_4 (\widetilde{V}_4^{-1}+\rho^2)^{-1}\bigg)$$
$$\Gamma\bigg(\frac{n+1+2h_3}{2}\bigg)(2S_4)^{\frac{n+1+2h_3}{2}} \frac{1}{\Gamma\big(\frac{n}{2}\big)2^{\frac{n}{2}}}$$

Therefore the likelihood for $t=4$ is as follows:

$$L(y_4|y_3,y_2,y_1) = \int \pi(y_4|k_4,y_3,y_2,y_1)\pi(k_4|y_3,y_2,y_1)dk_4$$

Thus we have:

$$L(y_4|y_3,y_2,y_1) = \int (2\pi)^{-\frac{1}{2}}\sqrt{B^2}\frac{1}{c_4} k_4^{\frac{n+1-2}{2}} \exp\bigg(-\frac{1}{2}k_4(B^2e_4^2+1)\bigg) \sum_{h_3=0}^{\infty} \widetilde{C}_{3,h_3}\Gamma\bigg(\frac{n+1+2h_3}{2}\bigg)$$
$${}_1F_1\bigg(\frac{n+1+2h_3}{2};\frac{n}{2};\frac{1}{2}k_4\rho^2 S_4\bigg)(2S_4)^{\frac{n+1+2h_3}{2}}\frac{1}{\Gamma\big(\frac{n}{2}\big)2^{\frac{n}{2}}}dk_4$$

This is similar to $t=3$ therefore we have:

$$L(y_4|y_3,y_2,y_1) = (2\pi)^{-\frac{1}{2}}\sqrt{B^2}\frac{1}{c_4}\sum_{h_3=0}^{\infty}\widetilde{C}_{3,h_3}\frac{\Gamma\Big(\frac{n+1+2h_3}{2}\Big)}{(B^2e_4^2+1)^{\frac{n+1}{2}}}(2S_4)^{\frac{n+1+2h_3}{2}}\frac{2^{\frac{n+1}{2}}}{2^{\frac{n}{2}}}\frac{\Gamma\big(\frac{n+1}{2}\big)}{\Gamma\big(\frac{n}{2}\big)}\hat{C}_4$$

and the likelihood for any t is:

$$L(y_t|y_{1:t-1}) = (2\pi)^{-\frac{1}{2}}\sqrt{B^2}\frac{1}{c_t}\sum_{h_{t-1}=0}^{\infty}\widetilde{C}_{t-1,h_{t-1}}\frac{\Gamma\Big(\frac{n+1+2h_{t-1}}{2}\Big)}{(B^2e_t^2+1)^{\frac{n+1}{2}}}(2S_t)^{\frac{n+1+2h_{t-1}}{2}}\frac{2^{\frac{n+1}{2}}}{2^{\frac{n}{2}}}\frac{\Gamma\big(\frac{n+1}{2}\big)}{\Gamma\big(\frac{n}{2}\big)}\hat{C}_t$$

where for $t \geq 4$:

$$
\begin{aligned}
\delta_t &= \left((1-\rho^2 S_t)^{-1} S_t \rho^2 (\widetilde{V}_{t-1}^{-1} + \rho^2)^{-1}\right) \\
Z_t &= (B^2 e_t^2 + 1)^{-1} S_t \rho^2 \\
\hat{C}_t &= {}_2F_1\left(\frac{n+1+2h_{t-1}}{2}, \frac{n+1}{2}; \frac{n}{2}; Z_t\right) \\
\widetilde{V}_t^{-1} &= 1 + B^2 e_{t-1}^2 \\
S_t &= (B^2 e_{t-1}^2 + 1 + \rho^2)^{-1} = (\widetilde{V}_t^{-1} + \rho^2)^{-1} \\
c_t &= \sum_{h_{t-1}=0}^{\infty} \widetilde{C}_{t-1,h_{t-1}} (1-\rho^2 S_t)^{-\frac{n+1+2h_{t-1}}{2}} \Gamma\left(\frac{n+1+2h_{t-1}}{2}\right) (2S_t)^{\frac{n+1+2h_{t-1}}{2}}
\end{aligned}
$$

$$
\widetilde{C}_{t-1,h_{t-1}} =
$$

$$
\sum_{h_{t-2}=0}^{\infty} \widetilde{C}_{t-2,h_{t-2}} \Gamma\left(\frac{n+1+2h_{t-2}}{2}\right) \frac{[(n+1)/2 + h_{t-2}]_{h_{t-1}}}{[n/2]_{h_{t-1}}} \left(\frac{1}{2}\rho^2 S_{t-1}\right)^{h_{t-1}} \frac{(2S_{t-1})^{\frac{n+1+2h_{t-2}}{2}}}{h_{t-1}!}
$$

