Bayesian Analysis (2019) **14**, Number 1, pp. 29–52

Filtering and Estimation for a Class of Stochastic Volatility Models with Intractable Likelihoods

Emilian R. Vankov[*§], Michele Guindani[†], and Katherine B. Ensor[‡§]

Abstract. We introduce a new approach to latent state filtering and parameter estimation for a class of stochastic volatility models (SVMs) for which the likelihood function is unknown. The α-stable stochastic volatility model provides a flexible framework for capturing asymmetry and heavy tails, which is useful when modeling financial returns. However, the α-stable distribution lacks a closed form for the probability density function, which prevents the direct application of standard Bayesian filtering and estimation techniques such as sequential Monte Carlo and Markov chain Monte Carlo. To obtain filtered volatility estimates, we develop a novel approximate Bayesian computation (ABC) based auxiliary particle filter, which provides improved performance through better proposal distributions. Further, we propose a new particle based MCMC (PMCMC) method for joint estimation of the parameters and latent volatility states. With respect to other extensions of PMCMC, we introduce an efficient single filter particle Metropolis-within-Gibbs algorithm which can be applied for obtaining inference on the parameters of an asymmetric α-stable stochastic volatility model. We show the increased efficiency in the estimation process through a simulation study. Finally, we highlight the necessity for modeling asymmetric α-stable SVMs through an application to propane weekly spot prices.

Keywords: particle Markov chain Monte Carlo, auxiliary particle filter, approximate Bayesian computation, stable distribution.

1 Introduction

Assessing the unobserved variability of financial returns is crucial to regulators and policy makers. Underestimating such variability can have detrimental effects to the individual investor as well as the stability of the entire global economy. In an effort to study the variability of returns, researchers have proposed various measures such as historical volatility, conditional heteroscedastic volatility models (Engle, 1982; Bollerslev, 1986; Nelson, 1991) and stochastic volatility models (SVM) (Taylor, 1994; Hull and

[*]Baker Institute for Public Policy at Rice University, P.O. Box 1892 Houston, Texas 77251-1892, emilian.r.vankov@rice.edu

[†]Department of Statistics, 2241 Bren Hall, University of California, Irvine, CA 92697-1250, michele.guindani@uci.edu

[‡]Department of Statistics, Rice University, 2053 Duncan Hall Rice University Houston, TX 77251-1892, ensor@rice.edu

[§]The authors would like to thank the Rice Center for Computational Finance and Economic Systems and the Baker Institute for Public Policy, Rice University, Houston, TX.

 https://doi.org/10.1214/18-BA1099

White, 1987). In these models, a Gaussian distribution is often assumed, (Kim et al., 1998; Kastner and Frühwirth-Schnatter, 2014).

While assuming Gaussian returns allows simpler and estimable models, this assumption excludes important features of financial returns, such as heavy tails and skewness. A natural framework for modeling such behavior has been provided by the α-stable distribution (Mandelbrot, 1963). In addition to scale and location parameters, this distribution has two additional parameters representing the heavy-tailedness (stability) and skewness of the distribution (Nolan, 1997). The Gaussian distribution can be obtained as a special case to the α-stable distribution, so that modeling returns with an α-stable distribution provides a flexible framework to estimate the volatility. A key challenge to models based on the α-stable distribution is the absence of closed form expression for the probability density function (p.d.f.). Standard estimation techniques, such as maximum likelihood and Markov chain Monte Carlo (MCMC) are unsuitable. In the last decade, advancements in approximate Bayesian computation (ABC), have aided the applications of α-stable distributions.

In this paper we develop an ABC based auxiliary particle filter (APF-ABC) for efficient latent state estimation in state-space models. In addition, we develop a single filter particle Metropolis-within-Gibbs (SF-PMwG) algorithm to efficiently estimate the volatility and parameters of the asymmetric α-stable stochastic volatility model. To avoid confusion, we note that we refer to asymmetry in the unconditional distribution of the returns rather than the correlation between volatility and lagged returns, also known as 'leverage' effect. Current methods for ABC filtering use proposal densities which do not consider available information from the data. While simple and easy to implement, such a strategy can lead to inefficiencies and poor estimates in the presence of heavy tails in the state-space model. Building upon ideas from Pitt and Shephard (1999) and Carpenter et al. (1999), we introduce a modified proposal density for the ABC filter, which uses the data to pre-weight the samples, also termed particles, used for approximating the latent states. We show through simulation studies that APF-ABC improves on standard ABC based filters.

The ABC based auxiliary particle filter only provides estimates of the latent volatility, conditional on the data and the parameters. The parameters are typically unknown and need to be estimated. Particle MCMC (PMCMC) algorithms (Andrieu et al., 2010), use a particle filter to build a proposal within standard MCMC techniques. These type of algorithms can be divided into particle marginal Metropolis-Hastings (PMMH) and particle Gibbs sampler (PG). Mendes et al. (2015) have recently proposed a sampler that combines PMMH and PG. As indicated by the authors there is an increased computational burden associated with running two filtering algorithms at each step of the sampler. In an effort to produce an efficient sampler, we develop a single filter particle Metropolis-within-Gibbs. Through simulations, we illustrate the significant increase in computational efficiency of our algorithm as compared to standard ABC based particle MCMC discussed in Jasra et al. (2013) and the ABC counterpart of Mendes et al. (2015). We show how our work can provide inference on a dataset involving weekly propane spot prices from Mount Belvieu, Texas. Our estimation confirms empirical findings that propane returns are heavy-tailed and skewed.

Our paper is organized as follows: in Section 2 we provide an overview of most used filtering and estimation procedures for the stochastic volatility model under different distributional assumptions. In Section 3 we introduce the ABC based auxiliary particle filter for filtering in state-space models and study its properties via a simulation study. In Section 4 we introduce the single filter particle Metropolis-within-Gibbs algorithm. We present the results obtained by applying the proposed algorithms to simulated data from the α-stable SVM along with comparison to other methods in Section 5. In Section 6 we study the volatility of weekly propane spot prices, and consider the economic implications. Section 7 concludes the paper.

2 Background

2.1 Stochastic Volatility Models

The simple stochastic volatility model for a time series of length T was originally introduced in Taylor (1994) and can be represented by the following set of equations:

$$\begin{aligned} y_t &= exp\,(x_t/2)\,v_t && 1 \le t \le T, \\ x_t &= \mu + \phi(x_{t-1} - \mu) + \sigma w_t && 1 \le t \le T, \end{aligned} \tag{1}$$

where $\{Y_t\}_{t\geq 1},\ Y_t \in \mathcal{Y} \subseteq \mathbb{R}^{\mathbb{N}}$ indicates the returns process with p.d.f. satisfying $Y_t|X_t = x_t \sim g(\cdot|x_t, \theta)$; and $\{X_t\}_{t\geq 1},\ X_t \in \mathcal{X} \subseteq \mathbb{R}^{\mathbb{N}}$ indicates the latent log-volatility process with p.d.f. $X_t|X_{t-1} = x_{t-1} \sim f(\cdot|x_{t-1}, \theta)$. The initial distribution of the latent process, $\eta(x_0|\theta)$, is $X_0 \sim N(\mu, \sigma^2/(1-\phi^2))$. The parameter vector is denoted as $\theta = \{\mu, \phi, \sigma, \lambda\} \subseteq \mathbb{R} \times (-1, 1) \times (0, \infty) \times \mathbb{R}$, where λ is a vector including any additional parameters associated with the distribution of the latent log-volatility and returns processes.

If we assume that θ is known, the volatility estimation is based on the filtering distribution

$$\begin{aligned} p(x_{0:T}|y_{1:T}, \theta) &= \frac{p(x_{0:T}, y_{1:T}|\theta)}{p(y_{1:T}|\theta)} \\ &\propto \eta(x_0|\theta) \prod_{t=1}^{T} g(y_t|x_t, \theta) f(x_t|x_{t-1}, \theta). \end{aligned} \tag{2}$$

In what follows we may equivalently set $p_\theta(\cdot) = p(\cdot|\theta)$ for any p.d.f. p, for notational simplicity when there is no possibility of confusion. Due to the inability to express the filtering distribution in closed form, the estimation of the unobserved volatility is typically based on numerical methods known as sequential Monte Carlo (SMC) or particle filters. Applications of those methods to SVMs are given in Doucet (2001), Pitt and Shephard (1999), and Creal (2012).

In most applications, however, θ is unknown; hence interest lies in joint estimation of the parameters as well as the filtered states given by $p(\theta, x_{0:T}|y_{1:T})$. Regardless of the particular choice for f and g, closed forms for the posterior densities associated with the SVM given in (1) are not available.

Different specifications of the distributional form for the noise terms v_t and w_t have appeared in the literature, the most common assuming that they are uncorrelated and follow a standard normal distribution (Shephard, 2005). In such cases, Bayesian estimation of the parameters and the filtering of states can rely on standard MCMC methods (Jacquier et al., 1994; Kim et al., 1998). Kastner and Frühwirth-Schnatter (2014) studied the efficiency of the MCMC estimates for the different parametrizations of the Gaussian SVM and introduced an interweaving strategy to improve efficiency of parameter estimation. However, to capture the heavy-tails empirically observed in financial returns Harvey et al. (1994) and Chib et al. (2002) have also suggested t-distributed error term v_t for modeling returns. Other attempts to capture the heavy-tails have been made via finite (Jacquier et al., 2004; Abanto-Valle et al., 2010) and infinite (Jensen and Maheu, 2010) mixture of distributions. Alternatively, one can use the α-stable distribution (Lombardi and Calzolari, 2009). Here we denote the α-stable distribution as $SD(\alpha, \beta, \zeta, \nu)$. The parameter vector $(\alpha, \beta, \zeta, \nu) \in (0, 2] \times [-1, 1] \times (-\infty, \infty) \times [0, \infty)$ gives a measure of heavy-tailedness, skewness, location and scale, respectively. When $\alpha < 2$, the distribution is heavy-tailed; negative values of the parameter β correspond to negatively skewed returns.

