

## Approaches Toward the Bayesian Estimation of the Stochastic Volatility Model with Leverage

Hosszejni, Darjus; Kastner, Gregor

*Published in:*

Bayesian Statistics and New Generations - Selected Contributions from BAYSM 2018

Published: 01/01/2019

*Document Version*

Early version, also known as preprint

[Link to publication](#)

*Citation for published version (APA):*

Hosszejni, D., & Kastner, G. (2019). Approaches Toward the Bayesian Estimation of the Stochastic Volatility Model with Leverage. In Raffaele Argiento, Daniele Durante, Sara Wade (Ed.), *Bayesian Statistics and New Generations - Selected Contributions from BAYSM 2018* (pp. 75 - 83). Springer.  
<https://arxiv.org/abs/1901.11491>

# Approaches Toward the Bayesian Estimation of the Stochastic Volatility Model with Leverage

Darjus Hosszejni\*      Gregor Kastner†

February 1, 2019

**With very minor editorial changes, this article is published as:**

Darjus Hosszejni and Gregor Kastner. Approaches toward the Bayesian estimation of the stochastic volatility model with leverage. In Raffaele Argiento, Daniele Durante, and Sara Wade, editors, *Bayesian Statistics and New Generations – Selected Contributions from BAYSM 2018*, volume 296 of *Springer Proceedings in Mathematics & Statistics*, pages 75–83, Cham, 2019. Springer. doi:10.1007/978-3-030-30611-3\_8.

## Abstract

The sampling efficiency of MCMC methods in Bayesian inference for stochastic volatility (SV) models is known to highly depend on the actual parameter values, and the effectiveness of samplers based on different parameterizations varies significantly. We derive novel algorithms for the centered and the non-centered parameterizations of the practically highly relevant SV model with leverage, where the return process and innovations of the volatility process are allowed to correlate. Moreover, based on the idea of ancillarity-sufficiency interweaving (ASIS), we combine the resulting samplers in order to guarantee stable sampling efficiency irrespective of the baseline parameterization. We carry out an extensive comparison to already existing sampling methods for this model using simulated as well as real world data.

**Keywords:** ancillarity-sufficiency interweaving strategy (ASIS), auxiliary mixture sampling, Bayesian inference, Markov chain Monte Carlo (MCMC), state-space model.

---

\*WU Vienna University of Economics and Business, Austria. Email: darjus.hosszejni@wu.ac.at.

†WU Vienna University of Economics and Business, Austria. Email: gregor.kastner@wu.ac.at.

# 1 Introduction and Model Specification

Stochastic volatility (SV) models (Taylor, 1982) are an increasingly popular choice for modeling financial return data. The basic SV model assumes an autoregressive structure for the log-volatility, and it is able to match the empirically observable low serial autocorrelation in the return series but high serial autocorrelation in the squared return series. The SV model with leverage (SVL, Harvey and Shephard, 1996) extends the SV model by allowing the return series and the increment of the log-volatility series to correlate. This correlation models a real world phenomenon, the asymmetric relationship between returns and their volatility.

SVL, in its centered parameterization (C), is typically formulated as

$$\begin{aligned}y_t &= \exp(h_t/2)\varepsilon_t, \\h_{t+1} &= \mu + \varphi(h_t - \mu) + \sigma\eta_t, \\ \text{cor}(\varepsilon_t, \eta_t) &= \rho,\end{aligned}\tag{1}$$

for  $t = 1, \dots, T$ , where  $\varepsilon_t, \eta_t \sim$  i.i.d.  $\mathcal{N}(0, 1)$ . The only observed variable is  $\vec{y} = (y_1, \dots, y_T)'$ , usually some de-measured financial return series. An AR(1) structure is assumed for the latent log variance  $\vec{h} = (h_1, \dots, h_T)'$ , with mean  $\mu$ , persistence  $\varphi$ , and increment volatility  $\sigma$ . The leverage effect is captured by  $\rho$ , which is zero in the basic SV model by assumption.

