

A particle-Kalman double filtering method for state space models with superposed latent states — Online appendix

Mengheng Li*

December 31, 2025

This online appendix accompanies: Li, Mengheng. (2026). “*A particle-Kalman double filtering method for state space models with superposed latent states*”. We provide some additional details of the simulation exercise, focusing on a case where the optimal proposal density is not available but approximated by the EIS method of [Scharth and Kohn \(2016\)](#) and [Li and Scharth \(2022\)](#). Also, some complementary estimation results of the empirical study in the main text are provided. This appendix provides only supporting materials and it is not meant for publication. Notations follow the main text, unless stated otherwise.

Keywords: trend inflation; sticky inflation; flexible inflation; stochastic volatility; dynamic factor model.

JEL Classification: C32; C53; E37.

*Senior Lecturer. University of Technology Sydney (UTS) Business School and Centre for Applied Macroeconomic Analysis (CAMA). Email: mengheng.li@uts.edu.au

1 Simulation studies with EIS proposal

In practical settings, the OPD is not available. Usually, such method involves computational overhead due to the necessary need to build an approximating proposal density. We consider the state-of-the-art EIS method originally developed by [Scharth and Kohn \(2016\)](#) and explored extensively by [Li and Scharth \(2022\)](#). In what follows, BS denotes the simple bootstrap filter. In the main text, it is denoted by BPF.

1.1 Statistical efficiency

Table [1](#) and [2](#) report the filtering accuracy results. Here, all model parameters are set equal to their true DGP values, ensuring that uncertainty arises solely from Monte Carlo variation in the particle filter. Relative to a PF implementation that does not make use of the superposition representation, PKDF with a basic BS scheme delivers substantial gains in estimating both the filtered mean and variance. The improvement scales with dimensionality, approximately threefold at $p = 2$, rising to nearly tenfold at $p = 8$. This illustrates the effectiveness of dimension reduction via superposition. When paired with EIS, PKDF still dominates PF, though the advantage is somewhat diminished. In line with earlier findings such as [Scharth and Kohn \(2016\)](#); [Li and Scharth \(2022\)](#), incorporating EIS helps alleviate particle degeneracy; across all p , the PKDF-filtered variance is estimated with near-perfect precision.

These patterns highlight a broader implication. PF combined with EIS remains less effective than PKDF with BS when the dimension grows. This reinforces the well-known deterioration in particle filtering performance caused by increasing state dimension and Monte Carlo noise, an instance of the curse of dimensionality documented by [Snyder et al. \(2008\)](#) and [Bengtsson et al. \(2008\)](#). Consequently, the benefits of methods like EIS may diminish rapidly without exploiting the superposition property. Further analysis of computational implications is given in Section [1.2](#).

Table [3](#) summarizes the RMSEs of parameter estimates over simulations using $M = M^* = 1,000$ and $p = 4$. Across all parameters, PKDF combined with BS achieves roughly a 50% reduction in RMSE relative to PF. Since mean estimates under both filters are nearly identical, this improvement largely reflects lower estimation variance under PKDF, a conclusion further supported by its visibly smaller dispersion in estimates.

One potential interpretation is that cross-replication variability in parameter estimates is tied

Table 1: RMSE OF PARTICLE ESTIMATE OF FILTERED MEAN

p	$E_t(z_{1,t})$		$E_t(z_{p,t})$	
	BS	EIS	BS	EIS
PKDF exploits superposition				
2	0.61 [0.5, 0.8]	0.18 [0.0, 0.4]	0.97 [0.5, 1.4]	0.33 [0.2, 0.5]
4	0.88 [0.6, 1.4]	0.22 [0.1, 0.4]	1.44 [0.9, 1.8]	0.58 [0.2, 1.0]
8	0.92 [0.6, 1.8]	0.30 [0.1, 0.5]	2.04 [1.5, 3.0]	0.90 [0.4, 1.6]
PF does not exploit superposition				
2	1.74 [1.1, 2.6]	0.72 [0.5, 0.9]	6.64 [3.7, 11]	2.36 [1.6, 3.7]
4	3.42 [2.2, 5.1]	1.08 [0.7, 1.4]	10.8 [6.1, 17]	4.82 [2.9, 7.4]
8	9.46 [6.1, 16]	2.48 [1.9, 3.0]	19.5 [11, 32]	6.20 [3.5, 10]