Due to the inability to express the p.d.f. in closed form, traditional estimation techniques, such as MCMC, are unavailable. Earlier estimation procedures for the parameters of the α-stable distribution can be attributed to Press (1972), McCulloch (1986), Buckle (1995), and more recently Lombardi (2007). The development of ABC in the last decade, which requires only that one is able to simulate auxiliary data from the model, has aided inference in α-stable models (Peters et al., 2012). This is possible due to the likelihood-free data generating algorithm of α-stable random samples of Chambers et al. (1976). For details see Appendix A of the Supplementary Material (Vankov et al., 2019).

2.2 Approximate Bayesian Computation and SMC

A standard SMC algorithm provides approximate samples from $p_\theta(x_{0:T}|y_{1:T})$ by sampling a vector of particles $\{X_t^i, i = 1, \ldots, N\}$ from an importance distribution $q(\cdot)$. To account for the difference between the target and the importance distributions, one typically needs to weight the samples sequentially through time.

In contrast, ABC based sequential Monte Carlo (Peters et al., 2012) targets an extended distribution given by $p_\theta^\epsilon(x_{0:T}, u_{1:T}|y_{1:T})$, where $u_{1:T}$ is a sequence of auxiliary data simulated from the model $p_\theta(y_{1:T}|x_{0:T})$. More specifically, the target is

$$p_\theta^\epsilon(x_{0:T}, u_{1:T}|y_{1:T}) = \frac{\prod_{t=1}^T K_\epsilon(y_t|u_t) g_\theta(u_t|x_t) f_\theta(x_t|x_{t-1})}{\int \prod_{t=1}^T K_\epsilon(y_t|u_t) g_\theta(u_t|x_t) f_\theta(x_t|x_{t-1}) dx_{1:t}}, \tag{3}$$

where $K_\epsilon(y) = K(y/\epsilon)/\epsilon$ is a general kernel, for a given bandwidth $\epsilon > 0$ with $K(\cdot)$ a density satisfying $\int K(y)dy = 1$. The use of an extended target circumvents the evaluation of the likelihood in the sampling algorithm. Furthermore, if the proposal $q(\cdot)$ is selected so that $X_t \sim f_\theta(\cdot|x_{t-1})$, and $U_t \sim g_\theta(\cdot|x_t)$, then the weights used for propagating the particles through time are updated simply as

$$w_t^i = K_\epsilon(y_t|u_t^i)w_{t-1}^i \ , \text{ for } \ i = 1, \dots, N.$$

This is desirable in models such as those based on the α-stable distribution, where the likelihood is computationally expensive to calculate, while simulating data can be achieved at a low computational cost.

There has been considerable literature expanding on the SMC-ABC idea. For example, Jasra et al. (2012) propose to use a uniform kernel in (3). A drawback, however, is that the uniform kernel produces binary weights, either 0 or 1. Hence, it is possible that at some point all weights will become zero. In order to prevent the algorithm from collapsing in such a manner, when using a uniform kernel, Jasra et al. (2013) propose the alive particle filter, which ensures that a pre-specified number of particles survives all steps of the algorithm. In particular, this filter resamples particles at each step, until there are exactly $N - 1$ of them with non-zero weights. This scheme comes at a computational cost, however, as the amount of time each step takes to ensure that the above condition holds is a random variable. For a review and further examination of SMC-ABC we refer the reader to Jasra (2015).

To estimate the latent states and parameters of a symmetric stochastic volatility model for stock returns, Jasra et al. (2013) propose to combine particle marginal Metropolis-Hastings with ABC based sequential Monte Carlo (PMMH-ABC). However, the parameter characterizing the stability of the distribution is set to a fixed value rather than being part of the estimation. Another application of the symmetric SVM can be found in Barthelmé and Chopin (2014), who extend the expectation propagation algorithm of Minka (2001) to ABC framework (EP-ABC). A simulation example is provided to compare their method to PMMH-ABC. The authors conclude that EP-ABC provides improvements in computational time by sacrificing some accuracy. The theoretical properties related to the loss in accuracy remain to be studied. All of these methods, however, do not account for asymmetry in the financial returns.

In the following sections we introduce a single filter particle Metropolis-within-Gibbs algorithm, combined with an ABC based auxiliary particle filter and show improved efficiency for simultaneously estimating all parameters of the asymmetric, heavy-tailed stochastic volatility model.

3 Auxiliary Particle Filter for ABC

All the SMC-ABC algorithms discussed in the previous section have relied on sampling $X_t \sim f_\theta(\cdot|x_{t-1})$, the state transition density, without accounting for the most current data point. While this is a convenient choice, in certain situations, especially when the data is very informative, such proposals can degenerate the performance of the algorithm. Being able to account for the data in the proposal distribution of the particles can improve the performance of the algorithm as the state space is explored more efficiently. Based on those considerations, we develop an ABC based auxiliary particle filter, which we present in the following subsection.

3.1 Algorithm Description

In this section we introduce an ABC based auxiliary particle filter, which is inspired by ideas in Pitt and Shephard (1999) and Carpenter et al. (1999), and then examine its application to α-stable stochastic volatility models. For propagating the particles $\{X_t^i,\ i = 1, \ldots, N\}$ through time in ABC based sequential Monte Carlo, one requires the evaluation of weights

$$\begin{aligned} w_t^i &\propto \frac{p_\theta^\epsilon(x_{0:t}^i, u_{1:t}^i | y_{1:t})}{q_\theta(x_{0:t}^i, u_{1:t}^i | y_{1:t})} \\ &\propto \frac{K_\epsilon(y_t|u_t^i) g_\theta(u_t^i|x_t^i) f_\theta(x_t^i|x_{t-1}^i)}{q_\theta(x_t^i, u_t^i | x_{0:t-1}^i, u_{1:t-1}^i, y_{1:t})} \frac{p_\theta^\epsilon(x_{0:t-1}^i, u_{1:t-1}^i | y_{1:t-1})}{q_\theta(x_{0:t-1}^i, u_{1:t-1}^i | y_{1:t})}. \end{aligned} \tag{4}$$

The optimal importance density, in the sense of minimizing the variance of the weights, $p_\theta^\epsilon(x_{0:t-1}, u_{1:t-1}|y_{1:t})$ is not available for the α-stable stochastic volatility model. However, we would like to choose $q_\theta(x_{0:t-1}, u_{1:t-1}|y_{1:t})$ as close as possible to the optimal density without severely increasing the computational complexity of the particle filter. Consider that we can write

$$\begin{aligned} & p_\theta^\epsilon(x_{0:t-1}, u_{1:t-1}|y_{1:t}) \\ & \propto p_\theta^\epsilon(x_{0:t-1}, u_{1:t-1}|y_{1:t-1}) \int\int K_\epsilon(y_t|u_t) g_\theta(u_t|x_t) f_\theta(x_t|x_{t-1}) dx_t du_t \\ & \approx \sum_{i=1}^N \hat{w}_{t-1}^i \delta_{X_{0:t-1}^i}, \end{aligned} \tag{5}$$

where

$$\hat{w}_{t-1}^i = \left(w_{t-1}^i \int\int K_\epsilon(y_t|u_t^i) g_\theta(u_t^i|x_t^i) f_\theta(x_t^i|x_{t-1}^i) dx_t^i du_t^i \right). \tag{6}$$

In (5) we have made use of the SMC approximation

$$\hat{p}_\theta^\epsilon(x_{0:t-1}, u_{1:t-1}|y_{1:t-1}) = \sum_{i=1}^N w_{t-1}^i \delta_{X_{0:t-1}^i}.$$

Therefore, we can choose the importance density

$$q_\theta(x_{0:t}, u_{1:t}|y_{1:t}) = q_\theta(x_t, u_t|x_{0:t-1}, u_{1:t-1}, y_{1:t}) \sum_{i=1}^N \hat{w}_{t-1}^i \delta_{X_{0:t-1}^i}$$

and substitute in (4) to obtain

$$w_t^i \propto \frac{w_{t-1}^i}{\hat{w}_{t-1}^i} \frac{K_\epsilon(y_t|u_t^i) g_\theta(u_t^i|x_t^i) f_\theta(x_t^i|x_{t-1}^i)}{q_\theta(x_t^i, u_t^i|x_{0:t-1}^i, u_{1:t-1}^i, y_{1:t})}.$$

As the integral in (6) does not typically admit a closed form, we must either numerically evaluate it or resort to alternative approximation methods. In the context of standard SMC, Johansen and Doucet (2008) suggest that having an importance function with heavier tails than the target keeps the variance of the estimates bounded. In concurrence with this guideline, in our work we chose to approximate the integral with the density of a t-distribution, but arguably other approximations can be employed.

Let $h : \mathbb{R} \to [0, \infty)$ be a density approximating the integral in (6). Further, let $\xi : \mathbb{R} \to \mathbb{R}$, typically chosen to be the mean or mode of the state. In our work we let ξ be the identity function. Finally, let $A_t \sim \mathcal{R}(\cdot|W_{t-1})$ be a random variable representing the index used for resampling the particles at each time step. For a review of commonly used resampling methods we refer to Douc and Cappé (2005). Here we use multinomial resampling.

We present our ABC based auxiliary particle filter (APF-ABC) in Algorithm 1. We should note that the algorithm can be applied to a wide range of state-space models, for which the p.d.f. of the data is not available in closed form, but simulation from the model is possible.