$\square$

6.3 Proof of Proposition 3.2

Proof. Combining the prior density for k_1 in (6.1) with the transition equation in (6.4) and the likelihood, we get:

$$
\begin{aligned}
\pi(k_1|k_{2:T}, y_{1:T}) &\propto |k_1|^{\frac{n+1-2}{2}} \exp\left(-\frac{1}{2}S_2^{-1}k_1\right) {}_0F_1\left(\frac{n}{2}; \frac{1}{4}\rho^2 k_1 k_2\right) \\
&= |k_1|^{\frac{n+1-2}{2}} \exp\left(-\frac{1}{2}S_2^{-1}k_1\right) \sum_{h=0}^{\infty} \left(C_{1,h}|k_1|^h\right)
\end{aligned}
\tag{6.18}
$$

with $C_{1,h} = \frac{1}{h!}\frac{1}{[n/2]_h}\left(\frac{1}{4}\rho^2 k_2\right)^h$.

The integral of (6.18) with respect to k_1 is proportional to:

$$
{}_1F_1\left(\frac{n+1}{2}; \frac{n}{2}; \frac{1}{2}\rho^2 k_2 S_2\right)
$$

and therefore:

$$
\begin{aligned}
&\pi(k_2|k_{3:T}, y_{1:T}) \\
&\propto |k_2|^{\frac{n+1-2}{2}} \exp\left(-\frac{1}{2}S_3^{-1}k_2\right) {}_1F_1\left(\frac{n+1}{2}; \frac{n}{2}; \frac{1}{2}\rho^2 k_2 S_2\right) {}_0F_1\left(\frac{n}{2}; \frac{1}{4}\rho^2 k_3 k_2\right)
\end{aligned}
\tag{6.19}
$$

where we have used that $S_3^{-1} = 1 + B^2 e_2^2 + \rho^2$. Combining the series we get that:

$$
\begin{aligned}
&{}_1F_1\left(\frac{n+1}{2};\frac{n}{2};\frac{1}{2}\rho^2 k_2 S_2\right){}_0F_1\left(\frac{n}{2};\frac{1}{4}\rho^2 k_3 k_2\right) = \\
&\left(\sum_{h_1=0}^{\infty}\frac{[(n+1)/2]_{h_1}}{[n/2]_{h_1}}\frac{(\frac{1}{2}\rho^2 S_2)^{h_1}k_2^{h_1}}{h_1!}\right)\left(\sum_{h_2=0}^{\infty}\frac{1}{h_2!}\frac{1}{[n/2]_{h_2}}\left(\frac{1}{4}\rho^2 k_3\right)^{h_2}k_2^{h_2}\right)
\end{aligned}
\tag{6.20}
$$

By making the change of variables $h = h_1 + h_2$ we get that (6.20) can be written as:

$$
\sum_{h=0}^{\infty}\sum_{h_2=0}^{h}\left(\left(\frac{[(n+1)/2]_{h-h_2}}{[n/2]_{h-h_2}}\frac{(\frac{1}{2}\rho^2 S_2)^{h-h_2}}{(h-h_2)!}\right)\frac{1}{h_2!}\frac{1}{[n/2]_{h_2}}\left(\frac{1}{4}\rho^2\right)^{h_2}k_3^{h_2}\right)k_2^h = \sum_{h=0}^{\infty}C_{2,h}k_2^h \tag{6.21}
$$

where:

$$
C_{2,h} = \sum_{h_2=0}^{h}\widetilde{C}_{2,h-h_2}\frac{1}{h_2!}\frac{1}{[n/2]_{h_2}}\left(\frac{1}{4}\rho^2\right)^{h_2}k_3^{h_2}
$$

and $\widetilde{C}_{2,h-h_2}$ has been defined in proposition 3.1 as:

$$
\widetilde{C}_{2,h-h_2} = \frac{[(n+1)/2]_{h-h_2}}{[n/2]_{h-h_2}}\frac{(\frac{1}{2}\rho^2 S_2)^{h-h_2}}{(h-h_2)!}
$$

Using (6.21) we obtain that:

$$
\pi(k_2|k_{3:T}, y_{1:T}) \propto |k_2|^{\frac{n+1-2}{2}}\exp\left(-\frac{1}{2}S_3^{-1}k_2\right)\sum_{h=0}^{\infty}\left(C_{2,h}k_2^h\right) \tag{6.22}
$$

as we wanted to prove.

The integral of (6.22) with respect to k_2 is proportional to:

$$
\sum_{h=0}^{\infty}\left(C_{2,h}\frac{\Gamma\left(\frac{n+1+2h}{2}\right)}{(S_3^{-1}/2)^{\frac{n+1+2h}{2}}}\right) = \sum_{h=0}^{\infty}\left(\sum_{h_2=0}^{h}\widetilde{C}_{2,h-h_2}\frac{1}{h_2!}\frac{1}{[n/2]_{h_2}}\left(\frac{1}{4}\rho^2\right)^{h_2}k_3^{h_2}\right)\frac{\Gamma\left(\frac{n+1+2h}{2}\right)}{(S_3^{-1}/2)^{\frac{n+1+2h}{2}}} \tag{6.23}
$$