Reported are the mean, the 2nd and the 3rd quartiles (in square brackets) of RMSEs for the particle estimate of the filtered mean across 100 simulations obtained by PKDF and PF. $E_t(\cdot)$ denotes conditional expectation. $z_{1,t}$ and $z_{p,t}$ respectively denote the most and the least persistent state in the model with p latent states. The number of particles is set to be 200 in all cases. Model parameters are fixed at their DGP values. Resampling takes place whenever the ESS drops below 75% of the ensemble size.

to the variance of the likelihood estimate formed from terminal particles, as emphasized in [Chopin \(2004\)](#) and [Shephard \(2020\)](#). Because EIS aims to stabilize likelihood estimation, one might expect it to eliminate differences in variance between methods. Yet the bottom panel of Table 3 shows mixed results: PKDF reduces variance for most parameters, although AR coefficients show little discrepancy across filters.

Taken together, the consistent RMSE gains under PKDF indicate that leveraging superposition enhances the stability of particle likelihood evaluation, regardless of the propagation scheme employed.

Table 4 reports the heuristic particle sizes required to achieve a unit variance in the particle-based log-likelihood estimate, alongside the corresponding variances. Under the BS scheme, PF demands a substantially larger number of particles relative to PKDF, echoing the dimensionality constraints documented in [Bengtsson et al. \(2008\)](#) and [Snyder et al. \(2008\)](#). As p increases from 2 to 8, the particle requirement for PF grows roughly sevenfold and displays high variability across replications, evidence of rising Monte Carlo noise. In contrast, PKDF shows minimal fluctuation in required particle size, as only one superposed state must be integrated regardless of p .

Introducing EIS reduces the particle requirement for both filters; however, PF still exhibits

Table 2: RMSE OF PARTICLE ESTIMATE OF FILTERED VARIANCE

p	$\text{Var}_t(z_{1,t})$		$\text{Var}_t(z_{p,t})$	
	BS	EIS	BS	EIS
PKDF exploits superposition				
2	0.13 [0.1, 0.2]	0.00 [0.0, 0.0]	0.30 [0.1, 0.5]	0.10 [0.0, 0.1]
4	0.22 [0.1, 0.4]	0.01 [0.0, 0.0]	0.56 [0.2, 0.8]	0.12 [0.0, 0.2]
8	0.34 [0.2, 0.4]	0.04 [0.0, 0.1]	0.88 [0.5, 1.5]	0.14 [0.0, 0.2]
PF does not exploit superposition				
2	0.41 [0.3, 0.6]	0.11 [0.0, 0.2]	0.80 [0.5, 1.1]	0.53 [0.4, 0.7]
4	1.24 [0.8, 1.9]	0.26 [0.1, 0.4]	2.33 [1.6, 3.1]	0.76 [0.4, 1.0]
8	2.77 [2.0, 3.5]	0.35 [0.2, 0.5]	5.65 [3.4, 8.0]	1.40 [1.0, 1.8]

Reported are the mean, the 2nd and the 3rd quartiles (in square brackets) of RMSEs for the particle estimate of the filtered variance across 100 simulations obtained by PKDF and PF. $\text{Var}_t(\cdot)$ denote conditional variance. See notes under Table 1.

growth in ensemble size as p rises, whereas PKDF remains largely unaffected. With ensemble size fixed at 500, PF becomes inefficient, while EIS helps stabilize variance and enhance performance, most notably for PKDF. Increasing the size to 1000 further smooths variance but leaves the comparative patterns unchanged. These results suggest that global approximation tools like EIS can curb particle degeneracy, though often at notable computational cost, whereas PKDF paired with BS offers a simpler and more computationally attractive alternative that maintains strong accuracy in our experiments.