Algorithm 1 ABC based auxiliary particle filter (APF-ABC)

Initialization: $t = 0$

Sample $x_0^{(1)}, \ldots, x_0^{(N)}$ from $\eta_\theta(\cdot)$

set $w_0^i = \frac{1}{N}$ for $i = 1, \ldots, N$

Update: $t \geq 1$

for $t = 1, \ldots, T$ **do**

for $i = 1, \ldots, N$ **do**

Set $\hat{w}_{t-1}^i \propto w_{t-1}^i h_\theta(y_t|\xi(x_{t-1}^i))$

Sample $A_{t-1}^i \sim \mathcal{R}(\cdot|\hat{W}_{t-1})$

Sample $(X_t^i, U_t^i) \sim q_\theta(\cdot|x_{0:t-1}^{a_{1:t-1}^i}, u_{1:t-1}^{a_{1:t-1}^i}, y_{1:t})$

Update the weights via:

$$w_t^i \propto \frac{w_{t-1}^i}{\hat{w}_{t-1}^i} \frac{K_\epsilon(y_t|u_t^{a_{1:t-1}^i}) g_\theta(u_t^i|x_t^{a_{1:t-1}^i}) f_\theta(x_t^i|x_{t-1}^{a_{1:t-1}^i})}{q_\theta(x_t^i, u_t^i|x_{t-1}^{a_{1:t-1}^i}, u_{t-1}^{a_{1:t-1}^i}, y_{1:t})} \tag{7}$$

Normalize the weights $w_t^i = w_t^i / \sum_{j=1}^N w_t^j$

end for

end for

If X_t is sampled from the transition density of the state, $f_\theta(\cdot|x_{t-1})$ and U_t from the observation density $g_\theta(\cdot|x_t)$, the weights in (7) become

$$w_t^i = \frac{w_{t-1}^i}{\hat{w}_{t-1}^i} K_\epsilon(y_t|u_t^i). \tag{8}$$

The typical choice of a uniform kernel for $K_\epsilon(y_t|u_t)$ can cause the integral in (6) to become zero and force the weights in (7) to be undefined. Therefore, we select a Gaussian kernel.

Algorithm 1 provides us with an approximation to the target ABC based density; $\hat{p}_\theta^\epsilon(x_{0:T}, u_{1:T}|y_{1:T}) = \sum_{i=1}^N w_T^i \delta_{X_{0:T}^i}$. To obtain the filtered volatility state associated with this approximation, Kitagawa (1996) has proposed to use ancestral sampling. This technique consists of drawing an index from the distribution $\mathcal{R}(\cdot|W_T)$ and tracing back the ancestor particles corresponding to that index. If we let I denote the random variable representing the sampled index, then we can mathematically describe this procedure by choosing a particle at time T, and defining the path of such a particle as $\{B_{t-1}^i = A_{t-1}^{B_t^i} : t = T, \ \ldots \ , 0\}$. Tracing the particles in such a manner provides us with $\{X_T^{B_T^i}, X_{T-1}^{B_{T-1}^i}, \ldots, X_0^{B_0^i}\}$ as described by Andrieu et al. (2010). Finally, the marginal likelihood can be approximated via:

$$\hat{p}_\theta^\epsilon(y_{1:T}) = \frac{1}{N^T} \prod_{t=1}^{T} \sum_{i=1}^{N} w_t^i. \tag{9}$$

The ABC based auxiliary particle filter allows for the best parent trajectories to be selected while taking the current data point y_t into consideration by pre-weighting the particles sampled at each iteration.

3.2 Simulation Study

We begin this section by first exploring the performance of APF-ABC on a dynamic Gaussian linear model (DGLM). For such a model, the exact filtering densities can be obtained via the Kalman filter. In particular, the DGLM is given by:

$$\begin{aligned} y_t &= x_t + \sigma_y v_t \qquad && 1 \le t \le T, \\ x_t &= \mu + \phi x_{t-1} + \sigma_x w_t \qquad && 1 \le t \le T. \end{aligned} \tag{10}$$

The initial distribution of the latent process, $\eta(x_0|\theta)$, is $X_0 \sim N(\mu/(1-\phi), \sigma^2/(1-\phi^2))$ where the parameter vector is denoted as $\theta = \{\mu, \phi, \sigma_y, \sigma_x\} \subseteq \mathbb{R} \times (-1, 1) \times (0, \infty) \times (0, \infty)$. The noise terms v_t and w_t are assumed to be uncorrelated and follow a standard normal distribution.

We study the performance of APF-ABC, the Kalman filter, as well as an implementation of SMC-ABC as in Jasra et al. (2012), based on root-mean-squared error (RMSE), for three different data sets with different signal-to-noise ratios (Tables 1–3 in Appendix C of the Supplementary Material (Vankov et al., 2019)). We find that while the Kalman filter has slightly lower RMSE, both APF-ABC and SMC-ABC can provide low errors even with as few as 1000 particles. We notice that as ϵ_g, the bandwidth of the Gaussian kernel used in APF-ABC, and ϵ_u, the bandwidth of the uniform kernel used in SMC-ABC, increase, the performance of the filters deteriorates. Similarly, the RMSE increases as the length of the time series data increases. We note that even when the data is generated from DGLM, APF-ABC can provide lower RMSE as compared to SMC-ABC. Increasing the signal-to-noise ratio, results in higher RMSE for all filters. The accuracy of the ABC based filters is comparable to the RMSE of the Kalman filter, despite being slightly lower.

We note here that the results presented are the estimates of the filtering distribution of the state. The Kalman filter can also provide the exact smoothing density of the unobserved state. Moreover, recently, Martin et al. (2014) considered an ABC extension of the forward only smoothing algorithm of Del Moral et al. (2010), which is an extension to the forward-filtering backward smoothing (FFBS) method by Godsill et al. (2004). In particular, to avoid a backward pass, an auxiliary function is introduced, which is updated at each step of the filter. The weights, W_t^i, from (7) and the corresponding particles X_t^i, for $i = 1, \dots, N$, computed at each step of Algorithm 1 can be used to update the auxiliary function, which in turn provides the smoothed estimates. The theoretical and empirical results presented in Martin et al. (2014) indicate that ABC based approximations to the smoothing density are feasible in the context of state space models.

In order to study the performance of our proposed APF-ABC filter on data characterized by heavy tails, we apply our proposed algorithm to a simulated data from the stochastic volatility model (1). We assume that $w_t \sim N(0,1)$ and $v_t \sim SD(\alpha, \beta, 0, 1)$. In our first simulation setting, we set the parameters of the model to the following values: $\alpha = 1.75,\ \beta = 0.1,\ \mu = -0.2,\ \phi = 0.95,\ \sigma = 0.2$. We simulate 350 observation (y_t) and state (x_t) data points. For model fitting, we implement APF-ABC with $K_\epsilon(y_t|u_t)$ to be Gaussian with bandwidth $\epsilon_g = 0.5$. The integral in (6) is approximated with the p.d.f. of a t-distribution. We choose a t-distribution with 2 degrees of freedom to allow for heavier tails in the proposal distribution. We tested allowing for more degrees of freedom, but did not find any significant impact on the performance of our algorithm. Figure 1 represents a plot of the true state values along with the mean estimates of the latent unobserved log-volatility given by $x_t,\ t = 1, \dots, 350$ from the APF-ABC filter with $N = 5 \times 10^3$ particles (left) as well as the simulated returns (right). APF-ABC captures the general trend very well and provides estimates of the log-volatility that are close to the true values.

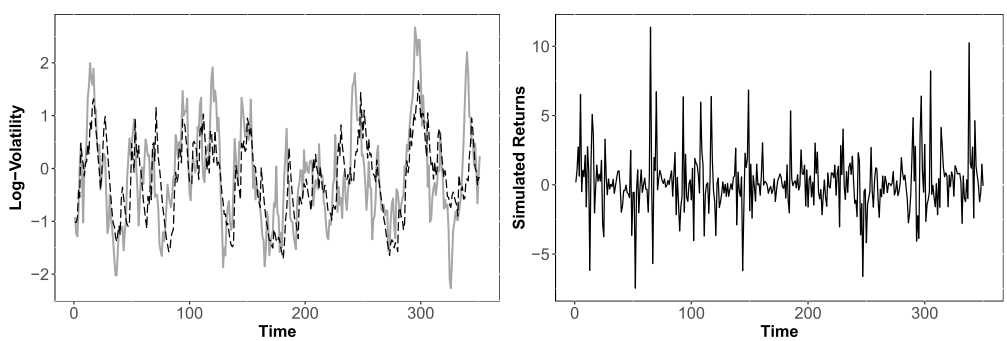

Figure 1: Simulation Study in Section 3.2. Left panel: APF-ABC mean volatility estimate (dashed line) and the true state volatility (solid line). The filtered values presented are averaged over 100 simulations of the APF-ABC. Right panel: simulated returns. The data is simulated from the model given by (1) with $w_t \sim N(0,1),\ \ v_t \sim SD(\alpha, \beta, 0, 1)$ where $\alpha = 1.75,\ \beta = 0.1,\ \mu = -0.2,\ \phi = 0.95,\ \sigma = 0.2$.

To study the differences between APF-ABC and SMC-ABC, we report in Figure 2 box plots of the RMSE for the two ABC based particle filters. The SMC-ABC and the APF-ABC were run with the same number of particles. Such a choice does not produce significant computational differences. The value of $\epsilon_u = 1.5$ is chosen as the minimum possible epsilon for which SMC-ABC does not completely collapse, which happens when all weights become zero. The advantage of our proposed filter in fitting an α-stable stochastic volatility model is highlighted by the fact that the maximum error of APF-ABC is less than the minimum possible error for SMC-ABC (Figure 2).