An equivalent specification, called the non-centered parameterization (NC), can be obtained by substituting  $\tilde{h}_t = (h_t - \mu)/\sigma$  into (1), thereby moving  $\mu$  and  $\sigma$  from the state equation to the observation equation. The resulting formulation is given by

$$\begin{aligned}y_t &= \exp((\mu + \sigma\tilde{h}_t)/2)\varepsilon_t, \\ \tilde{h}_{t+1} &= \varphi\tilde{h}_t + \eta_t.\end{aligned}\tag{2}$$

Common priors are chosen from the literature:  $(\varphi + 1)/2 \sim \text{Beta}(a_\varphi, b_\varphi)$ ,  $(\rho + 1)/2 \sim \text{Beta}(a_\rho, b_\rho)$ ,  $\sigma^2 \sim \text{Gamma}(\alpha_\sigma, \beta_\sigma)$ ,  $\mu \sim \mathcal{N}(\mu_\mu, \sigma_\mu^2)$ ,  $h_1 \sim \mathcal{N}(\mu, \sigma^2/(1 - \varphi^2))$  (Frühwirth-Schnatter and Wagner, 2010; Kastner and Frühwirth-Schnatter, 2014; Omori et al., 2007).

While the SV model is accessible through the R (R Core Team, 2018) package `stochvol` (Kastner, 2016), it does not cater for the leverage effect, and, to the best of our knowledge, there is no implementation of SVL that works out-of-the-box in a free, open source environment. Our goal is to extend the package with an easy-to-use MCMC sampler that performs reasonably well on a diverse variety of data sets. To this end, we compare various sampling algorithms through a large simulation study from a practical viewpoint. In doing so, it is important to note that `stochvol` is often used as a subsampler for hierarchical models such as (vector auto) regressions or multivariate (factor) SV models. Consequently, in order to use the extended package in a similar manner, adaptive

algorithms are not preferred, as their adaptation state can be cumbersome to implement within a larger MCMC scheme.

## 2 Estimation Strategies

The state-of-the-art solution (Omori et al., 2007) for estimating  $\vec{h}$  is based on linearizing the observation equation in (2), and employing a ten-component bivariate Gaussian mixture approximation to the joint law of  $(\log \varepsilon_t^2, \eta_t)$  separately for each time point, thus introducing a new array of latent variables  $s_t \in \{1, \dots, 10\}$ ,  $t = 1, \dots, T$ , encoding the mixture components. The resulting conditionally Gaussian state space can be written as

$$\begin{aligned} y_t^* &= \mu + \sigma \tilde{h}_t + m_{s_t}^{(1)} + v_{s_t}^{(1)} w_t, \\ \tilde{h}_{t+1} &= \varphi \tilde{h}_t + \sqrt{1 - \rho^2} z_t + d_t \rho (m_{s_t}^{(2)} + v_{s_t}^{(2)} w_t), \end{aligned} \quad (3)$$

where  $y_t^* = \log(y_t^2)$ ,  $d_t = \text{sgn}(y_t)$ ,  $w_t, z_t \sim$  i.i.d.  $\mathcal{N}(0, 1)$  for  $t = 1, \dots, T$ , and  $m_j^{(i)}, v_j^{(i)}$  are model-independent constants for  $i = 1, 2$ , and  $j = 1, \dots, 10$ , defined in Omori et al. (2007).

Let  $\vec{\vartheta} = (\varphi, \rho, \sigma, \mu)'$ , and  $\vec{s} = (s_1, \dots, s_T)'$ . The sampling algorithm of the auxiliary model (AUX) consists of repeating the steps below.

- Step 1: Draw  $\vec{s} \mid \vec{y}, \vec{h}, \vec{\vartheta}$  using inverse transform sampling with the posterior probabilities calculated as in Section 2.3.2 of Omori et al. (2007).
- Step 2: Draw  $\varphi, \rho, \sigma \mid \vec{y}, \vec{s}$  via an independent Metropolis-Hastings (MH) step utilizing the Laplace approximation of the collapsed distribution of  $\varphi, \rho, \sigma \mid \vec{y}, \vec{s}$  as the proposal. The calculation of the acceptance ratio includes Kalman filter evaluations, numerical optimization, and numerical differentiation.
- Step 3: Draw  $\mu \mid \vec{y}, \vec{s}, \varphi, \rho, \sigma$ , and then  $\vec{h} \mid \vec{y}, \vec{s}, \vec{\vartheta}$ , using Gaussian simulation smoothing (Carter and Kohn, 1994; Frühwirth-Schnatter, 1994).