1.2 Computational efficiency

Figure 1 plots the effective sample size ESS_t from a representative simulation. Resampling is triggered whenever ESS_t falls below 75% of the particle population. By collapsing the state into a superposed representation, PKDF experiences much slower particle degeneracy than PF. When parameters are estimated and the simple BS is used, PKDF requires nearly twenty times fewer resampling steps than PF. A more sophisticated proposal, such as EIS, further moderates degeneracy, reinforcing the appeal of well-crafted proposal densities. Nonetheless, PKDF with BS already performs on par with PF equipped with EIS in terms of resampling frequency, suggesting that complexity is not always necessary for robust filtering.

Figure 2 compares the computation time of both filters with parameter estimation, using the

Table 3: RMSE OF PARTICLE ML ESTIMATES OF MODEL PARAMETERS

	ϕ_1	ϕ_4	σ_1	σ_4	σ_ϵ
Coupled with the bootstrap filter (BS)					
PKDF	0.084	0.040	0.025	0.103	0.142
PF	0.121	0.066	0.051	0.273	0.360
$\Delta \widehat{E}(\hat{\theta}_i)$	0.022 [0.42]	-0.007 [0.31]	-0.021 [0.18]	0.112 [0.03]	0.202 [0.45]
$\Delta \widehat{\text{Var}}(\hat{\theta}_i)$	-0.627 [0.00]	-0.126 [0.02]	-0.084 [0.06]	-0.306 [0.00]	-3.183 [0.01]
Coupled with efficient importance sampling (EIS)					
PKDF	0.085	0.063	0.018	0.211	0.0878
PF	0.092	0.066	0.070	0.245	0.132
$\Delta \widehat{E}(\hat{\theta}_i)$	-0.032 [0.13]	-0.043 [0.29]	0.003 [0.74]	0.055 [0.50]	-0.104 [0.22]
$\Delta \widehat{\text{Var}}(\hat{\theta}_i)$	-0.102 [0.42]	-0.083 [0.20]	-0.192 [0.03]	-0.511 [0.00]	-12.10 [0.00]

Reported are the RMSEs of a selection of model parameters estimated by particle ML estimation. Entries in the row with header $\Delta \widehat{E}(\hat{\theta}_i)$ ($\Delta \widehat{\text{Var}}(\hat{\theta}_i)$) show differences in the sample averages (variances) of particle ML estimates from PKDF and PF. Each difference is followed by the p -value of a paired two-sample t (F) test for equal mean (variance). All entries in the row of $\Delta \widehat{\text{Var}}(\hat{\theta}_i)$ are multiplied by 100.

Table 4: HEURISTIC ENSEMBLE SIZE AND VARIANCE OF LOG-LIKELIHOOD

BS			EIS		
$p = 2$	$p = 4$	$p = 8$	$p = 2$	$p = 4$	$p = 8$
Ensemble size M and M^* for a unit variance of $\log \hat{p}_\theta(y_{1:T})$ (divided by 1,000)					
1.5 [1.0, 2.2]	1.2 [0.9, 1.7]	2.5 [2.2, 2.8]	0.6 [0.5, 0.8]	0.7 [0.5, 0.9]	0.7 [0.5, 1.0]
3.6 [3.2, 4.2]	8.4 [7.6, 12]	17 [11, 27]	1.4 [1.1, 1.9]	5.9 [5.3, 6.6]	8.3 [7.3, 9.5]
Variance of $\log \hat{p}_\theta(y_{1:T})$ with $M = M^* = 500$					
2.6 [2.1, 4.4]	3.7 [3.0, 5.9]	11 [8.7, 14]	1.0 [0.7, 1.3]	1.2 [0.9, 1.6]	1.2 [0.8, 1.7]
7.6 [6.1, 9.5]	21 [15, 29]	45 [33, 82]	2.4 [2.1, 2.8]	11 [8.7, 14]	18 [12, 24]
Variance of $\log \hat{p}_\theta(y_{1:T})$ with $M = M^* = 1000$					
1.1 [0.9, 1.4]	1.0 [0.8, 1.5]	1.6 [1.3, 1.9]	0.7 [0.4, 1.1]	0.8 [0.4, 1.3]	0.8 [0.3, 1.5]
3.3 [2.4, 3.9]	15 [10, 21]	37 [22, 51]	1.1 [0.8, 1.6]	6.7 [5.4, 8.3]	26 [16, 40]

