

Online Appendix for
Numerically Accelerated Importance Sampling for
Nonlinear Non-Gaussian State Space Models

Siem Jan Koopman^(a,b) André Lucas^(a,b,c) Marcel Scharth^(d)

^(a) VU University Amsterdam, The Netherlands

^(b) Tinbergen Institute, The Netherlands

^(c) Duisenberg School of Finance, The Netherlands

^(d) Australian School of Business, University of New South Wales

February 2014

A Numerical integration

When a continuous function $\varphi(x)$ is known analytically for any x , we can efficiently evaluate integrals of the form

$$\int \varphi(x) dx, \tag{1}$$

by numerical integration methods which are fast, reliable, and accurate. Numerical integration is not prone to simulation uncertainty and conforms to any desired degree of precision.

The numerical evaluation of the integral in (1) via a Gauss-Hermite quadrature designates a set of M abscissae z_j and associated weights $h(z_j)$ with $j = 1, \dots, M$. We compute the numerical approximation as

$$\int_{-\infty}^{\infty} \varphi(x) dx = \int_{-\infty}^{\infty} e^{-x^2} [e^{x^2} \varphi(x)] dx \approx \sum_{j=1}^M h(z_j) e^{z_j^2} \varphi(z_j), \tag{2}$$

where M is typically between 20 and 30, and where the weights $h(z_j)$ can be tabulated for different values of M . For a more detailed discussion on Gauss-Hermite quadrature, we refer to Monahan (2001). We refer to numerical integration in a general way in our discussions below but where possible we aim to work with Gaussian integrals for which analytical solutions may be available in specific applications. In such cases, the use of the analytical expression will bring further efficiency to our importance sampling method. Here we focus on the state space model of Section 2 with scalar signal θ_t , that is $q = 1$. Although all results are valid for a high dimensional vector θ_t , this setting brings additional and specific computational challenges that are beyond the scope of this paper. However, our treatment still allows for a high-dimensional state vector α_t .

B The NAIS, EIS and SPKD algorithms

Algorithms 1 and 2 outline the full procedure for likelihood estimation using the NAIS method. Algorithm 1 shows how we compute the likelihood estimate given the importance parameters (without control variables). Algorithm 2 reviews the method for obtaining the importance parameters. Algorithms 3 and 4 review the Shephard and

Pitt (1997) and Durbin and Koopman (1997) (SPDK) and EIS methods for obtaining the importance parameters as implemented in our Monte Carlo and empirical studies. Given the importance parameters, the two methods estimate the likelihood using Algorithm 1.

Algorithm 1 Importance sampling using an approximating linear state space model (with antithetic variables for variance reduction)

- ▷ Select the importance parameters b and C .
 - ▷ Obtain the smoothed mean $\tilde{\theta}_t$ of signal θ_t ($t = 1, \dots, n$) for the importance model using either the KFS or the backward-forward (BF) smoother.
 - ▷ Use a simulation smoothing method for linear state space models to sample $S/2$ trajectories $\theta^{(s)}$ for the signal. Two options are the simulation smoother of de Jong and Shephard (1995) and the BF method.
 - ▷ Compute the antithetic draws $\theta_t^{(s)} = 2\tilde{\theta}_t - \theta_t^{(s-S/2)}$, for $s = S/2 + 1 \dots, S$ and $t = 1, \dots, n$.
 - ▷ Compute the likelihood of the approximating linear state space model using the Kalman filter or the BF method of Appendix A.
 - ▷ Compute the likelihood estimate. The normalising constant a_t is given by (10) in the main text under the KFS based method and zero when using the BF method.
-

C Monte Carlo study: further results

C.1 Three other stochastic dynamic models

We first provide details of three additional stochastic dynamic models which are special cases of the general nonlinear and non-Gaussian state space model as introduced in Section 2 of the main paper. The model specifications will be used in our Monte Carlo and empirical studies below.