Making the change of variables $h_1 = h - h_2$, equation (6.23) can be written as:

$$
\sum_{h_1=0}^{\infty}\sum_{h_2=0}^{\infty}\left(\widetilde{C}_{2,h_1}\frac{1}{h_2!}\frac{1}{[n/2]_{h_2}}\left(\frac{1}{4}\rho^2\right)^{h_2}k_3^{h_2}\right)\frac{\Gamma\left(\frac{n+1}{2}+h_1+h_2\right)}{(S_3^{-1}/2)^{\frac{n+1}{2}+h_1+h_2}} \tag{6.24}
$$

Note that $\Gamma\left(\frac{n+1}{2}+h_1+h_2\right) = \Gamma\left(\frac{n+1+2h_1}{2}\right)\left[\frac{n+1+2h_1}{2}\right]_{h_2}$. Then (6.24) can be written as:

$$
\sum_{h_2=0}^{\infty}\sum_{h_1=0}^{\infty}\widetilde{C}_{2,h_1}\Gamma\left(\frac{n+1+2h_1}{2}\right)\frac{\left[(n+1)/2+h_1\right]_{h_2}}{[n/2]_{h_2}}\left(\frac{1}{2}\rho^2 S_3\right)^{h_2}\frac{1}{h_2!}(2S_3)^{\frac{n+1+2h_1}{2}}k_3^{h_2} \tag{6.25}
$$

Using the definition of $\widetilde{C}_{3,h_2}$ in proposition 3.1, we can write (6.25) as:

$$\sum_{h_2=0}^{\infty} \widetilde{C}_{3,h_2} k_3^{h_2}$$

Recall that the transition density is in (6.4). Therefore, we have:

$$\pi(k_3|k_{4:T}, y_{1:T}) \propto \left(\sum_{h_2=0}^{\infty} \widetilde{C}_{3,h_2} k_3^{h_2} \right) {}_0F_1\left(\frac{n}{2}; \frac{1}{4}\rho^2 k_3 k_4\right) |k_3|^{\frac{n+1-2}{2}} \exp\left(-\frac{1}{2} S_4^{-1} k_3 \right)$$

with $S_4^{-1} = 1 + B^2 e_3^2 + \rho^2$. As before, we can multiply the two series as follows:

$$\left(\sum_{h_2=0}^{\infty} \widetilde{C}_{3,h_2} k_3^{h_2} \right) {}_0F_1\left(\frac{n}{2}; \frac{1}{4}\rho^2 k_3 k_4\right) = \left(\sum_{h_2=0}^{\infty} \widetilde{C}_{3,h_2} k_3^{h_2} \right) \left(\sum_{h_3=0}^{\infty} \frac{1}{[n/2]_{h_3}} \left(\frac{1}{4}\rho^2 k_3\right)^{h_3} k_4^{h_3} \frac{1}{h_3!} \right)$$
$$= \sum_{h=0}^{\infty} \sum_{h_3=0}^{h} |k_3|^h \widetilde{C}_{3,h-h_3} \frac{1}{[n/2]_{h_3}} \left(\frac{1}{4}\rho^2\right)^{h_3} k_4^{h_3} \frac{1}{h_3!} = \sum_{h=0}^{\infty} |k_3|^h C_{3,h}$$

where

$$C_{3,h} = \sum_{h_3=0}^{\infty} \widetilde{C}_{3,h-h_3} \frac{1}{[n/2]_{h_3}} \left(\frac{1}{4}\rho^2\right)^{h_3} \frac{k_4^{h_3}}{h_3!}$$

and therefore, $\pi(k_3|k_{4:T}, y_{1:T})$ can be written as:

$$\pi(k_3|k_{4:T}, y_{1:T}) \propto |k_3|^{\frac{n+1-2}{2}} \exp\left(-\frac{1}{2} S_4^{-1} k_3 \right) \sum_{h=0}^{\infty} |k_3|^h C_{3,h} \tag{6.26}$$

as we wanted to prove.

Since $\pi(k_3|k_{4:T}, y_{1:T})$ in (6.26) and $\pi(k_2|k_{3:T}, y_{1:T})$ in (6.22) have the same structure, and, since the transition density of k_t is always the same, we get analogous results for any $t < T$, as we wanted to prove. For $t = T$ the only difference is that there is no transition density from k_T to k_{T+1}. For this reason we do not need to multiply two series, and hence $C_{T,h} = \widetilde{C}_{T,h}$ and $S_{T+1} = (1 + B^2 e_T^2)^{-1}$