We also investigated the performance of the two algorithms for alternative scenarios, characterized by different parameter values, sample sizes, number of particles N, and ϵ_u and ϵ_g values (see Tables 4 and 5 in Appendix C of the Supplementary Material (Vankov et al., 2019)). All simulations confirm the conclusions obtained by comparing the RMSE between the two ABC based filters. Further, we notice that increasing the number of particles improves the accuracy of the filters. Smaller values of the stability parameter, α, which allow for heavier tails, and larger sample sizes result in higher RMSE for both filters. Finally, a larger value of ϵ_u and ϵ_g decreases the accuracy of the filters, consistently with results in the literature (Jasra, 2015). We note that direct comparison of the two filters based on the magnitude of each bandwidth ϵ_g and ϵ_u is not possible due to their different interpretation within the Gaussian and uniform kernels.

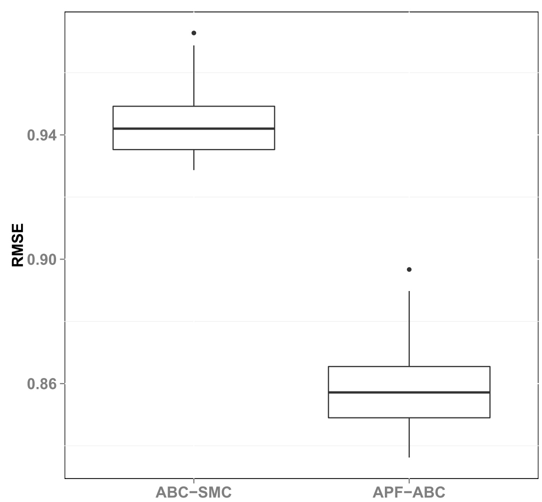

Figure 2: Simulation Study in Section 3.2. Box plots of the root-mean-square error for APF-ABC and ABC-SMC. The box plots were obtained from 100 runs of each filter.

So far we have considered only the problem of obtaining estimates of the latent volatility states assuming that all parameters are known. While APF-ABC proves to be a powerful tool for obtaining inference on the filtered states, the assumption of known parameters is typically not true in practice. As a matter of fact, the main interest of the α-stable SVM lies in estimating the *joint* distribution $p(\theta, x_{0:T}|y_{1:T})$. In the remainder of this paper we develop a novel algorithm for filtering and estimation of the asymmetric α-stable stochastic volatility model, which makes use of the APF-ABC algorithm developed in this section.

4 Single Filter Particle Metropolis-within-Gibbs

In this section we introduce a single filter particle Metropolis-within-Gibbs sampler, which we then use to estimate the asymmetric, heavy-tailed stochastic volatility model (1) with $w_t \sim N(0,1)$ and $v_t \sim SD(\alpha, \beta, 0, 1)$. Denote the parameter vector in a state-space model as $\theta^{1:p} = (\theta_1, \ldots, \theta_p) \in \Theta^p$. An ideal sampler for this problem would require to

$$\textbf{1.} \text{ sample } \ (\theta^j, x_{0:T}) \sim p(\cdot|y_{1:T}, \theta^{(-j)}), \ \text{ for } \ j = 1, \ldots, m < p, \text{ and} \tag{11}$$

$$\textbf{2.} \text{ sample } \ \theta^j \sim p(\cdot|y_{1:T}, x_{0:T}, \theta^{(-j)}), \ \text{ for } \ j = m+1, \ldots, p, \tag{12}$$

where we have used the short hand notation $\theta^{(-j)}$ to indicate all components in θ except the j-th one. However, direct sampling from $p(x_{0:T}|y_{1:T}, \theta^{1:p})$ and hence from $p(\cdot|y_{1:T}, \theta^{(-j)})$ is not possible. To address this, Andrieu et al. (2010) have proposed two particle based MCMC algorithms, the particle Gibbs sampler (PG) and the particle marginal Metropolis-Hastings algorithm (PMMH). Both algorithms rely on constructing an extended target distribution,

$$p^N(\theta^{1:p}, x_{0:T}^{1:N}, a_{0:T-1}^{1:N}, I|y_{1:T}), \tag{13}$$

where $(X_{0:T}^{1:N}, A_{0:T-1}^{1:N})$ is defined on the space $\mathcal{X}^{(T+1)N} \times \{1, \ldots N\}^{TN}$. If direct sampling from the full conditional distribution of the parameters $p(\theta^{1:p}|y_{1:T}, x_{0:T})$ is available, one can construct a Gibbs sampler on an extended target, where at each iteration the sampler alternates among the following steps:

$$\begin{aligned}
&\textbf{1.} \quad \text{sample } \ \theta^{1:p} \sim p^N(\cdot|y_{1:T}, x_{0:T}^{1:N}, a_{0:T-1}^{1:N}, I),\\
&\textbf{2.} \quad \text{sample } \ \{X_{0:T}^{1:N}, A_{0:T-1}^{1:N}\} \sim p^N(\cdot|y_{1:T}, \theta^{1:p}, I),\\
&\textbf{3.} \quad \text{sample } I \sim p^N(\cdot|y_{1:T}, x_{0:T}^{1:N}, a_{0:T-1}^{1:N}, \theta^{1:p}) = w_T^i.
\end{aligned}$$

Andrieu et al. (2010) propose a conditional sequential Monte Carlo algorithm to obtain samples in step 2. In the conditional SMC one particle is guaranteed to survive all resampling steps and hence we resample only $N-1$ particles. For a more detailed discussion we refer the readers to Andrieu et al. (2010). Alternatively, one could use the procedures proposed by Lindsten et al. (2014) and the references therein for sampling from a conditional SMC.

If the full conditional distributions of the parameters are not available, Andrieu et al. (2010) have proposed a particle marginal Metropolis-Hastings algorithm, which as the name suggests, relies on a particular form of a proposal density $q^N(\theta^{1:p}, x_{0:T}^{1:N}, a_{0:T-1}^{1:N}, I|y_{1:T})$ to form a Metropolis-Hastings (MH) ratio targeting (13). In particular, if at iteration k, for $j = 1, \ldots, p$ one sets:

$$\begin{aligned}
q^N(\theta^{j*}, x_{0:T}^{*1:N}, a_{0:T-1}^{*1:N}, I^*|\theta^{(-j)}) \quad &= \quad q(\theta^{j*}|\theta^j(k-1), \theta^{(-j)}) \\
&\times \quad p(x_{0:T}^{1:N}, a_{0:T-1}^{1:N}|y_{1:T}, \theta^{j*}, \theta^{(-j)}) P(I = i|W_T),
\end{aligned} \tag{14}$$

then the Metropolis-Hastings ratio is given by:

$$1 \wedge \frac{\hat{p}^N(y_{1:T}|\theta^{j*}, \theta^{(-j)})}{\hat{p}^N(y_{1:T}|\theta^j(k-1), \theta^{(-j)})} \times \frac{q(\theta^j(k-1)|\theta^{j*}, \theta^{(-j)})}{q(\theta^{j*}|\theta^j(k-1), \theta^{(-j)})} \times \frac{p(\theta^{j*}|\theta^{(-j)})}{p(\theta^j(k-1)|\theta^{(-j)})}. \tag{15}$$

Upon inspection of (15), we note that the result resembles the ratio of a standard MH step with target given by (13), and likelihood approximation as in (9). Furthermore, Andrieu et al. (2010) prove that the marginal distribution of $(\theta^{1:p}(k), X_{0:T}(k))$, $\mathcal{F}^N$, for the PG and PMMH converges to the target distribution in total variation norm:

$$\left\| \mathcal{F}^N(\theta^{1:p}(k), X_{0:T}(k)|y_{1:T}) - p(\cdot|y_{1:T}) \right\| \to 0, \quad \text{as} \quad k \to \infty. \tag{16}$$

After describing the PMCMC algorithms, we now introduce our proposed sampler. Our motivation for the new sampler stems from the difficulty in certain situations to build adequate proposals for some of the parameters in PMMH algorithms as well as the intractability of the full conditionals for others. Here, we propose to sample (11) via a particle marginal Metropolis-Hastings step and (12) via a standard Gibbs step. By combining the two we show we can obtain more efficient parameter sampling. In contrast to Mendes et al. (2015) which incorporate the idea of sampling the parameters of a state-space model in blocks by PG and PMMH, we propose an algorithm which relies on using SMC only once – in the PMMH step of the algorithm. Our single filter particle Metropolis-within-Gibbs algorithm is given in Algorithm 2.