C.1.1 Stochastic conditional duration model

Bauwens and Veredas (2004) propose the stochastic conditional duration model for modelling durations between high-frequency financial transactions. Bauwens and Galli

Algorithm 2 Efficient importance parameters using NAIS

- ▷ Initialise the iteration index $k \leftarrow 0$.
- ▷ Set the initial values for the importance parameters $b^{[0]}, C^{[0]}$. We can do this using a local approximation technique.
- ▷ Choose M and obtain the Gauss-Hermite nodes z_j with associated weights $h(z_j)$ for $j = 1, \dots, M$.

while convergence criterion is not met **do**

▷ $k \leftarrow k + 1$

▷ Given the importance parameters $b^{[k-1]}$ and $C^{[k-1]}$, obtain the smoothed means $\tilde{\theta}_t$ and variances V_t of the signal θ_t ($t = 1, \dots, n$) under the approximating linear state space model. We can use either the KFS or the BF smoother for this purpose.

for $t=1:n$ **do**

▷ Generate $\tilde{\theta}_{tj} = \tilde{\theta}_t + V_t^{1/2} z_j$ for $j = 1, \dots, n$.

▷ Run a weighted least squares regression with

– Dependent variable: $\log p(y_t | \tilde{\theta}_{tj}; \psi)$.

– Explanatory vector: $(1, \tilde{\theta}_{tj}, -0.5 \tilde{\theta}_{tj}^2)'$.

– Weight: $\exp(\frac{1}{2} z_j^2) h(z_j)$ (fast version) or $\exp(\frac{1}{2} z_j^2) h(z_j) \omega(\tilde{\theta}_{tj}, y_t; \psi)$.

▷ The new importance parameters $b^{[k]}$ and $C^{[k]}$ are the coefficients associated with the second and third independent variables of the WLS regression.

end for

end while

▷ Set the efficient importance parameters as $b = b^{[k]}$ and $C = C^{[k]}$.

Algorithm 3 Importance parameters using SPDK

- ▷ Initialise the iteration index $k \leftarrow 0$.
- ▷ Set the initial value for the mode estimate $\widehat{\theta}_t^{[0]}$, $t = 1, \dots, n$.

while convergence criterion is not met **do**

▷ $k \leftarrow k + 1$

for $t=1:n$ **do**

- ▷ Obtain (12) in the main text as second order Taylor series approximation of the log observation density around the current mode estimate

$$b_t^{[k]} = \left. \frac{\partial \log p(y_t | \theta_t; \psi)}{\partial \theta_t} \right|_{\theta_t = \widehat{\theta}_t^{[k-1]}}, \quad C_t^{[k]} = - \left. \frac{\partial^2 \log p(y_t | \theta_t; \psi)}{\partial \theta_t \partial \theta_t'} \right|_{\theta_t = \widehat{\theta}_t^{[k-1]}}$$

end for

- ▷ Given the importance parameters $b^{[k]}$ and $C^{[k]}$, update the mode estimate $\widehat{\theta}_t^{[k]}$ by running the KFS under the approximating linear state space model.

end while

- ▷ Set the importance parameters as $b = b^{[k]}$ and $C = C^{[k]}$.
-

Algorithm 4 Efficient importance parameters using the standard EIS method

- ▷ Initialise the iteration index $k \leftarrow 0$.
- ▷ Set the initial values for the importance parameters $b^{[0]}, C^{[0]}$.
- ▷ Choose the number of simulations S and a draw set of $N(0, 1)$ common random numbers (CRN) $z_t^{(s)}$, $t = 1, \dots, n$ and $s = 1, \dots, S/2$. Set the antithetic draws as $z_t^{(s)} = -z_t^{(s-S/2)}$ for $s = S/2 + 1, \dots, S$.

while convergence criterion is not met **do**

▷ $k \leftarrow k + 1$

▷ Given the importance parameters $b^{[k-1]}$ and $C^{[k-1]}$ and the CRNs, use a simulation smoothing method for linear state space models to sample S trajectories $\theta^{(s)}$ for the the signal.

for $t=1:n$ **do**

▷ Run a weighted least squares regression with

- Dependent variable: $\log p(y_t | \theta_t^{(s)}; \psi)$.
- Explanatory vector: $(1, \theta_t^{(s)}, -0.5 \theta_t^{(s)2})'$.
- Weight: 1 (fast version) or $\omega(\theta_t^{(s)}, y_t; \psi)$.