□

6.4 Proof of Proposition 3.3

Proof. We need to integrate $\pi(k_{1:T})\pi(y_{1:T}|k_{1:T})$ with respect to k_T first. The terms that depend on k_T are the following:

$$\exp\Bigg(-\frac{1}{2}e_T^2B^2k_T\Bigg)|k_T|^{\frac{1}{2}}|k_T|^{\frac{n-2}{2}}\exp\Bigg(-\frac{1}{2}k_T\Bigg){}_0F_1\Bigg(\frac{n}{2};\frac{1}{4}\rho^2k_Tk_{T-1}\Bigg) = \exp\Bigg(-\frac{1}{2}S_{T+1}^{-1}k_T\Bigg)|k_T|^{\frac{n+1-2}{2}}\sum_{h=0}^{\infty}a_{T,h}|k_T|^h \tag{6.27}$$

with $a_{T,h}=\frac{1}{h!}\frac{1}{[n/2]_h}\big(\frac{1}{4}\rho^2k_{T-1}\big)^h$. This proves the result for $s=0$. The integral of (6.27) with respect to k_T is proportional to:

$${}_1F_1\Bigg(\frac{n+1}{2};\frac{n}{2};\frac{1}{2}\rho^2k_{T-1}S_{T+1}\Bigg)$$

Therefore, the terms that depend on k_{T-1} in $\pi(k_{1:T})\pi(y_{1:T}|k_{1:T})$ after integrating out k_T are the following:

$$|k_{T-1}|^{\frac{n+1-2}{2}}\exp\Bigg(-\frac{1}{2}S_T^{-1}k_{T-1}\Bigg){}_1F_1\Bigg(\frac{n+1}{2};\frac{n}{2};\frac{1}{2}\rho^2k_{T-1}S_{T+1}\Bigg){}_0F_1\Bigg(\frac{n}{2};\frac{1}{4}\rho^2k_{T-1}k_{T-2}\Bigg) \tag{6.28}$$

Equation (6.28) has the product of two series, that can be written as:

$$\begin{aligned}&{}_1F_1\Bigg(\frac{n+1}{2};\frac{n}{2};\frac{1}{2}\rho^2k_{T-1}S_{T+1}\Bigg){}_0F_1\Bigg(\frac{n}{2};\frac{1}{4}\rho^2k_{T-1}k_{T-2}\Bigg)=\\&=\Bigg(\sum_{h_T=0}^{\infty}\frac{[(n+1)/2]_{h_T}}{[n/2]_{h_T}}\frac{(\frac{1}{2}\rho^2S_{T+1})^{h_T}k_{T-1}^{h_T}}{h_T!}\Bigg)\Bigg(\sum_{h_{T-1}=0}^{\infty}\frac{1}{(h_{T-1})!}\frac{1}{[n/2]_{h_{T-1}}}\Bigg(\frac{1}{4}\rho^2k_{T-2}\Bigg)^{h_{T-1}}k_{T-1}^{h_{T-1}}\Bigg)\end{aligned} \tag{6.29}$$

Making the change of variables $h=h_T+h_{T-1}$ we get that (6.29) is equal to:

$$\sum_{h=0}^{\infty}\Bigg(\sum_{h_{T-1}=0}^{h}\frac{[(n+1)/2]_{h-h_{T-1}}}{[n/2]_{h-h_{T-1}}}\frac{(\frac{1}{2}\rho^2S_{T+1})^{h-h_{T-1}}}{(h-h_{T-1})!}\frac{1}{(h_{T-1})!}\frac{1}{[n/2]_{h_{T-1}}}\Bigg(\frac{1}{4}\rho^2\Bigg)^{h_{T-1}}k_{T-2}^{h_{T-1}}\Bigg)k_{T-1}^h = \sum_{h=0}^{\infty}a_{T-1,h}k_{T-1}^h$$

where:

$$a_{T-1,h}=\sum_{h_{T-1}=0}^{h}\Bigg(\frac{[(n+1)/2]_{h-h_{T-1}}}{[n/2]_{h-h_{T-1}}}\frac{(\frac{1}{2}\rho^2S_{T+1})^{h-h_{T-1}}}{(h-h_{T-1})!}\frac{1}{(h_{T-1})!}\frac{1}{[n/2]_{h_{T-1}}}\Bigg(\frac{1}{4}\rho^2\Bigg)^{h_{T-1}}k_{T-2}^{h_{T-1}}\Bigg)$$

which can be written as:

$$a_{T-1,h} = \sum_{h_{T-1}=0}^{h} \widetilde{a}_{T-1,h-h_{T-1}} \frac{1}{(h_{T-1})!} \frac{1}{[n/2]_{h_{T-1}}} \left(\frac{1}{4}\rho^2\right)^{h_{T-1}} k_{T-2}^{h_{T-1}}$$

and:

$$\widetilde{a}_{T-1,h} = \frac{[(n+1)/2]_h}{[n/2]_h} \frac{(\frac{1}{2}\rho^2 S_{T+1})^h}{h!}$$

Therefore, $\pi(k_{T-1}|k_{1:T-2}, y_{1:T})$ which is given by (6.28), can be written as:

$$\pi(k_{T-1}|k_{1:T-2}, y_{1:T}) \propto |k_{T-1}|^{\frac{n+1-2}{2}} \exp\left(-\frac{1}{2} S_T^{-1} k_{T-1}\right) \sum_{h=0}^{\infty} \left(a_{T-1,h} k_{T-1}^h\right) \tag{6.30}$$

which proves the result for $s = 1$.