Algorithm 2 Single Filter PMwG

$k = 0$

Initialize $\theta^j(0)$ for $j = 1, \ldots, p$

Run SMC targeting $p(x_{0:T}|y_{1:T}, \theta^{1:p}(0))$

Sample $X_{0:T}(0) \sim p^N(x_{0:T}|y_{1:T}, \theta^{1:p}(0))$ and calculate $\hat{p}^N(y_{1:T}|\theta^{1:p}(0))$

$k = 1, \ldots, K$

PMMH step

For $j = 1, \ \ldots \ , m < p$

Sample $\theta^{j*} \sim q(\cdot|\theta^j(k-1))$

Run SMC targeting $p(x_{0:T}|y_{1:T}, \theta^{j*}, \theta^{(-j)})$

With probability

$$1 \wedge \frac{\hat{p}^N(y_{1:T}|\theta^{j*}, \theta^{(-j)})}{\hat{p}^N(y_{1:T}|\theta^j(k-1), \theta^{(-j)})} \frac{q(\theta^j(k-1)|\theta^{j*})}{q(\theta^{j*}|\theta^j(k-1))} \frac{p(\theta^{j*}|\theta^{(-j)})}{p(\theta^j(k-1)|\theta^{(-j)})} \tag{17}$$

Set $\theta^j(k) = \theta^{j*}$, $X_{0:T}(k) = X^*_{0:T}$ and $\hat{p}^N(y_{1:T}|\theta^j(k), \theta^{(-j)}) = \hat{p}^N(y_{1:T}|\theta^{j*}, \theta^{(-j)})$; otherwise set $\theta^j(k) = \theta^j(k-1)$, $X_{0:T}(k) = X_{0:T}(k-1)$ and $\hat{p}^N(y_{1:T}|\theta^j(k), \theta^{(-j)}) = \hat{p}^N(y_{1:T}|\theta^j(k-1), \theta^{(-j)})$

Gibbs step

For $j = m+1, \ \ldots \ , p$

Sample $\theta^j(k) \sim p(\cdot|y_{1:T}, \theta^{(-j)}, X_{0:T}(k))$

We present SF-PMwG in general form, where K is the number of iterations. Each parameter is sampled individually conditionally on all other parameters at each iteration. However, in some applications block sampling can improve the efficiency of the algorithm, as we shall see in the following section. To estimate the α-stable stochastic volatility model, we replace the SMC step of the algorithm with APF-ABC. Additionally, if further smoothing of the states is desired, one can adopt ABC based smoothing strategies similar to the forward smoothing discussed in Martin et al. (2014).

More specifically, the first step in Algorithm 2 is a PMMH step targeting the density $p^N(\theta^j, x_{0:T}^{1:N}, a_{0:T-1}^{1:N}, I|y_{1:T}, \theta^{(-j)})$ for $j = 1, \ldots, m < p$. A proper full conditional which will leave the extended target invariant is $p^N(\theta^j|y_{1:T}, x_{0:T}^{1:N}, a_{0:T-1}^{1:N}, I, \theta^{(-j)})$ for $j = m+1, \ldots, p$. However, the Gibbs step in Algorithm 2 samples from a marginalized full conditional $p^N(\theta^j|y_{1:T}, x_{0:T}^i, b_{0:T-1}^i, I, \theta^{(-j)})$, which could potentially cause the sampler to converge to a stationary distribution different from (13). Following Dyk and Jiao (2015) one must sample $\{X_{0:T}^{(-i)}, B_{0:T-1}^{(-i)}\} \sim p^N(\cdot|y_{1:T}, \theta^{1:p}, I)$ to ensure that SF-PMwG leaves (13) invariant. However, $\{X_{0:T}^{(-i)}, B_{0:T-1}^{(-i)}\}$ does not appear in (17) and only modifies the PMMH kernel following a rejection. Moreover, the inferential interest is in the marginal distribution $\mathcal{F}^N$, which does not depend on $\{X_{0:T}^{(-i)}, B_{0:T-1}^{(-i)}\}$. Hence, the additional sampling step may be omitted with no impact on the posterior samples from $\mathcal{F}^N(\theta^{1:p}, X_{0:T}|y_{1:T})$.

5 Simulation Study: Estimation of the α-Stable SVM

In this section we compare the accuracy and efficiency of three algorithms for estimating the asymmetric α-stable SVM. More specifically, we compare SF-PMwG, the algorithm of Mendes et al. (2015) extended to a likelihood free setting and referred to here as DF-PMwG, and PMCMC-ABC (Jasra et al., 2013). The latter algorithm employs SMC-ABC. The single and double filter PMwG algorithms use APF-ABC. We use the following parametrization for the α-stable stochastic volatility model:

$$y_t = \exp\left(\frac{\tilde{x}_t \sigma}{2}\right) v_t, \qquad v_t \sim SD(\alpha, \beta, 0, 1), \tag{18}$$

$$\tilde{x}_t = \gamma + \phi(\tilde{x}_t - \gamma) + w_t, \qquad w_t \sim N(0, 1), \tag{19}$$

$$x_0 \sim N\left(\gamma, \frac{1}{1-\phi^2}\right), \tag{20}$$

where we have let $\tilde{x}_t = x_t/\sigma$ and $\gamma = \mu/\sigma$. We choose this parametrization following Roberts et al. (2004) for efficiency purposes, see also Kastner and Frühwirth-Schnatter (2014).

We assume $\gamma \sim N(\gamma_0, \tau_0)$, $(\phi+1)/2 \sim Beta(a_0, b_0)$, $\sigma \sim IG(c_0, d_0)$, $\alpha \sim U(0, 2)$, and $\beta \sim U(-1, 1)$ to be the prior distributions for all model parameters. To our knowledge, all available methods for the α-stable SVM have used the ABC version of PMMH to estimate all parameters. However, including an ABC step inside PMMH can further affect the accuracy of the approximation in (9), which can lead to higher inefficiency in the estimation procedure. The Gaussian assumption for the distribution of the latent volatility states, and the dependence of the state parameters on $y_{1:T}$ only through $x_{0:T}$, enables us to sample γ and ϕ, respectively, via a full conditional step and a MH step as detailed in Appendix B of the Supplementary Material (Vankov et al., 2019). However, the parameters α, β and σ in (18) can not be sampled via a Gibbs step, because their full conditional distributions depend on the data $y_{1:T}$, for which the p.d.f. is not known in closed form. Therefore, we use SF-PMwG to improve sampling efficiency.

We consider a normal random walk proposal. To account for the support of the parameters we transform and define the new set of parameters as $\theta = \{\log\sigma, \Phi(\alpha/2), \Phi([\beta+1]/2)\} \in \mathbb{R}^3$. The prior used on the transformed parameters is $N(0_{1\times 3}, I_{3\times 3})$ as in Barthelmé and Chopin (2014). For PMCMC-ABC the transformed parameter vector is $\theta = \{\log\sigma, \Phi(\alpha/2), \Phi([\beta+1]/2), \Phi([\phi+1]/2), \mu\} \in \mathbb{R}^5$. To facilitate comparison with literature, nevertheless, we report the final posterior estimates in terms of the original parameters.

The choice of the covariance matrix in the proposal distribution can have a large impact on the performance of the Metropolis-Hastings algorithm. If we set the covariance too large, the algorithm will have very few acceptances. If the covariance is too small, the parameter space is explored poorly. Furthermore, tuning the covariance of the proposal can be very cumbersome and require many test runs. This has lead researchers to develop adaptive MH algorithms, which propose values of the covariance based on the history of the Markov chain. Haario et al. (2001) propose to use a scaled version of the empirical covariance calculated from the Markov chain until the current iteration, as follows,

$$\hat{\Sigma}_k = \begin{cases} \Sigma_0 & \text{if } k \le k_0, \\ c_d\left(\hat{\Sigma}_{\theta(1:k-1)} + \zeta I_d\right) & \text{if } k > k_0, \end{cases} \tag{21}$$

where $\hat{\Sigma}_{\theta(1:k-1)}$ is the empirical covariance of θ until iteration k, and d denotes the dimension of the parameter vector. We set $\Sigma_0 = I_{3\times 3}$, $k_0 = 500$ and $\zeta = 0.0001$. Since $d = 3$ for both SF-PMwG and DF-PMwG, and $d = 5$ for PMCMC-ABC, the scaling constant is set to $c_3 = 2.38^2/3$ and $c_5 = 2.38^2/5$, respectively (Haario et al., 2001; Roberts and Rosenthal, 2009).

To compare the efficiency of the different algorithms, we consider the integrated autocorrelation time, defined by $\text{IACT} = 1 + 2\sum_{l=1}^{\infty}\rho_l$, where ρ_l denotes the autocorrelation function at lag l. Let $\widehat{\text{IACT}} = 1 + 2\sum_{l=1}^{L}\hat{\rho}_l(\theta(1:K))$ be an estimate of IACT, where L is the lag at which the empirical autocorrelation function becomes statistically insignificant. This condition is satisfied whenever $|\hat{\rho}_L(\theta(1:K))| < 2/\sqrt{K}$. Low IACT values indicate that we have more uncorrelated samples generated from the Markov chain implying that we have good mixing. In all of our results, we normalize the estimated IACT by the run time of the algorithms. To compare the accuracy of the algorithms we consider the posterior mean, standard deviation and the 95% credible interval for θ.

We simulate $T = 350$ data points from the model in (18)–(20) with $(\alpha, \beta, \gamma, \phi, \sigma) = (1.7, 0.3, -2, 0.95, 0.2)$. Hence, the value of the mean of the latent state corresponding to (1) is $\mu = -0.4$. The hyperparameters for the prior are set to $a_0 = 40, b_0 = 80/(1+\phi_{true})$, $\gamma_0 = 0, \tau_0 = 10$ as in Kastner and Frühwirth-Schnatter (2014) and $c_0 = 5/2, d_0 = 0.01c_0/2$ as in Kim et al. (1998).