The top (bottom) row in each of the three panels indicates the results of PKDF (PF). Reported at are mean values across simulations, with the 2nd and the 3rd quartiles in brackets. M and M^* are the number of particles (ensemble size) of PF and PKDF, respectively.

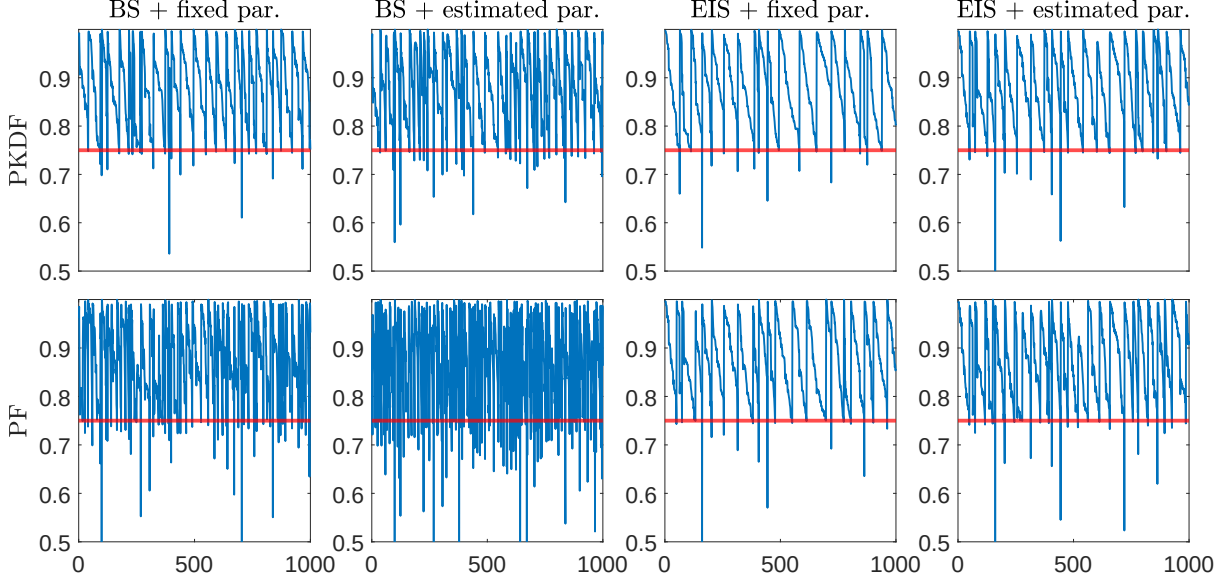


Figure 1: Effective sample size. Top and bottom panels show the effective sample size (in percentage point relative to the heuristic ensemble size) of PKDF and PF, respectively, from one simulation. Columns of figures show results under BS or EIS with fixed or estimated parameters, as indicated by the column names. The red line in each figure indicates 75%, the resampling threshold.

heuristic ensemble sizes. All simulation runs are parallelized externally, but within each run, further parallelization is restricted by the sequential nature of PF, the Kalman filter, and EIS.