▷ The new importance parameters $b^{[k]}$ and $C^{[k]}$ are the coefficients associated with the second and third independent variables of the WLS regression.

end for

end while

▷ Set the efficient importance parameters as $b = b^{[k]}$ and $C = C^{[k]}$.

(2009) consider the efficient importance sampling method for the maximum likelihood estimation of parameters. For an univariate time series of durations y_t , we consider the model

$$y_t \sim \text{Weibull}(\lambda_t, \xi), \quad \lambda_t = \exp(c + \alpha_t),$$

for $t = 1, \dots, n$, where λ_t is the time varying scale parameter and ξ is the shape parameter of the Weibull distribution. The scalar state variable α_t is specified as an autoregressive process of order 1. The set of true parameters is chosen to approximately reflect the estimation results of Bauwens and Galli (2009) for price durations. In terms of the state space model specification in Section 2, we have $c = 0$, $T_t = 0.98$, $Q_t = 0.15^2$ and $\xi = 1.2$.

C.1.2 Stochastic copula

Hafner and Manner (2011) formulate the stochastic copula model for estimating and forecasting time-varying and possibly nonlinear dependence between the time series. We consider a dynamic stochastic bivariate t -copula. Let u_{1t} and u_{2t} be two random variables with uniform $(0, 1)$ marginal distributions. In our simulation study, we take $u_t = (u_{1t}, u_{2t})'$ as probability integral transforms of two independent univariate series. We denote t_ν as the standardised Student's t distribution and 2×2 matrix P as the correlation matrix with unity values on the main diagonal and the correlation coefficient ρ on the two off-diagonal elements. The t -copula function $C_{\nu, P}(u_t)$ describes the dependence structure within the vector u_t and is given by

$$C_{\nu, P}(u_t) = T_{\nu, P} [t_\nu^{-1}(u_{1t}), t_\nu^{-1}(u_{2t})],$$

where $T_{\nu, P}(a, b)$ is the cumulative density function for the standardised bivariate Student's t distribution with degrees of freedom ν and correlation matrix P , for any set of variables $\{a, b\}$. The copula is invariant under any standardisation of the marginal distributions. It follows that

$$C_{\nu, P}(u_t) = \int_{-\infty}^{t_\nu^{-1}(u_{1t})} \int_{-\infty}^{t_\nu^{-1}(u_{2t})} \frac{\Gamma(\frac{\nu+2}{2})}{\Gamma(\nu/2)\sqrt{(\pi\nu)^2|P|}} \left(1 + \frac{x_t' P^{-1} x_t}{\nu}\right)^{-\frac{\nu+2}{2}} dx_{1t} dx_{2t}, \quad (3)$$

where $x_t = (x_{1t}, x_{2t})'$. A possible state space model for the stochastic copula with a time-varying correlation coefficient ρ_t , and hence a time-varying correlation matrix

$$P_t = \begin{bmatrix} 1 & \rho_t \\ \rho_t & 1 \end{bmatrix},$$

is given by

$$u_t \sim C_{\nu, P_t}(u_t), \quad \rho_t = 2[(1 + \exp(-c - \alpha_t))^{-1} - 0.5],$$

for $t = 1, \dots, n$, where the scalar state vector α_t is specified as an autoregressive process of order 1.

We take the set of true parameters from the empirical study of bivariate time series of financial log-returns in Hafner and Manner (2011). Hence the true parameter values in terms of the state space model in Section 2 are set to $c = 1$, $T_t = 0.98$ and $Q_t = 0.01$. We set d_t in the state space model such that the unconditional expectation of the correlation coefficient ρ_t is approximately 0.7. Since Hafner and Manner (2011) do not consider a t -copula, we take the degrees of freedom $\nu = 5$ to obtain sufficient tail dependence in our simulations.