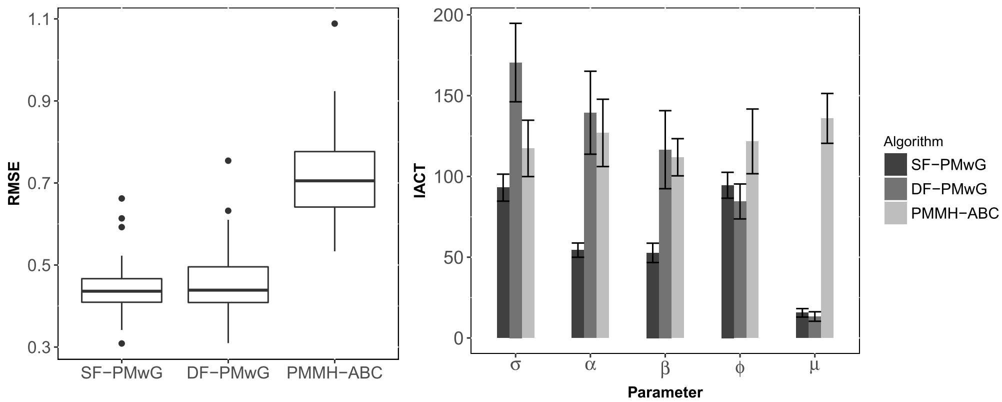

Figure 3: Simulation study in Section 5. Box plots of the root-mean-square error for the latent volatility estimates for SF-PMwG, DF-PMwG and PMMH-ABC (left). Comparison of IACT values of the three algorithms for estimating the α-stable SVM (right). The bars represent average IACT values over 30 runs for each parameter. The standard errors are shown as black lines.

To ensure that all algorithms run for the same amount of time we use $N = 10^4$ particles for the standard PMMH-ABC and SF-PMwG. For DF-PMwG we use $N = 5 \times 10^3$ particles for APF-ABC and $N_c = 1500$ particles for the conditional filter. All algorithms run for $K = 3 \times 10^4$ iterations with the first 5000 discarded as burn-in. The tuning parameters for APF-ABC and SMC-ABC are the same as in Section 3 except we set $\epsilon_g = \epsilon_u = 0.5$ to improve on the accuracy of the filter.

SF-PMwG results in decreased RMSE for the estimation of the latent volatility when compared to PMMH-ABC (Figure 3, left panel). The performance of DF-PMwG is similar to SF-PMwG with the latter having narrower RMSE range. Further, SF-PMwG leads to significant gains in efficiency for σ, α and β (Figure 3, right panel). The efficiency for the parameters sampled by the Gibbs step of the algorithms, μ and ϕ, is similar for the single and double filter PMwG. Both PMwG algorithms yield a lower IACT than PMCMC-ABC with respect to the estimation of ϕ. In this scenario, the inclusion of a Gibbs step has allowed very efficient sampling of μ compared to the PMCMC-ABC algorithm. Moreover, the standard error of the estimated integrated autocorrelation time of SF-PMwG, is significantly smaller than the other two algorithms.

The higher IACT values for σ and ϕ compared to the remaining parameters seen in SF-PMwG are consistent with the findings of Kastner and Frühwirth-Schnatter (2014) for the Gaussian stochastic volatility model.

All algorithms perform similarly for α, β and σ in terms of accuracy (Table 1). The proximity of the posterior mean to the true value and the smaller 95% credible intervals for μ and ϕ indicate the improvement in accuracy resulting from the addition of a Gibbs step.

Table 1: Simulation study in Section 5. Summary of the posterior distribution generated from SF-PMwG, DF-PMwG and PMCMC-ABC for all parameters of the α-stable stochastic volatility model. We present the posterior mean, standard deviation (SD) and 95% credible interval. All quantities are averaged over 30 runs.

Algorithm	SF-PMwG		DF-PMwG		PMMH-ABC	
Parameters	Mean(SD)	95% CI	Mean(SD)	95% CI	Mean(SD)	95% CI
$\alpha = 1.7$	1.691 (0.09)	$(1.51, 1.86)$	1.691 (0.11)	$(1.507, 1.857)$	1.71 (0.1)	$(1.52, 1.867)$
$\beta = 0.3$	0.322 (0.221)	$(-0.119, 0.701)$	0.327 (0.218)	$(-0.116, 0.74)$	0.324 (0.223)	$(-0.128, 0.75)$
$\sigma = 0.2$	0.251 (0.09)	$(0.114, 0.472)$	0.254 (0.09)	$(0.114, 0.475)$	0.256 (0.11)	$(0.09, 0.524)$
$\phi = 0.95$	0.941 (0.04)	$(0.848, 0.992)$	0.94 (0.039)	$(0.841, 0.992)$	0.91 (0.08)	$(0.74, 0.99)$
$\mu = -0.4$	-0.404 (0.331)	$(-1.08, 0.312)$	-0.39 (0.338)	$(-1.02, 0.335)$	-0.42 (0.392)	$(-1.21, 0.33)$

Further, we study the impact of different values of the bandwidths of the Gaussian and uniform kernels used in filtering, ϵ_g and ϵ_u, respectively, different number of particles, different sample sizes and different SVM parameter values (Tables 6–9 in Appendix D of the Supplementary Material (Vankov et al., 2019)). Considering the efficiency of the algorithms (Tables 6 and 7), as measured by IACT, increasing the value of the bandwidth ϵ_g and ϵ_u or increasing the number of particles results in increases efficiency for all algorithms. However, increasing the sample size or decreasing the value of α for the α-stable SVM results in decreased efficiency for all algorithms. Similar to our results in the right panel of Figure 3, when $\epsilon_g = \epsilon_u = 0.5$, the proposed SF-PMwG algorithm has lower IACT for α, β and σ over all simulation scenarios. Moreover, SF-PMwG and DF-PMwG perform similarly for μ and ϕ, outperforming PMMH-ABC. When the bandwidth, ϵ, is increased SF-PMwG continues to outperform for α, β and σ. For the parameters calculated in the Gibbs step, μ and ϕ, SF-PMwG performs similarly or better to the other two algorithms.

Comparing the accuracy of the parameter estimates, as measured by their posterior means and 95% credible intervals, for the different scenarios, we find that increasing the value of ϵ_g and ϵ_u results in less accurate posterior mean estimates for all algorithms. Further, increasing the number of particles can lead to more accurate posterior mean estimates. Increasing the magnitude of the tails of the returns distribution, by lowering the value of α in the SVM, has little effect on the posterior means. As previously noted, direct comparison between the performance of SF-PMwG, DF-PMwG and the PMMH-ABC algorithms for different magnitudes of ϵ_g and ϵ_u should be attempted with caution due to their different interpretation within the context of the Gaussain and uniform density kernels.

6 An Application to the Study of Weekly Spot Prices for Propane

In this section, we apply SF-PMwG to an α-stable stochastic volatility model for weekly propane spot prices.

Propane is a very clean fuel that can be produced from natural gas or crude oil. In the United States it is heavily used in the agricultural and industrial sectors, and as a source for residential and commercial heating. It was estimated that in 2014, the industrial and agricultural sectors, combined for 65% of the total propane used in the United States (Kanderdine, 2014).

Key drivers for propane prices are weather, inventory capacity and transportation infrastructure. Inventories for propane are typically accumulated over the summer months when demand is low. During the winter, the stored propane is withdrawn, transported to the locations in high demand and used for heating. If the winter is severe, the inventory levels are low and the demand increases sharply, large increases in prices can occur. If the infrastructure in place does not allow for timely delivery of propane to the locations in high demand, price fluctuations are additionally exacerbated. For example, in the winter of 2014, the Northeast and Midwest regions of the United States were faced with extreme winter conditions, the stored propane in the region was not sufficient and transportation was difficult, which resulted in propane shortages and large price increases (Kanderdine, 2014). Thus, the stochastic volatility model we consider here can serve policy makers in their effort to minimize the impact of large energy shocks.

To study the distribution of the returns and the volatility, Charfeddine (2014), considered commodity future prices for the period 1994–2009, including propane, with one and three month maturities. More specifically, propane daily returns are found to be negatively skewed and have significant excess kurtosis. When studying month-spot prices for propane for the period 1994–2005, Elder and Serletis (2008), find the distribution to be positively skewed.

We have available weekly spot prices for propane for Mount Belviue, Texas from the Energy Information Administration (EIA) for the time period 10/01/2007 through 04/12/2015 (393 data points), allowing us to study volatility in propane prices for several winter seasons. We should also note that this time period covers the financial crisis of 2008. We select weekly data, as we are interested in the long term behavior of volatility, while discarding as much of the noise inherent in daily trading activities. We calculate the weekly demeaned returns as

$$y_t = 100 \left(\log(P_t/P_{t-1}) - \frac{1}{T} \sum_{t=1}^{T} \log(P_t/P_{t-1}) \right).$$

In order to check for normality of the returns we perform a Jarque–Bera test, for which we reject the null-hypothesis ($p-$ value < 0.001) and conclude that the returns are not normal. Moreover, high sample skewness (-1.44) and kurtosis (8.90) suggest that a Gaussian SVM may not capture all distributional features present in the data (Figure 4).

The prior distribution of the parameters of the α-stable SVM is characterized as in Section 5. Similarly, all hyperparameters are unchanged except we fix $a_0 = 20, b_0 = 1.5$ as in Kim et al. (1998). We apply SF-PMwG to the weekly spot propane prices with $\Sigma_0 = 0.5 I_{3\times 3}$, $c_3 = 2.38^2/3$, and $N = 10^4$ particles for APF-ABC. We run multiple chains of the algorithm for $K = 10^5$ iterations. To assess convergence we use the Gelman–Rubin statistic (after 5×10^4 burn-in). The potential scale reduction factors for all parameters are between 1 and 1.01, indicating that our chains have converged. We combine the chains for posterior parameter inference.