At least three issues arise with this method. First, due to the involvement of Kalman filter evaluations and the numerical optimization part in Step 2, the execution time of the sampler is significantly worse than the runtime of methods with more naïve proposals, e.g. MH algorithms based on sampling from the full conditional distribution. According to our measurements, Step 2 requires around 80% of the total runtime. Second, for extreme data sets, the sampler might get stuck in a state and be unable to accept a new state for many iterations. Third, and finally, the numerical optimization step is sensitive to its configuration, possibly returning a negative semi-definite Hessian matrix at the found mode.

Hence, for parameter sampling, we replace Step 2 by a random-walk MH (RWMH) method which estimates (1) or (2) without resorting to the auxiliary mixture approximation. For the latent vector, we again use the highly efficient Step 3 of AUX as a proposal, followed by an MH acceptance-rejection step to correct for the difference between models (1) and (3).

As already shown for SV (Kastner and Frühwirth-Schnatter, 2014; McCausland et al., 2011), samplers based on different parameterizations can have substantially different sampling efficiencies on the same data set due to the altered dependence structure. To exploit this phenomenon, the ancillarity-sufficiency interweaving strategy (ASIS, Yu and Meng, 2011) can utilize samplers of both C and NC, and thus ASIS may be able to deliver a markedly higher effective sample size than samplers based on a single parameterization only.

The RWMH sampling algorithm estimates SVL by repeating the steps below.

- Step 1: Draw  $\vec{h} \mid \vec{y}, \vec{\vartheta}$ . A candidate  $\vec{h}^*$  is proposed using the AUX sampler by drawing  $\vec{s} \mid \vec{y}, \vec{h}, \vec{\vartheta}$  and then drawing  $\vec{h} \mid \vec{y}, \vec{s}, \vec{\vartheta}$  as explained in Steps 1 and 3 of algorithm AUX. Subsequently,  $\vec{h}^*$  is accepted with probability

$$\min \left\{ 1, \frac{p_C(\vec{h}^* \mid \vec{y}, \vec{\vartheta}) p_A(\vec{h} \mid \vec{y}, \vec{\vartheta})}{p_C(\vec{h} \mid \vec{y}, \vec{\vartheta}) p_A(\vec{h}^* \mid \vec{y}, \vec{\vartheta})} \right\},$$

where  $p_C$  and  $p_A$  denote the corresponding posteriors resulting from specifications (1) and (3), respectively.

- Step 2: Draw  $\vec{\vartheta} \mid \vec{y}, \vec{h}$ . In order to avoid a possibly problematic truncation of the proposal distribution, the parameter vector  $\vec{\vartheta}$  is transformed from  $(-1, 1) \times (-1, 1) \times (0, \infty) \times \mathbb{R}$  to  $\mathbb{R}^4$  by applying the transformation  $x \mapsto 0.5 \log((1+x)/(1-x))$  to  $\varphi$  and  $\rho$ , and by taking the logarithm of  $\sigma^2$ . Then, in the resulting unbounded space, a simple four-dimensional Gaussian random walk is proposed. Its innovation covariance matrix elements are fixed at 0.1 on the diagonal and zero elsewhere.
- If ASIS is applied, then, after Step 2,  $\tilde{\vec{h}} = (\tilde{h}_1, \dots, \tilde{h}_T)'$  is calculated using the new values of  $\sigma$ ,  $\mu$ , followed by a new draw from  $\vec{\vartheta} \mid \vec{y}, \tilde{\vec{h}}$ . Finally, in order to move back to the original parameterization,  $\vec{h}$  is recalculated from  $\tilde{\vec{h}}$  and the new values of  $\sigma$ ,  $\mu$ .

ASIS is a natural extension to the RWMH samplers for the centered and the non-centered parameterizations. However, in the case of AUX, resampling in a different parameterization is detrimental to sampling efficiency for two reasons. First, in Step 2, the parameters  $\varphi$ ,  $\rho$  and  $\sigma$  are drawn from a collapsed distribution that is independent of  $\vec{h}$ . Consequently, ASIS provides only negligible gains in sampling efficiency. Second, if

ASIS were applied to AUX, the computationally most expensive parts of Step 2 would be repeated, thereby increasing the execution time by around 80%.