The results in Figure 2 illustrate that treating the state jointly via superposition yields sizeable computational savings. As dimensionality increases, the runtime of PKDF combined with BS grows only modestly, whereas PF exhibits near exponential growth due to the escalating particle requirement needed to stabilize $\log \hat{p}_{\theta}(y_{1:T})$. Under EIS, computation times surge for both methods because the optimal proposal requires an intensive forward-filtering-backward-smoothing step, with the added cost scaling roughly linearly in p . Considering that PKDF with BS delivers comparable filtering quality to PF with EIS while being substantially faster and easier to implement, exploiting the superposition structure is advisable whenever the SSM accommodates it.

2 Estimation results of the MC-SV model in the empirical study

We apply our PKDF method to the MC-SV model as given in the main text. Estimations of static parameters and filtered states are done stock-wise and in parallel.

The estimated regression coefficients in Equation (13) in the main text replicate standard

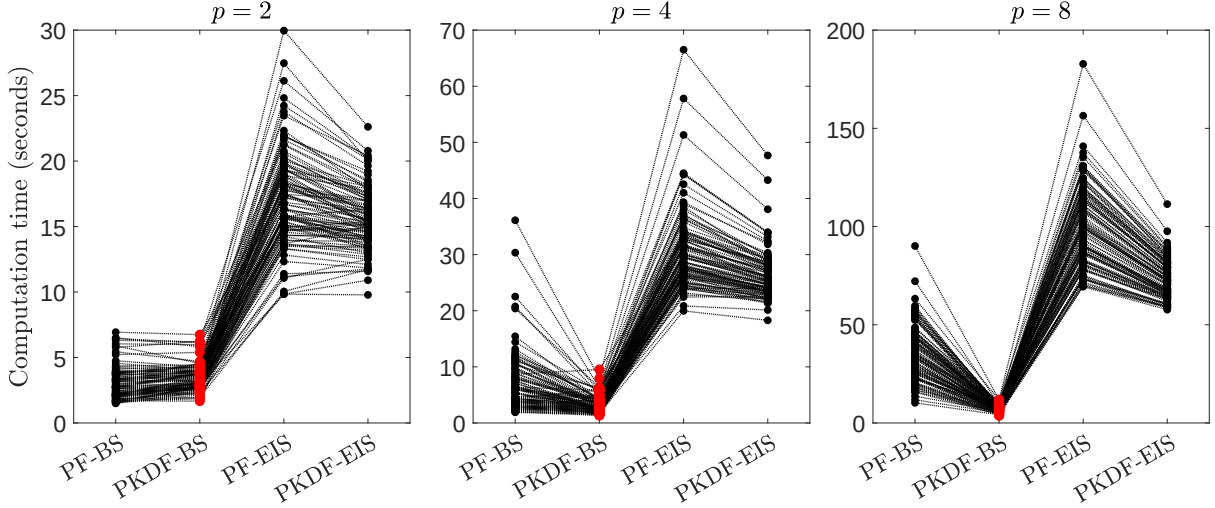


Figure 2: Computation time. Computation time includes the time needed for estimating parameters via the particle ML estimation of [Poyiadjis et al. \(2011\)](#) and for estimating the filtering density of individual states. For each p , 100 simulations are kept the same across methods.

findings in the literature and are not our focus. Instead, we examine the idiosyncratic component, corresponding to Step 1 in Section 2 in the main text. Figure 3 plots the estimated AR coefficients of each log volatility component against their respective log volatility-of-volatility coefficients on the ϕ_{ij} - $\log v_{ij}$ plane for $j = 1, \dots, 4$.