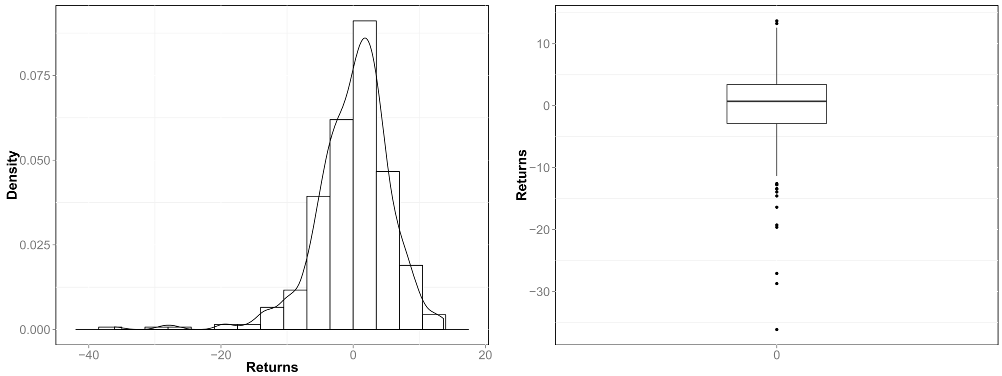

Figure 4: Application to propane prices study in Section 6. Histogram (left) and box-plot (right) for the demeaned weekly spot returns for propane 10/01/2007–04/12/2015, Mount Belvieu TX.

The algorithm captures the heavy tails and the skewness in the returns, as can be seen from the posterior means of α and β respectively (Table 2). The persistence in the volatility is relatively high, indicating high relationship between volatility in successive periods. A moderately varying latent process can be detected from the posterior distribution of σ. Even though, the 95% credible interval for β contains positive values, the posterior probability that the skewness is negative is $P(\beta < 0|y_{1:T}) = 0.948$.

Table 2: Summary of the posterior distribution for $\alpha, \beta, \sigma, \mu, \phi$ for the α-stable SVM applied to weekly returns for propane.

Algorithm	SF-PMwG		DF-PMwG		PMMH-ABC	
Parameters	Mean(SD)	95% CI	Mean(SD)	95% CI	Mean(SD)	95% CI
α	1.88 (0.05)	$(1.76, 1.98)$	1.89 (0.06)	$(1.77, 1.99)$	1.89 (0.06)	$(1.76, 1.98)$
β	−0.58 (0.31)	$(-0.98, 0.22)$	−0.56 (0.34)	$(-0.98, 0.31)$	−0.56 (0.31)	$(-0.97, 0.15)$
σ	0.38 (0.07)	$(0.25, 0.55)$	0.42 (0.08)	$(0.29, 0.61)$	0.42 (0.08)	$(0.29, 0.62)$
μ	2.19 (0.28)	$(1.62, 2.64)$	2.18 (0.24)	$(1.67, 0.2.64)$	2.31 (0.21)	$(1.86, 2.73)$
ϕ	0.89 (0.04)	$(0.8, 0.96)$	0.88 (0.05)	$(0.77, 0.95)$	0.86 (0.06)	$(0.72, 0.95)$

For the time period under consideration, the largest variability in propane prices is observed during the financial crisis of 2008 (Figure 5). The α-stable SVM correctly depicts those large changes in the returns as can be seen from the filtered volatility estimates. Two other volatile periods are the winter of 2014 and 2015. They can be attributed to price increases due to severe winter conditions.

In order to determine the quality of the fit of the α-stable SVM, and to compare it to SVMs with Gaussian and Student's t return errors, we use approximate sequential Bayes factors. The Bayes factor for two competing models, a and b, is calculated via, $BF_{a,b} = p(y_{1:t}|\mathcal{M}_a)/p(y_{1:t}|\mathcal{M}_b)$. The log Bayes factor for comparing the α-stable and

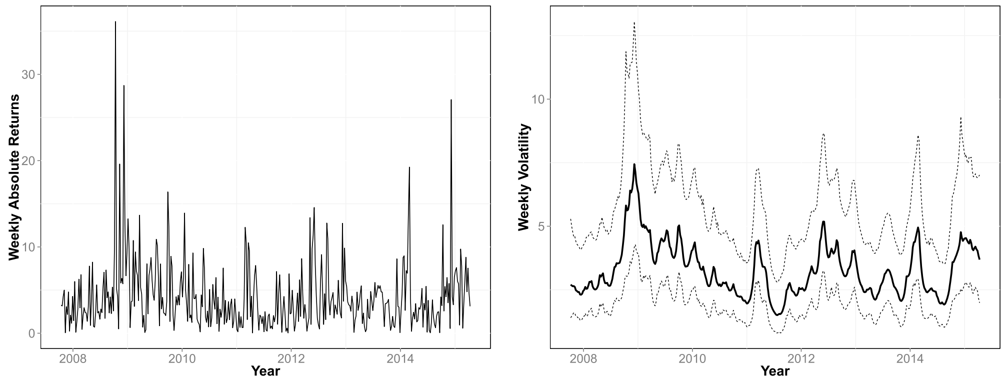

Figure 5: Application to propane prices study in Section 6. Absolute weekly returns for propane October 1, 2007–April 12, 2015 (left) and filtered volatility (right), $\exp(x_t/2)$, with the 95% CI (dashed-lines).

Gaussian SVM is $BF_{\alpha,Gauss} = 20.68$. The log Bayes factor between the α-stable and the Student's t SVM is $BF_{\alpha,t} = 17.27$.

7 Conclusion

In this article we present novel methods for filtering and estimation in state-space models, with a particular application to asymmetric, heavy-tailed α-stable stochastic volatility models. The α-stable stochastic volatility model does not exclude the possibility of Gaussian returns, but allows for more flexible modeling. Wrongly assuming that the returns are normally distributed can lead to underestimation of volatility. Therefore the use of α-stable distribution allows for fewer distributional assumptions prior to estimating the unobserved volatility.

We introduce an ABC based auxiliary particle filter, providing better proposals in a sequential Monte Carlo framework. In particular, by pre-weighting the particles in ABC based SMC using information from the data, we improve on the accuracy of the approximation of the latent volatility states. Such a strategy can prove very useful especially if there are heavy tails present in the data. Even though the proposed filter applies to a richer class of state-space models, we show its application for the asymmetric, heavy-tailed α-stable stochastic volatility model. It is shown through simulation studies that APF-ABC has better performance than the method of Jasra et al. (2012).

We also propose a novel method for joint posterior estimation of the latent states and parameters. Based on ideas from particle Markov chain Monte Carlo methods, we develop a single filter particle Metropolis-within-Gibbs algorithm. The proposed SF-PMwG algorithm is applied to an asymmetric α-stable stochastic volatility model. The inability to express the p.d.f. for the returns for this class of volatility models, highlights the necessity to develop better methods for filtering, such as the proposed APF-ABC.

We find through a simulation study that the proposed SF-PMwG algorithm performs similarly or better than the DF-PMwG of Mendes et al. (2015) and PMMH-ABC of Jasra et al. (2013) under "reasonable" bandwidth values (e.g. $\epsilon_g \leq 0.5$). We apply the single filter PMwG to weekly spot prices for propane for the time period 2007–2015. The algorithm suggests that the returns are negatively skewed and heavy-tailed and hence assuming Gaussian returns could be misleading and lead to underestimation of the inherent returns deviations.

Some inherent drawbacks within PMCMC methods, including the single filter particle Metropolis-within-Gibbs algorithm, are the computational times required for the chain to converge. One way to further increase the efficiency of our procedure could be to employ other adaptive MH procedures in the PMMH step of SF-PMwG (see for example, Dahlin et al., 2015). In addition, in the APF-ABC filter we use multinomial resampling; however, it will be interesting to implement other resampling strategies (e.g. systematic, residual etc.) and study their impact on the filter. We also plan to investigate how to extend the methods proposed here to multivariate settings, accounting for the presence of contemporaneous correlations (Chib et al., 2006).

Supplementary Material

Supplementary Material for Filtering and Estimation for a Class of Stochastic Volatility Models with Intractable Likelihoods (DOI: 10.1214/18-BA1099SUPP; .pdf).

References

Abanto-Valle, C. A., Bandyopadhyay, D., Lachos, V. H., and Enriquez, I. (2010). "Robust Bayesian analysis of heavy-tailed stochastic volatility models using scale mixtures of normal distributions." *Computational Statistics & Data Analysis*, 54(12): 2883–2898. MR2727721. doi: https://doi.org/10.1016/j.csda.2009.06.011. 32

Andrieu, C., Doucet, A., and Holenstein, R. (2010). "Particle Markov chain Monte Carlo methods." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(3): 269–342. MR2758115. doi: https://doi.org/10.1111/j.1467-9868.2009.00736.x. 30, 36, 39, 40

Barthelmé, S. and Chopin, N. (2014). "Expectation propagation for likelihood-free inference." *Journal of the American Statistical Association*, 109(505): 315–333. MR3180566. doi: https://doi.org/10.1080/01621459.2013.864178. 33, 42

Bollerslev, T. (1986). "Generalized autoregressive conditional heteroskedasticity." *Journal of Econometrics*, 31: 307–327. MR0853051. doi: https://doi.org/10.1016/0304-4076(86)90063-1. 29

Buckle, D. (1995). "Bayesian inference for stable distributions." *Journal of the American Statistical Association*, 90(430): 605–613. MR1340513. 32

Carpenter, J., Clifford, P., and Fearnhead, P. (1999). "Improved particle filter for nonlinear problems." *IEE Proceedings-Radar, Sonar and Navigation*, 146(1): 2–7. 30, 34

Chambers, J. M., Mallows, C. L., and Stuck, B. (1976). "A method for simulating stable random variables." *Journal of the American Statistical Association*, 71(354): 340–344. MR0415982. 32

Charfeddine, L. (2014). "True or spurious long memory in volatility: Further evidence on the energy futures markets." *Energy Policy*, 71: 76–93. 45

Chib, S., Nardari, F., and Shephard, N. (2002). "Markov chain Monte Carlo methods for stochastic volatility models." *Journal of Econometrics*, 108(2): 281–316. MR1894758. doi: `https://doi.org/10.1016/S0304-4076(01)00137-3`. 32