The integral of (6.30) with respect to k_{T-1} gives:

$$\begin{aligned} &\sum_{h=0}^{\infty} \left(a_{T-1,h} \frac{\Gamma\left(\frac{n+1+2h}{2}\right)}{(S_T^{-1}/2)^{\frac{n+1+2h}{2}}}\right) = \\ &= \sum_{h=0}^{\infty} \sum_{h_{T-1}=0}^{h} \widetilde{a}_{T-1,h-h_{T-1}} \frac{1}{(h_{T-1})!} \frac{1}{[n/2]_{h_{T-1}}} \left(\frac{1}{4}\rho^2\right)^{h_{T-1}} k_{T-2}^{h_{T-1}} \frac{\Gamma\left(\frac{n+1+2h}{2}\right)}{(S_T^{-1}/2)^{\frac{n+1+2h}{2}}} \end{aligned} \tag{6.31}$$

Making a change of variables $h = h_T + h_{T-1}$, equation (6.31) can be written as:

$$\sum_{h_T=0}^{\infty} \sum_{h_{T-1}=0}^{\infty} \left(\widetilde{a}_{T-1,h_T} \frac{1}{(h_{T-1})!} \frac{1}{[n/2]_{h_{T-1}}} \left(\frac{1}{4}\rho^2\right)^{h_{T-1}} k_{T-2}^{h_{T-1}}\right) \frac{\Gamma\left(\frac{n+1}{2} + h_T + h_{T-1}\right)}{(S_T^{-1}/2)^{\frac{n+1}{2} + h_T + h_{T-1}}} \tag{6.32}$$

Noting that $\Gamma\left(\frac{n+1}{2} + h_T + h_{T-1}\right) = \Gamma\left(\frac{n+1}{2} + h_T\right)\left[\frac{n+1}{2} + h_T\right]_{h_{T-1}}$, (6.32) can be written as:

$$\begin{aligned} &\sum_{h_{T-1}=0}^{\infty} \left(\sum_{h_T=0}^{\infty} \widetilde{a}_{T-1,h_T} \Gamma\left(\frac{n+1}{2} + h_T\right) \frac{\left[(n+1)/2 + h_T\right]_{h_{T-1}}}{[n/2]_{h_{T-1}}} \frac{\left(\frac{1}{2}\rho^2 S_T\right)^{h_{T-1}}}{(h_{T-1})!} (2S_T)^{\frac{n+1+2h_T}{2}} \right) k_{T-2}^{h_{T-1}} \\ &= \sum_{h_{T-1}=0}^{\infty} \widetilde{a}_{T-2,h_{T-1}} k_{T-2}^{h_{T-1}} \end{aligned} \tag{6.33}$$

where:

$$\widetilde{a}_{T-2,h_{T-1}} = \sum_{h_T=0}^{\infty} \widetilde{a}_{T-1,h_T}\Gamma\Big(\frac{n+1}{2}+h_T\Big)\frac{\big[(n+1)/2+h_T\big]_{h_{T-1}}}{[n/2]_{h_{T-1}}}\frac{\big(\frac{1}{2}\rho^2 S_T\big)^{h_{T-1}}}{(h_{T-1})!}(2S_T)^{\frac{n+1+2h_T}{2}}$$

Therefore, we have that the integral of (6.30) with respect to k_{T-1} gives (6.33).Therefore, collecting the terms that depend on k_{T-2} we have that:

$$\pi(k_{T-2}|k_{1:(T-3)},y_{1:T}) \propto$$
$$|k_{T-2}|^{\frac{n+1-2}{2}}\exp\Big(-\frac{1}{2}S_{T-1}^{-1}k_{T-2}\Big)\Big(\sum_{h_{T-1}=0}^{\infty}\widetilde{a}_{T-2,h_{T-1}}k_{T-2}^{h_{T-1}}\Big){}_0F_1\Big(\frac{n}{2};\frac{1}{4}\rho^2 k_{T-2}k_{T-3}\Big) \tag{6.34}$$

Equation (6.34) depends on the product of two series, which can be written as follows:

$$\Big(\sum_{h_{T-1}=0}^{\infty}\widetilde{a}_{T-2,h_{T-1}}k_{T-2}^{h_{T-1}}\Big){}_0F_1\Big(\frac{n}{2};\frac{1}{4}\rho^2 k_{T-2}k_{T-3}\Big) =$$
$$\Big(\sum_{h_{T-1}=0}^{\infty}\widetilde{a}_{T-2,h_{T-1}}k_{T-2}^{h_{T-1}}\Big)\Big(\sum_{h_{T-2}=0}^{\infty}\frac{\Big(\frac{1}{4}\rho^2 k_{T-2}k_{T-3}\Big)^{h_{T-2}}}{(h_{T-2})!}\frac{1}{[n/2]_{h_{T-2}}}\Big) =$$
$$\sum_{h=0}^{\infty}\Big(\sum_{h_{T-2}=0}^{h}\widetilde{a}_{T-2,h-h_{T-2}}\frac{1}{(h_{T-2})!}\frac{1}{[n/2]_{h_{T-2}}}\Big(\frac{1}{4}\rho^2 k_{T-3}\Big)^{h_{T-2}}\Big)k_{T-2}^{h} = \sum_{h=0}^{\infty}a_{T-2,h}k_{T-2}^{h}$$