### 3 Simulation Study

In order to assess the efficiency of our estimation algorithms for the parameter vector  $\vec{\vartheta}$ , we simulate data using SVL from an extensive grid of data generating processes (DGPs). The parameters  $\varphi_{\text{true}}, \rho_{\text{true}}, \sigma_{\text{true}}$  vary on a  $\{0, 0.5, 0.9, 0.95, 0.99\} \times \{-0.6, -0.3, 0, 0.3, 0.6\} \times \{0.1, 0.3, 0.5\}$  grid. For the sake of readability,  $\mu_{\text{true}}$  is set to  $-9$  in all cases, resulting in 75 distinct parameter settings. This choice covers previously investigated ranges (Jacquier et al., 2004; Kastner and Frühwirth-Schnatter, 2014). After the burn-in, respectively, adaptation phase, 50 000 MCMC draws are obtained from the posterior distribution. We repeat this exercise for ten data sets of length 300, and ten data sets of length 3000, for eight sampling algorithms: AUX, Stan-C, Stan-N, JAGS-C, JAGS-N, RWMH-C, RWMH-N, and RWMH-ASISx5, where C and N stand for the centered and, respectively, non-centered parameterization, while ASISx5 denotes the algorithm repeating the two steps of ASIS five times after each draw of  $\vec{h}$ , which in general we found to be superior to executing the two ASIS steps only once. Note, that, although they do not fit our needs due to their adaptation phase, we include Stan (Carpenter et al., 2017) and JAGS (Plummer, 2003) as benchmarks, and all reported results are based on the chain after adaptation has stopped. We fix the priors throughout the simulation study to  $a_\varphi = 20, b_\varphi = 1.5, a_\rho = 3, b_\rho = 6, \alpha_\sigma = 0.5, \beta_\sigma = 0.5, \mu_\mu = -10$ , and  $\sigma_\mu^2 = 100$ . The prior hyperparameters of  $\varphi, \sigma^2$ , and  $\mu$  are chosen from previous studies (Kastner and Frühwirth-Schnatter, 2014; Nakajima and Omori, 2009), and the slightly informative prior on  $\rho$  is chosen to improve the estimation process of Stan-C and of AUX in the extremes of the parameter grid. However, results not reported here due to space constraints indicate that when  $T = 300$ , the posterior distribution of  $\rho$  is only mildly affected by this choice compared to a uniform prior, whereas when  $T = 3000$  the differences are barely noticeable.

The resulting 12 000 MCMC chains were computed on a cluster of computers consisting of 400 Intel E5 2.3 GHz cores running R version 3.4.3. The Stan and the JAGS models were estimated using `rstan` (Carpenter et al., 2017) version 2.17.3 and `rjags` (Plummer, 2003) version 4-6. The RWMH samplers and AUX were based on our computationally efficient `Rcpp` (Eddelbuettel and François, 2011) implementation. The typical runtime of the samplers is summarized in Table 1. Inefficiency factors and effective sample sizes were calculated using the `coda` (Plummer et al., 2006) package, data analysis and visualization was done with the help of the `tidyverse` (Wickham, 2017) package.

We assess the statistical efficiency of the different competitors through their inefficiency factor (IF), an estimator for the integrated autocorrelation time  $\tau$ , given by  $\tau = 1 + 2 \sum_{t=1}^{\infty} \rho_{\text{auto}}(t)$ , where  $\rho_{\text{auto}}(t)$  denotes the autocorrelation function at lag  $t$ . For

Stan-C	Stan-N	JAGS-C	JAGS-N	RWMH	RWMH-ASISx5	AUX
90–642	59–441	22–31	50–106	6–21	14–29	44–86

Table 1: Typical execution times (in minutes) for 50 000 draws after the burn-in when  $T = 3000$ . The displayed values correspond to the first and the ninth deciles of all wall clock times. The choice of the parameterization affects the execution time when JAGS or Stan is used and thus in these cases runtimes are shown separately for C and NC.

an MCMC sample  $S$ , the IFs reported here are calculated as  $\text{IF}_S = n_S/\text{ESS}_S$  (Kastner and Frühwirth-Schnatter, 2014), where  $n_S$  is the size of  $S$ , and  $\text{ESS}_S$  stands for the effective sample size of  $S$ , the size of a serially uncorrelated sample having the same Monte Carlo standard error as  $S$ . A good sampler has low serial correlation, thus the aim is to provide samples with low IF, or, in other words, high ESS. In practice, computational speed is comparably important to computational efficiency. Hence, the final assessment is based on the effective sampling rate (ESR), defined as the ESS divided by the execution time. We note that incorporating runtime in the assessment of algorithms may be problematic due to inconsistent implementations (Kriegel et al., 2016); however, as one of our objectives is a software package, we consider the computational speed an essential part of our study.