Across all stocks, $(\phi_{ij}, \log v_{ij})$ pairs form clear clusters for each volatility component. For instance, $h_{i1,t}$ exhibits high persistence across all i , whereas $h_{i4,t}$ is the least persistent, though still positive, for most stocks. A few stocks appear to have only three idiosyncratic volatility components, as their insignificant ϕ_{i4} values correspond to negligible volatility-of-volatility coefficients. Since all ϕ_{i4} values are close to zero, modeling idiosyncratic volatility as the superposition of four latent components strikes a balance between complexity and parsimony.

For models where an intractable likelihood function is approximated by the average of importance weights, [Geweke \(1989\)](#) show that estimator consistency and asymptotic normality require the importance weights to have a finite second moment. In our setting, this condition implies the existence of a finite variance for terminal particles or particle-estimated log-likelihood functions. To evaluate this, we apply the ML-based test of [Koopman et al. \(2009\)](#), which analyzes the asymptotic tail behavior of particle ensembles. Table 5 reports the results, including three tail index estimators—Pickands, Hill, and Dekkers-Einmahl-de Haan ([Embrechts et al., 2013](#))—which assess the highest finite moment of the distribution of

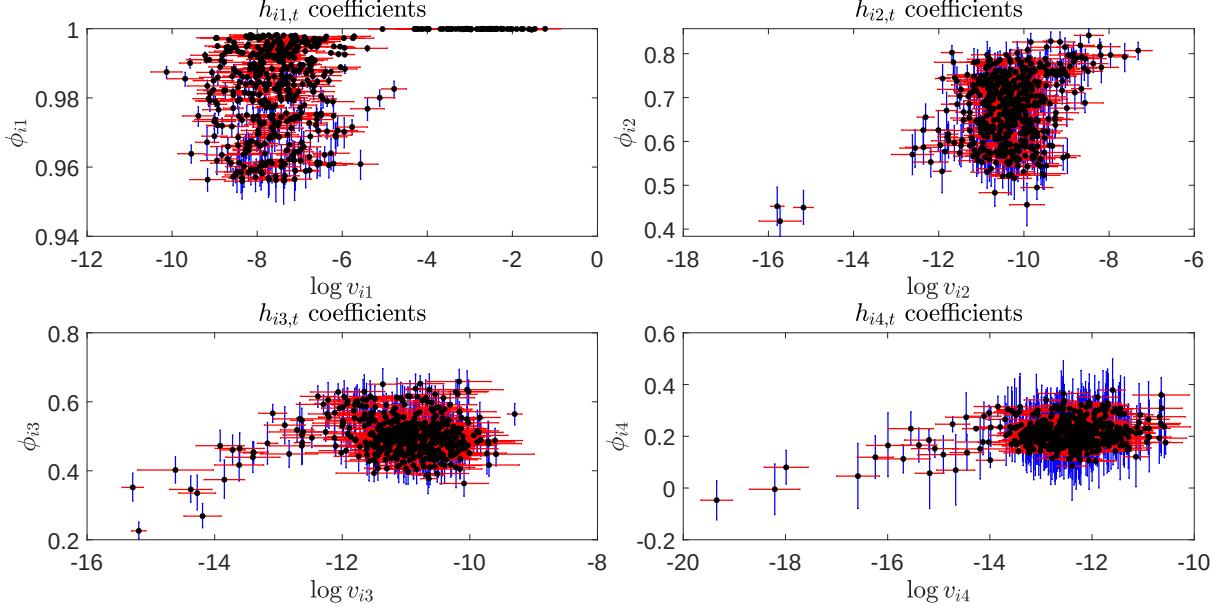


Figure 3: Parameter estimates for idiosyncratic volatility components. Each dot indicates a $(\phi_{ij}, \log v_{i,j})$ pair, $j = 1, \dots, 4$. The volatility-of-volatility coefficients are in logarithmic terms for readability. The associated blue and red bars correspond to the 95% confidence interval of ϕ_{ij} and $\log v_{ij}$, respectively. Standard errors are computed via the delta method.

maximized log-likelihood functions.