where:

$$a_{T-2,h} = \sum_{h_{T-2}=0}^{h}\widetilde{a}_{T-2,h-h_{T-2}}\frac{1}{(h_{T-2})!}\frac{1}{[n/2]_{h_{T-2}}}\Big(\frac{1}{4}\rho^2 k_{T-3}\Big)^{h_{T-2}}$$

Therefore, we can write (6.34) as:

$$\pi(k_{T-2}|k_{1:(T-3)},y_{1:T}) \propto |k_{T-2}|^{\frac{n+1-2}{2}}\exp\Big(-\frac{1}{2}S_{T-1}^{-1}k_{T-2}\Big)\sum_{h=0}^{\infty}a_{T-2,h}k_{T-2}^{h} \tag{6.35}$$

which proves the result for $s=2$.

Because $\pi(k_{T-2}|k_{1:(T-3)},y_{1:T})$ in (6.35) and $\pi(k_{T-1}|k_{1:T-2},y_{1:T})$ in (6.30) have the same structure, and because the transition density is always the same, we can conclude the result is proven for any $s=0,\dots,T-2$. For $s=T-1$ there is no transition density from k_0 to k_1, therefore there is no need to multiply two series, so we get $a_{1,h}=\widetilde{a}_{1,h}$ and $S_2=(1+B^2e_1^2)^{-1}$. □

6.5 Proof of Proposition 3.4

Proof. To find $\pi(k_t|y_{1:T})$ we need to integrate $\pi(k_{1:T})\pi(y_{1:T}|k_{1:T})$ with respect to $k_{1:(t-1)}$ and with respect to $k_{(t+1):T}$. From the proofs of propositions 3.2 and 3.3, we have that when $2 \leq t < (T-1)$:

$$\int\int \pi(k_{1:T})\pi(y_{1:T}|k_{1:T})dk_{1:(t-1)}dk_{(t+2):T} \propto |k_t|^{\frac{n+1-2}{2}} \exp\left(-\frac{1}{2}S_{t+1}^{-1}k_t\right)\left(\sum_{h=0}^{\infty} C_{t,h}|k_t|^h\right) \times |k_{t+1}|^{\frac{n+1-2}{2}} \exp\left(-\frac{1}{2}S_{t+2}^{-1}k_{t+1}\right)\left(\sum_{h=0}^{\infty} a_{t+2,h}\frac{\Gamma\left(\frac{n+1}{2}+h\right)}{(S_{t+3}^{-1}/2)^{\frac{n+1+2h}{2}}}\right) \tag{6.36}$$

In the proof of proposition 3.3, it is shown that:

$$\sum_{h=0}^{\infty} a_{t+2,h}\frac{\Gamma\left(\frac{n+1}{2}+h\right)}{(S_{t+3}^{-1}/2)^{\frac{n+1+2h}{2}}} = \sum_{h=0}^{\infty}\widetilde{a}_{t+1,h}k_{t+1}^h$$

Therefore (6.36) can be written as:

$$\pi(k_t, k_{t+1}|y_{1:T}) \propto |k_t|^{\frac{n+1-2}{2}} \exp\left(-\frac{1}{2}S_{t+1}^{-1}k_t\right)\left(\sum_{h=0}^{\infty} C_{t,h}|k_t|^h\right) \times |k_{t+1}|^{\frac{n+1-2}{2}} \exp\left(-\frac{1}{2}S_{t+2}^{-1}k_{t+1}\right)\left(\sum_{h=0}^{\infty}\widetilde{a}_{t+1,h}k_{t+1}^h\right) \tag{6.37}$$

The product of the two series can be written as:

$$\left(\sum_{h=0}^{\infty} C_{t,h}|k_t|^h\right)\left(\sum_{h=0}^{\infty}\widetilde{a}_{t+1,h}k_{t+1}^h\right) = \sum_{h_{t+1}=0}^{\infty}\sum_{h=0}^{\infty}\sum_{h_t=0}^{h}\widetilde{C}_{t,h-h_t}\frac{1}{[n/2]_{h_t}}\left(\frac{1}{4}\rho^2\right)^{h_t}\frac{(k_{t+1})^{h_t+h_{t+1}}}{h_t!}|k_t|^h\widetilde{a}_{t+1,h_{t+1}} \tag{6.38}$$