### 3.1 Collapsed vs. Full Conditional Sampling

AUX employs a well-known technique for improving the statistical efficiency of MCMC simulations by using a collapsed distribution for sampling  $\varphi$ ,  $\rho$ , and  $\sigma$ . This means that some variables are marginalized out in order to decrease the serial dependence in the chain (Liu, 1994). ASIS, on the other hand, takes advantage of being able to reorganize the dependence. Which technique is superior in practice largely depends on computational aspects. Figure 1 exemplifies the problem by displaying the autocorrelograms of the outputs of RWMH-ASISx5 and AUX, for the parameters  $\mu$ ,  $\varphi$ ,  $\rho$ , and  $\sigma$ , based on a selected data set. The figure illustrates the execution time as well: in both columns, the number of solid lines indicates the average number of samples drawn in 0.1 seconds. Thus, in each facet, the height of the rightmost solid line visualizes the ESR for the given parameter and sampler. Although the autocorrelation functions of AUX decay faster than the ones of RWMH-ASISx5, the latter counterbalances its disadvantages by its speed. Note, however, that different DGPs tend to produce qualitatively different pictures, making the choice between AUX and RWMH-ASISx5 non-trivial.

### 3.2 Efficiency Overview

The minimal ESR is the minimum taken over the ESRs of  $\varphi$ ,  $\rho$ ,  $\sigma$ , and  $\mu$ , and, thus, it measures the speed of discovering the joint posterior  $p(\vec{\vartheta} \mid \vec{y})$ . In order to provide

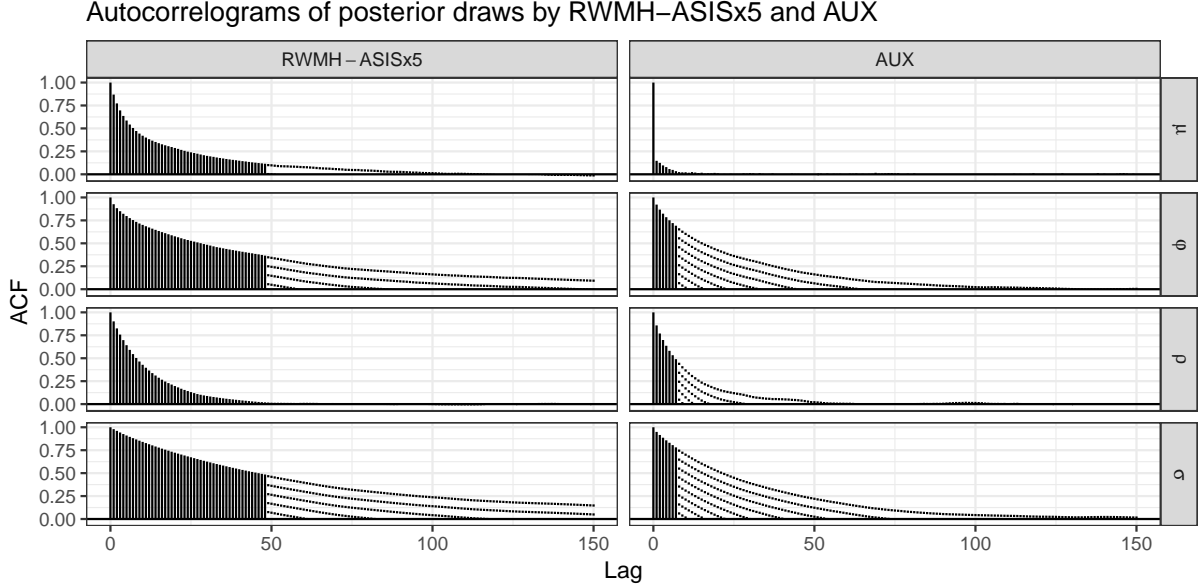


Figure 1: Autocorrelation functions of the posterior draws for  $\mu$ ,  $\varphi$ ,  $\rho$ , and  $\sigma$ , using RWMH-ASISx5 and AUX, for an illustrative example where  $\varphi_{\text{true}} = 0.95$ ,  $\rho_{\text{true}} = -0.3$ ,  $\sigma_{\text{true}} = 0.3$ , and  $T = 300$ . The line type indicates the speed of the Monte Carlo simulation: the number of solid lines equals the average number of samples drawn in 0.1 seconds.

an overview, Figure 2 displays the minimal ESRs for each sampler and strategy, and for all DGPs with  $\rho_{\text{true}} = -0.3$  and  $T = 3000$ . Taking into account that Stan and JAGS are general-purpose probabilistic modeling frameworks, they perform surprisingly well compared to AUX and our RWMH implementations which have been developed specifically for the model at hand. However, the absence of a best performing method is eye-catching. In particular, the choice between AUX and RWMH is noticeably difficult.