Chib, S., Nardari, F., and Shephard, N. (2006). "Analysis of high dimensional multivariate stochastic volatility models." *Journal of Econometrics*, 134(2): 341–371. MR2328413. doi: `https://doi.org/10.1016/j.jeconom.2005.06.026`. 48

Creal, D. (2012). "A survey of sequential Monte Carlo methods for economics and finance." *Econometric Reviews*, 31(3): 245–296. 31

Dahlin, J., Lindsten, F., and Schön, T. B. (2015). "Particle Metropolis–Hastings using gradient and Hessian information." *Statistics and Computing*, 25(1): 81–92. MR3304908. doi: `https://doi.org/10.1007/s11222-014-9510-0`. 48

Del Moral, P., Doucet, A., and Singh, S. (2010). "Forward smoothing using sequential Monte Carlo." *arXiv preprint arXiv:1012.5390*. 37

Douc, R. and Cappé, O. (2005). "Comparison of resampling schemes for particle filtering." In *Image and Signal Processing and Analysis, 2005. ISPA 2005. Proceedings of the 4th International Symposium on*, 64–69. IEEE. 35

Doucet, A. (2001). *Sequential Monte Carlo methods*. Wiley Online Library. 31

Dyk, D. A. V. and Jiao, X. (2015). "Metropolis-Hastings within partially collapsed Gibbs samplers." *Journal of Computational and Graphical Statistics*, 24(2): 301–327. MR3357383. doi: `https://doi.org/10.1080/10618600.2014.930041`. 41

Elder, J. and Serletis, A. (2008). "Long memory in energy futures prices." *Review of Financial Economics*, 17(2): 146–155. 45

Engle, R. F. (1982). "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation." *Econometrica*, 50(4): 987–1007. MR0666121. doi: `https://doi.org/10.2307/1912773`. 29

Godsill, S. J., Doucet, A., and West, M. (2004). "Monte Carlo smoothing for nonlinear time series." *Journal of the American Statistical Association*, 99(465): 156–168. MR2054295. doi: `https://doi.org/10.1198/016214504000000151`. 37

Haario, H., Saksman, E., and Tamminen, J. (2001). "An adaptive Metropolis algorithm." *Bernoulli*, 223–242. MR1828504. doi: `https://doi.org/10.2307/3318737`. 42

Harvey, A., Ruiz, E., and Shephard, N. (1994). "Multivariate stochastic variance models." *The Review of Economic Studies*, 61(2): 247–264. 32

Hull, J. and White, A. (1987). "The pricing of options on assets with stochastic volatilities." *The Journal of Finance*, 42(2): 281–300. 29

Jacquier, E., Polson, N. G., and Rossi, P. E. (1994). "Bayesian analysis of stochastic volatility models." *Journal of Business & Economic Statistics*, 12(4): 371–389. MR1940631. doi: https://doi.org/10.1198/073500102753410408. 32

Jacquier, E., Polson, N. G., and Rossi, P. E. (2004). "Bayesian analysis of stochastic volatility models with fat-tails and correlated errors." *Journal of Econometrics*, 122(1): 185–212. MR2083256. doi: https://doi.org/10.1016/j.jeconom.2003.09.001. 32

Jasra, A. (2015). "Approximate Bayesian computation for a class of time series models." *International Statistical Review*, 83(3): 405–435. MR3429284. doi: https://doi.org/10.1111/insr.12089. 33, 38

Jasra, A., Lee, A., Yau, C., and Zhang, X. (2013). "The alive particle filter." *arXiv preprint arXiv:1304.0151*. 30, 33, 41, 48

Jasra, A., Singh, S., S Martin, J. S., and McCoy, E. (2012). "Filtering via approximate Bayesian computation." *Statistics and Computing*, 22(6): 1223–1237. MR2992296. doi: https://doi.org/10.1007/s11222-010-9185-0. 33, 36, 47

Jensen, M. J. and Maheu, J. M. (2010). "Bayesian semiparametric stochastic volatility modeling." *Journal of Econometrics*, 157(2): 306–316. MR2661603. doi: https://doi.org/10.1016/j.jeconom.2010.01.014. 32

Johansen, A. M. and Doucet, A. (2008). "A note on auxiliary particle filters." *Statistics & Probability Letters*, 78(12): 1498–1504. MR2528342. doi: https://doi.org/10.1016/j.spl.2008.01.032. 34

Kanderdine, M. (2014). "*Hearing on short on gas: a look into the propane shortages this winter*." United States. Senate. Committee on Energy and Natural Resources 113 Cong. 2nd sess. Washington: GPO 2014. 45

Kastner, G. and Frühwirth-Schnatter, S. (2014). "Ancillarity-sufficiency interweaving strategy (ASIS) for boosting MCMC estimation of stochastic volatility models." *Computational Statistics & Data Analysis*, 76: 408–423. MR3209449. doi: https://doi.org/10.1016/j.csda.2013.01.002. 30, 32, 41, 42, 43

Kim, S., Shephard, N., and Chib, S. (1998). "Stochastic volatility: likelihood inference and comparison with ARCH models." *The Review of Economic Studies*, 65(3): 361–393. 30, 32, 42, 45

Kitagawa, G. (1996). "Monte Carlo filter and smoother for non-Gaussian nonlinear state space models." *Journal of Computational and Graphical Statistics*, 5(1): 1–25. MR1380850. doi: https://doi.org/10.2307/1390750. 36

Lindsten, F., Jordan, M. I., and Schön, T. B. (2014). "Particle Gibbs with ancestor sampling." *The Journal of Machine Learning Research*, 15(1): 2145–2184. MR3231604. 39

Lombardi, M. J. (2007). "Bayesian inference for alpha-stable distributions: a random walk MCMC approach." *Computational Statistics & Data Analysis*, 51(5): 2688–2700. MR2338996. doi: https://doi.org/10.1016/j.csda.2006.01.009. 32

Lombardi, M. J. and Calzolari, G. (2009). "Indirect estimation of α-stable stochastic volatility models." *Computational Statistics & Data Analysis*, 53(6): 2298–2308. MR2665829. doi: https://doi.org/10.1016/j.csda.2008.11.016. 32

Mandelbrot, B. (1963). "The Variation of Certain Speculative Prices." *The Journal of Business*, 36(4): 394–419. 30

Martin, J. S., Jasra, A., Singh, S. S., Whiteley, N., Del Moral, P., and McCoy, E. (2014). "Approximate Bayesian computation for smoothing." *Stochastic Analysis and Applications*, 32(3): 397–420. MR3197814. doi: https://doi.org/10.1080/07362994.2013.879262. 36, 37, 41

McCulloch, J. H. (1986). "Simple consistent estimators of stable distribution parameters." *Communications in Statistics-Simulation and Computation*, 15(4): 1109–1136. MR0876783. doi: https://doi.org/10.1080/03610918608812563. 32

Mendes, E. F., Carter, C. K., and Kohn, R. (2015). "On general sampling schemes for particle Markov chain Monte Carlo methods." *arXiv preprint arXiv:1401.1667*. 30, 40, 41, 48

Minka, T. P. (2001). "Expectation propagation for approximate Bayesian inference." In *Proceedings of the Seventeenth conference on Uncertainty in artificial intelligence*, 362–369. Morgan Kaufmann Publishers Inc. 33

Nelson, D. B. (1991). "Conditional heteroskedasticity in asset returns: a new approach." *Econometrica*, 59: 347–370. MR1097532. doi: https://doi.org/10.2307/2938260. 29

Nolan, J. P. (1997). "Numerical calculation of stable densities and distribution functions." *Communications in Statistics. Stochastic models*, 13(4): 759–774. MR1482292. doi: https://doi.org/10.1080/15326349708807450. 30

Peters, G. W., Fan, Y., and Sisson, S. A. (2012). "On sequential Monte Carlo, partial rejection control and approximate Bayesian computation." *Statistics and Computing*, 22(6): 1209–1222. MR2992295. doi: https://doi.org/10.1007/s11222-012-9315-y. 32

Pitt, M. K. and Shephard, N. (1999). "Filtering via simulation: auxiliary particle filters." *Journal of the American Statistical Association*, 94(446): 590–599. MR1702328. doi: https://doi.org/10.2307/2670179. 30, 31, 34

Press, S. J. (1972). "Estimation in univariate and multivariate stable distributions." *Journal of the American Statistical Association*, 67(340): 842–846. MR0362666. 32

Roberts, G. O., Papaspiliopoulos, O., and Dellaportas, P. (2004). "Bayesian inference for non-Gaussian Ornstein–Uhlenbeck stochastic volatility processes." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(2): 369–393. MR2062382. doi: https://doi.org/10.1111/j.1369-7412.2004.05139.x. 41

Roberts, G. O. and Rosenthal, J. S. (2009). "Examples of adaptive MCMC." *Journal of Computational and Graphical Statistics*, 18(2): 349–367. MR2749836. doi: https://doi.org/10.1198/jcgs.2009.06134. 42

Shephard, N. (ed.) (2005). *Stochastic Volatility: Selected Readings*. Advanced Texts in Econometrics. Oxford University Press. MR2203295. 32

Taylor, S. J. (1994). "Modeling stochastic volatility: a review and comparative study." *Mathematical Finance*, 4(2): 183–204. 29, 31

Vankov, E. R., Guindani, M., and Ensor, K. B. (2019). "Supplementary Material for Filtering and Estimation for a Class of Stochastic Volatility Models with Intractable Likelihoods." *Bayesian Analysis*. doi: https://doi.org/10.1214/18-BA1099SUPP. 32, 36, 38, 42, 44