where neither $\widetilde{a}_{t+1,h_{t+1}}$ nor $\widetilde{C}_{t,h-h_t}$ depend on k_{t+1}. Therefore, we can integrate out k_{t+1} from (6.37) using (6.38) to obtain:

$$\pi(k_t|y_{1:T}) \propto |k_t|^{\frac{n+1-2}{2}} \exp\left(-\frac{1}{2}S_{t+1}^{-1}k_t\right)\left(\sum_{h=0}^{\infty}\widetilde{D}_{t,h}|k_t|^h\right)$$

where

$$\widetilde{D}_{t,h} = \sum_{h_t=0}^{h} \sum_{h_{t+1}=0}^{\infty} \widetilde{C}_{t,h-h_t} \frac{1}{[n/2]_{h_t}} \left(\frac{1}{4}\rho^2\right)^{h_t} \frac{1}{h_t!} \frac{\Gamma\left(\frac{n+1}{2} + h_t + h_{t+1}\right)}{(S_{t+2}^{-1}/2)^{\frac{n+1}{2}+h_t+h_{t+1}}} \widetilde{a}_{t+1,h_{t+1}}$$

as we wanted to prove.

In the case $t = T-1$, expression (6.37) becomes:

$$\pi(k_{T-1}, k_T | y_{1:T}) \propto |k_{T-1}|^{\frac{n+1-2}{2}} \exp\left(-\frac{1}{2}S_T^{-1}k_{T-1}\right)\left(\sum_{h=0}^{\infty} C_{T-1,h}|k_{T-1}|^h\right)|k_T|^{\frac{n+1-2}{2}} \exp\left(-\frac{1}{2}S_{T+1}^{-1}k_T\right) \tag{6.39}$$

Thus, in this case we only have one series, not the product of two. Integrating with respect to k_T we get:

$$\pi(k_{T-1}|y_{1:T}) \propto |k_{T-1}|^{\frac{n+1-2}{2}} \exp\left(-\frac{1}{2}S_T^{-1}k_{T-1}\right) \sum_{h=0}^{\infty} \widetilde{D}_{T-1,h}|k_{T-1}|^h$$

with

$$\widetilde{D}_{T-1,h} = \sum_{h_{T-1}=0}^{h} \widetilde{C}_{T-1,h-h_{T-1}} \frac{1}{[n/2]_{h_{T-1}}} \left(\frac{1}{4}\rho^2\right)^{h_{T-1}} \frac{1}{(h_{T-1})!} \frac{\Gamma\left(\frac{n+1}{2} + h_{T-1}\right)}{(S_{T+1}^{-1}/2)^{\frac{n+1}{2}+h_{T-1}}}$$

as we wanted to prove.

□

6.6 Proof of Local Scale Model Likelihood

To facilitate the reading we do not explicitly write x_t as a conditioning argument. Given that we have a gamma distribution for the initial condition (2.2) and a Gaussian error term, we have that the joint density (y_1, h_1, ν_1) is :

$$\pi(y_1, h_1, \nu_1) = \frac{1}{\sqrt{2\pi}}(h_1)^{\frac{1}{2}} \exp\left(-\frac{1}{2}(y_1 - x_1\beta)^2 h_1\right) f(h_1|S_1) \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \nu_1^{\alpha_1-1}(1-\nu_1)^{\alpha_2-1}$$

where $f(h_1|S_1)$ is the density of the initial condition given as:

$$f(h_1|S_1) = h_1^{\frac{\nu}{2}-1} \exp\left(-\frac{h_1}{2S_1}\right) \frac{1}{\Gamma(\nu/2)(2S_1)^{\frac{\nu}{2}}} \tag{6.40}$$

The volatility process is represented by a non stationary process as in (2.1). We make a change of variables from (y_1, h_1, ν_1) to (y_1, Z, h_2) where $Z = h_1 - \lambda h_2$, and $v_1 = \frac{h_2\lambda}{h_1}$. The

Jacobian of this transformation is $\lambda/(Z+\lambda h_2)$. Therefore $\pi(y_1, Z, h_2)$ can be written as:

$$\pi(y_1,Z,h_2) = \frac{1}{\sqrt{2\pi}}(Z+\lambda h_2)^{\frac{1}{2}}\exp\left(-\frac{1}{2}(y_1-x_1\beta)^2(Z+\lambda h_2)\right)(Z+\lambda h_2)^{\frac{\nu}{2}-1}\exp\left(-\frac{1}{2S_1}(Z+\lambda h_2)\right)$$
$$\times\left(\frac{(Z+\lambda h_2)}{\lambda}\right)^{-1}\frac{1}{\Gamma(\nu/2)(2S_1)^{\frac{\nu}{2}}}\frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\left(\frac{h_2\lambda}{Z+\lambda h_2}\right)^{\alpha_1-1}\left(\frac{Z}{Z+\lambda h_2}\right)^{\alpha_2-1}$$

which simplifies to:

$$\pi(y_1,Z,h_2) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left((y_1-x_1\beta)^2+\frac{1}{S_1}\right)(Z+\lambda h_2)\right)\frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\frac{(Z+\lambda h_2)^{\frac{\nu}{2}+\frac{1}{2}-(\alpha_1+\alpha_2)}}{\Gamma(\nu/2)(2S_1)^{\frac{\nu}{2}}}\lambda_1^{\alpha}Z^{\alpha_2-1}h_2^{\alpha_1-1}$$