In terms of variability, note that the ESRs of AUX range from below 0.001 to above 1, while the ESRs of RWMH-ASISx5 fall between 0.01 and 0.1. This renders the latter more stable by around two orders of magnitude.

## 4 Application to Financial Data Sets

We apply the eight estimation methods to seven univariate time series of daily financial log-returns covering four asset types and two economic periods. The first time interval is a booming, pre-crisis period starting from 2005-01-01 and ending on 2007-12-31, including a total of 872 business days. The second interval is a more recent, more volatile period between 2015-01-01 and 2018-06-30, including 1014 business days. The series under consideration are the Bitcoin price in USD (ticker: BTCUSD=X), hereafter BTC, the German DAX index (ticker: ^GDAXI), the EUR/USD exchange rate (ticker: EURUSD=X), hereafter EUR, and a large German company, the Merck KG's equity (ticker: MRK.DE), hereafter MRK. The data is provided by Yahoo! Finance.



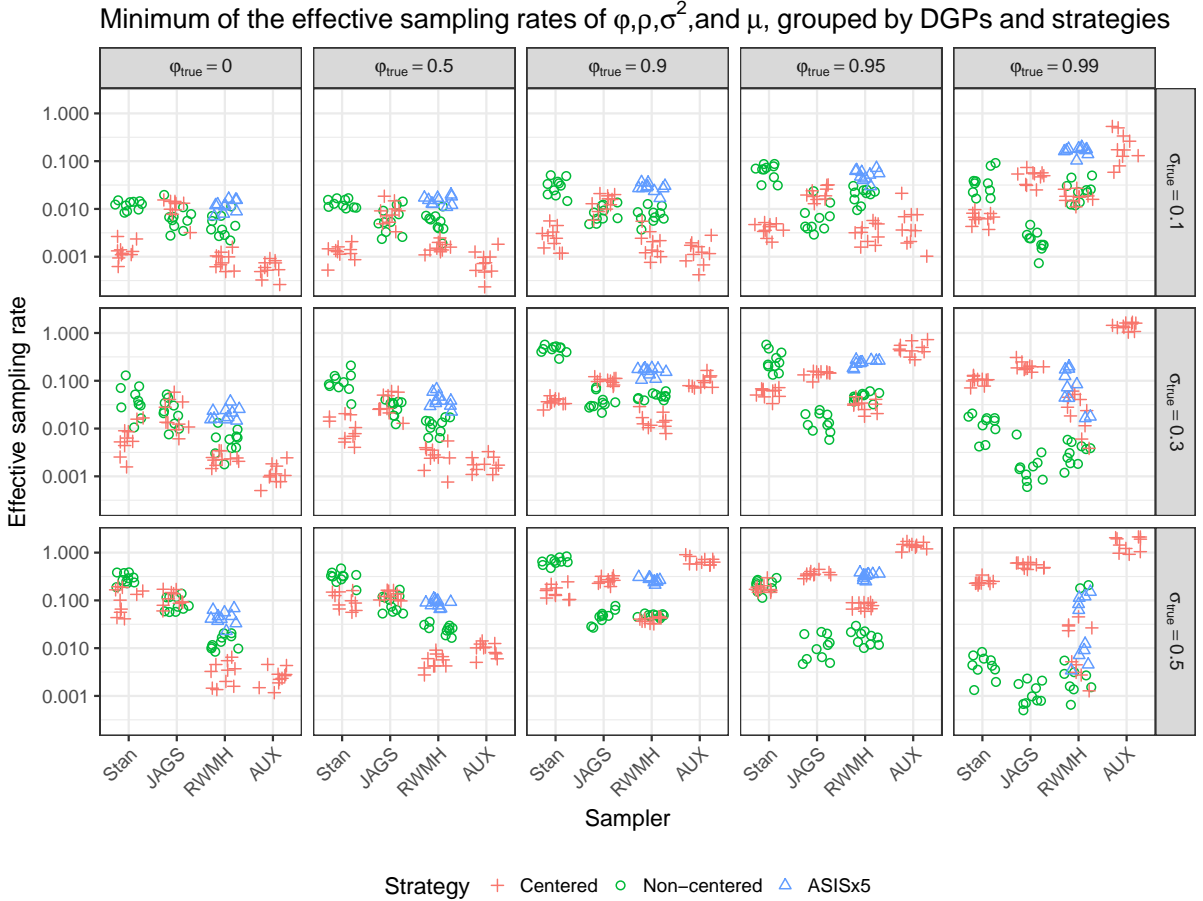


Figure 2: Minimal effective sampling rates of all the examined samplers and strategies, for the whole range of  $\varphi_{\text{true}}$  and  $\sigma_{\text{true}}$  values, while, for the sake of readability,  $\rho_{\text{true}}$  is set to  $-0.3$ , and  $T$  to 3000. In each facet, there are 10 data points plotted for each sampler and strategy, corresponding to the 10 repetitions of the simulation exercise. A small horizontal noise has been applied to the position of the points to increase their visibility.