When PKDF is coupled with BS, the number of particles needed for valid inference is at least 1,500. Among the 489 idiosyncratic error series, there are still few that do not have a finite second moment even when the ensemble size is 1,000. This is also shown by the log-likelihood variance and tail indexes. Admittedly, one can employ well-designed particle filters, such as the EIS used in the previous section, to reduce the required ensemble size to few hundreds. However, as shown previously, EIS brings marginal improvement in terms of statistical efficiency when the main gain comes from the proposed PKDF method that exploits the superposition structure.

Based on Table 5 and our discussion in Section 1.1, we choose the number of particles to be 2,000, following the rule-of-thumb that the variance of log-likelihoods across all idiosyncratic errors is made approximately one.

Table 5: DIAGNOSTICS AND VALIDITY OF PARTICLE ML ESTIMATION

Ensemble size	500	1000	1500	2000	2500
$\text{Var}(\log \hat{p}_{\hat{\theta}}(\hat{\epsilon}_{i,1:T}))$	31 [11, 134]	8.6 [4.5, 38]	2.3 [1.4, 5.7]	1.2 [0.7, 1.9]	0.9 [0.5, 1.4]
Tail index estimators					
Pickands	2.5 [0.8, 4.5]	3.1 [2.0, 5.9]	4.5 [2.8, 7.7]	4.4 [2.8, 10.6]	4.6 [3.4, 13.7]
Hill	2.3 [0.7, 8.3]	4.5 [2.8, 8.1]	5.6 [2.4, 11.8]	4.8 [3.3, 7.9]	5.0 [3.1, 11.4]
DEH	2.3 [0.5, 6.6]	5.8 [2.4, 11.3]	4.6 [2.4, 8.8]	6.2 [3.8, 11.2]	5.4 [2.8, 14.7]
Finite variance test					
min. p -value	0.00	0.03	0.17	0.21	0.15

For the variance of log-likelihood functions and the three tail index estimators, we report the sample mean across $N = 489$ stock return idiosyncratic error series, followed by the minimum and the maximum (*i.e.* the range) shown in brackets. The bottom panel shows the smallest p -value among all series of the [Koopman et al. \(2009\)](#) test with the null hypothesis H_0 : the distribution has a finite variance.

References

- Bengtsson, T., P. Bickel, B. Li, et al. (2008). Curse-of-dimensionality revisited: Collapse of the particle filter in very large scale systems. *Probability and statistics: Essays in honor of David A. Freedman* 2, 316–334.
- Chopin, N. (2004). Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference. *Annals of Statistics* 32(6), 2385–2411.
- Embrechts, P., C. Klüppelberg, and T. Mikosch (2013). *Modelling extremal events: for insurance and finance*, Volume 33. Springer Science & Business Media.
- Geweke, J. (1989). Bayesian inference in econometric models using Monte Carlo integration. *Econometrica: Journal of the Econometric Society*, 1317–1339.
- Koopman, S. J., N. Shephard, and D. Creal (2009). Testing the assumptions behind importance sampling. *Journal of Econometrics* 149(1), 2–11.
- Li, M. and M. Scharth (2022). Leverage, asymmetry, and heavy tails in the high-dimensional factor stochastic volatility model. *Journal of Business & Economic Statistics* 40(1), 285–301.
- Poyiadjis, G., A. Doucet, and S. S. Singh (2011). Particle approximations of the score and observed information matrix in state space models with application to parameter estimation. *Biometrika* 98(1), 65–80.
- Scharth, M. and R. Kohn (2016). Particle efficient importance sampling. *Journal of Econometrics* 190(1), 133–147.
- Shephard, N. (2020). Statistical aspects of ARCH and stochastic volatility. In *Time series models*, pp. 1–67. Chapman and Hall/CRC.
- Snyder, C., T. Bengtsson, P. Bickel, and J. Anderson (2008). Obstacles to high-dimensional particle filtering. *Monthly Weather Review* 136(12), 4629–4640.