Note that for mathematical convenience, α_1 is restricted as $\alpha_1 = \frac{\nu}{2}$ and $\alpha_2 = \frac{1}{2}$. Therefore, $\frac{\nu}{2}+\frac{1}{2}-(\alpha_1+\alpha_2)=0$, and $\pi(Z|y_1,h_2)$ is a gamma distribution. Using the properties of the gamma distribution, we can integrate over the state variable Z:

$$\pi(y_1,h_2) = \int \pi(y_1,Z,h_2)dZ$$
$$= \frac{\Gamma(\alpha_2)}{\sqrt{2\pi}}\frac{\exp\left(-\frac{\lambda h_2}{2}\left((y_1-x_1\beta)^2+\frac{1}{S_1}\right)\right)}{\left(2\left((y_1-x_1\beta)^2+\frac{1}{S_1}\right)^{-1}\right)^{-\alpha_2}}\frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\frac{\lambda_1^{\alpha}h_2^{\nu/2-1}}{\Gamma(\nu/2)(2S_1)^{\frac{\nu}{2}}} \tag{6.41}$$

From equation (6.41) we can see that $h_2|y_1$ is a gamma distribution with parameters $\left(\frac{\nu}{2}, 2S_2\right)$, where $S_2 = \left((y_1-x_1\beta)^2+\frac{1}{S_1}\right)^{-1}\frac{1}{\lambda}$. Let $f(h_2|S_2)$ be defined as in (6.40), that is, the density of a gamma distribution:

$$f(h_2|S_2) = h_2^{\frac{\nu}{2}-1}\exp\left(-\frac{h_2}{2S_2}\right)\frac{1}{\Gamma(\nu/2)(2S_2)^{\frac{\nu}{2}}}. \tag{6.42}$$

Then equation (6.41) can be written as follows:

$$\pi(y_1,h_2) = \frac{\Gamma(\alpha_2)}{\Gamma(\alpha_2)}\frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)}\frac{\lambda^{\alpha_1}}{\Gamma(\nu/2)(2S_1)^{\frac{\nu}{2}}}\frac{1}{\sqrt{2\pi}}\left(2\left((y_1-x_1\beta)^2+\frac{1}{S_1}\right)^{-1}\right)^{\alpha_2}f(h_2|S_2)\Gamma(\nu/2)(2S_2)^{\frac{\nu}{2}}$$

Therefore, $h_2|y_1$ is a gamma distribution, such that $\pi(h_2|y_1) = f(h_2|S_2)$, which is defined in (6.42).

From these derivations we can get the likelihood as follows. First, for $t=1$, we have that

$$\pi(y_1|h_1) = \frac{1}{(\sqrt{2\pi})}h_1^{\frac{1}{2}}\exp\left(-\frac{1}{2}h_1(y_1-x_1\beta)^2\right)$$

and the initial condition for h_1 is a gamma distribution given in (6.40). Therefore, $\pi(y_1)$ is a student-t and we have:

$$\pi(y_1) = \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)}\lambda^{\alpha_1}\left(\frac{S_2}{S_1}\right)^{\alpha_1}\frac{1}{\sqrt{2\pi}}\left(2\left((y_1-x_1\beta)^2+\frac{1}{S_1}\right)^{-1}\right)^{\alpha_2} \tag{6.43}$$

For $t = 2$, $\pi(y_2|h_2)$ is also a normal. Thus the conditional distribution for the second observation given h_2 is as follows:

$$\pi(y_2|h_2) = \frac{1}{(\sqrt{2\pi})}h_2^{\frac{1}{2}}\exp\left(-\frac{1}{2}h_2(y_2-x_2\beta)^2\right)$$

and $\pi(h_2|y_1)$ is the gamma distribution defined in (6.42). Therefore, we have the same structure as in $t = 1$, and using the properties of the gamma distribution , we get that the likelihood $\pi(y_2|y_1)$ is a student-t as follows:

$$\pi(y_2|y_1) = \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)}\lambda^{\alpha_1}\left(\frac{S_3}{S_2}\right)^{\alpha_1}\frac{1}{\sqrt{2\pi}}\left(2\left((y_2-x_2\beta)^2+\frac{1}{S_2}\right)^{-1}\right)^{\alpha_2} \tag{6.44}$$

where $S_3 = \left((y_2-x_2\beta)^2+\frac{1}{S_2}\right)^{-1}\frac{1}{\lambda}$.

Because the kernels are the same for $t = 1$ and for $t = 2$, then we have proved it for every t.