Figure 3 summarizes the results of the exercise, carried out under the same prior specification as in Section 3. It is interesting to note that Stan generally shows high ESRs with the only exception of BTC where RWMH-ASISx5 excels. Focusing on the comparison of RWMH and AUX it stands out that without interweaving, AUX is generally to be preferred, whereas RWMH-ASISx5 tends to outperform AUX in all scenarios but one. The overall picture is similar to Figure 2, as there is no single algorithm that dominates on all data sets.

## 5 Conclusion

The paper at hand contributes to the literature on MCMC sampling algorithms by investigating the efficiency of several competing methods for the stochastic volatility model with leverage. We derived an RWMH sampler and improved it through ASIS and an efficient

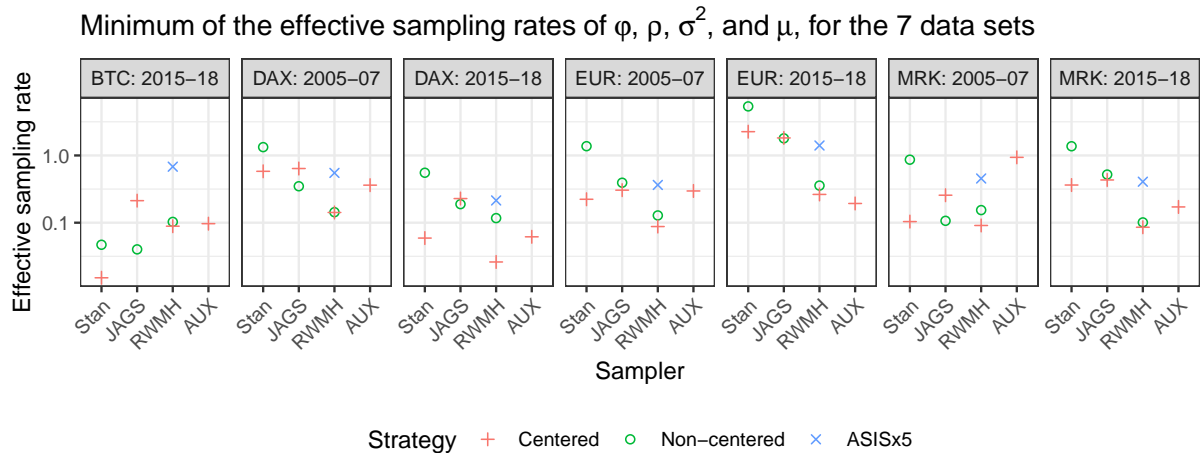


Figure 3: Minimal effective sampling rate for seven data sets. Facets correspond to time series, where “2005–07” and “2015–18” denote the first and the second time period, respectively. In each facet, each point corresponds to a certain sampler and strategy.

latent state sampler. Moreover, we carried out a computational experiment to compare our novel method to the state-of-the-art approach, an auxiliary mixture sampler, and to Stan and JAGS implementations as benchmarks. Based on our results, we conclude that employing our boosted naïve estimator for the latent space stabilizes the effective sampling rate of the algorithm by avoiding numerical optimization and differentiation.

Current research is directed towards further financial applications including factor models (Kastner et al., 2017), and extending the R package `stochvol` to allow for leverage.

## References

- Bob Carpenter, Andrew Gelman, Matthew Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1):1–32, 2017. doi:10.18637/jss.v076.i01.
- Chris K Carter and Robert Kohn. On Gibbs sampling for state space models. *Biometrika*, 81(3):541–553, 1994. doi:10.1093/biomet/81.3.541.
- Dirk Eddelbuettel and Romain François. Rcpp: Seamless R and C++ integration. *Journal of Statistical Software*, 40(8):1–18, 2011. doi:10.18637/jss.v040.i08.
- Sylvia Frühwirth-Schnatter. Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, 15(2):183–202, 1994. doi:10.1111/j.1467-9892.1994.tb00184.x.
- Sylvia Frühwirth-Schnatter and Helga Wagner. Stochastic model specification search for Gaussian and partial non-Gaussian state space models. *Journal of Econometrics*, 154(1):85–100, 2010. doi:10.1016/j.jeconom.2009.07.003.

- Andrew C. Harvey and Neil Shephard. Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business & Economic Statistics*, 14(4):429–434, 1996. doi:10.1080/07350015.1996.10524672.
- Eric Jacquier, Nicholas G. Polson, and Peter E. Rossi. Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics*, 122(1): 185–212, 2004. doi:10.1016/j.jeconom.2003.09.001.
- Gregor Kastner. Dealing with stochastic volatility in time series using the R package stochvol. *Journal of Statistical Software*, 69(5):1–30, 2016. doi:10.18637/jss.v069.i05.
- Gregor Kastner and Sylvia Frühwirth-Schnatter. Ancillarity-sufficiency interweaving strategy (ASIS) for boosting MCMC estimation of stochastic volatility models. *Computational Statistics and Data Analysis*, 76:408–423, 2014. doi:10.1016/j.csda.2013.01.002.
- Gregor Kastner, Sylvia Frühwirth-Schnatter, and Hedibert Freitas Lopes. Efficient Bayesian inference for multivariate factor stochastic volatility models. *Journal of Computational and Graphical Statistics*, 26(4):905–917, 2017. doi:10.1080/10618600.2017.1322091.
- Hans-Peter Kriegel, Erich Schubert, and Arthur Zimek. The (black) art of runtime evaluation: Are we comparing algorithms or implementations? *Knowledge and Information Systems*, 52:341–378, 2016. doi:10.1007/s10115-016-1004-2.
- Jun S. Liu. The collapsed Gibbs sampler in Bayesian computations with applications to a gene regulation problem. *Journal of the American Statistical Association*, 89(427): 958–966, 1994. doi:10.1080/01621459.1994.10476829.
- William J. McCausland, Shirley Miller, and Denis Pelletier. Simulation smoothing for state-space models: A computational efficiency analysis. *Computational Statistics and Data Analysis*, 55(1):199–212, 2011. doi:10.1016/j.csda.2010.07.009.
- Jouchi Nakajima and Yasuhiro Omori. Leverage, heavy-tails and correlated jumps in stochastic volatility models. *Computational Statistics and Data Analysis*, 53(6):2335–2353, 2009. doi:10.1016/j.csda.2008.03.015.
- Yasuhiro Omori, Siddhartha Chib, Neil Shephard, and Jouchi Nakajima. Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics*, 140(2):425–449, 2007. doi:10.1016/j.jeconom.2006.07.008.
- Martyn Plummer. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, volume 124, page 125, 2003.

- Martyn Plummer, Nicky Best, Kate Cowles, and Karen Vines. CODA: Convergence diagnosis and output analysis for MCMC. *R News*, 6(1):7–11, 2006.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, 2018. URL <https://www.R-project.org/>.
- Stephen J. Taylor. Financial returns modeled by the product of two stochastic processes: A study of daily sugar prices 1961-75. In *Time Series Analysis, Theory and Practice*, pages 203–226. North-Holland, 1982.
- Hadley Wickham. *tidyverse: Easily Install and Load 'tidyverse' Packages*, 2017. URL <https://CRAN.R-project.org/package=tidyverse>. R package version 1.1.1.
- Yaming Yu and Xiao-Li Meng. To center or not to center: That is not the question— An ancillarity-sufficiency interweaving strategy (ASIS) for boosting MCMC efficiency. *Journal of Computational and Graphical Statistics*, 20(3):531–570, 2011. doi:10.1198/jcgs.2011.203main.