C.1.3 Dynamic Factor Model for Multivariate Poisson Counts

The dynamic factor model for multivariate counts of Jung, Liesenfeld, and Richard (2011) is an example of a multivariate non-Gaussian state space model with possibly a large state vector but with a scalar signal function. Similar multivariate and discrete time series models are treated by Koopman and Lucas (2008). Here we show that we can implement the NAIS method can be implemented for this class of models as well, achieving substantial computational gains. Let y_t denote a vector of J counts. The observations are independently distributed conditional on the state vector α_t with Poisson density as given by

$$p(y_{j,t} | \lambda_{j,t}) = \frac{\exp(-\lambda_{j,t}) \lambda_{j,t}^{y_{j,t}}}{y_{j,t}!}, \quad j = 1, \dots, J, \quad t = 1, \dots, n,$$

where the intensity $\lambda_{j,t}$ is a nonlinear function of the $(J + 1) \times 1$ state vector $\alpha_t = (\alpha_{0,t}, \alpha_{1,t}, \dots, \alpha_{J,t})'$. In particular, the log of $\lambda_{j,t}$ is specified as the sum of a common

autoregressive factor $\alpha_{0,t}$ and an idiosyncratic component $\alpha_{j,t}$. The signal is given by

$$\theta_{j,t} = \log(\lambda_{j,t}) = c_j + \gamma_j \alpha_{0,t} + \alpha_{j,t}, \quad j = 1, \dots, J,$$

where c_j is the intercept and γ_j is the common factor loading. Each state element follows an autoregressive process of order 1 as given by

$$\alpha_{j,t+1} = \phi_j \alpha_{j,t} + \eta_{j,t}, \quad \eta_{j,t} \sim N(0, \sigma_{\eta,j}^2), \quad j = 0, 1, \dots, J,$$

where $|\phi_j| < 1$ is the autoregressive coefficient and $\sigma_{\eta,j}^2 > 0$ is the variance while the disturbances $\eta_{j,t}$ are mutually and serially uncorrelated, for $j = 0, 1, \dots, J$.

We implement the NAIS method for the multivariate model by formulating a state space representation as in Section 2 for the univariate time series

$$y_{1,1}, y_{2,1}, \dots, y_{J,1}, y_{1,2}, \dots, y_{J,n}.$$

The indices $j = 1, \dots, J$ and $t = 1, \dots, n$ follow the data sequence in corresponding order. Here we adopt the univariate approach of multivariate filtering and smoothing as discussed in Koopman and Durbin (2000). It is based on the univariate state space model with scalar signal $\theta_{j,t} = c_j + Z_{j,t} \alpha_t$, $j = 1, \dots, J$, and the two updating equations $\alpha_{j+1,t} = \alpha_{j,t}$, $j = 1, \dots, J - 1$, and $\alpha_{1,t+1} = T_{J,t} \alpha_{J,t} + \eta_{J,t}$ with $\eta_{J,t} \sim N(0, Q_{J,t})$, for $t = 1, \dots, n$. The associating time-varying (or index-varying) system matrices are given by

$$Z_{J,t} = \begin{pmatrix} \gamma_J \\ \iota_J \end{pmatrix}', \quad T_{J,t} = \begin{bmatrix} \phi_0 & 0 & 0 & 0 \\ 0 & \phi_1 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \phi_J \end{bmatrix}, \quad Q_{J,t} = \begin{bmatrix} \sigma_{\eta,0}^2 & 0 & 0 & 0 \\ 0 & \sigma_{\eta,1}^2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \sigma_{\eta,J}^2 \end{bmatrix},$$

for each t , and

$$Z_{j,t} = \begin{pmatrix} \gamma_j \\ \iota_j \end{pmatrix}', \quad T_{j,t} = I_{J+1}, \quad Q_{j,t} = 0,$$

for each t and where ι_j is the j th column of the $J \times J$ identity matrix I_J for $j = 1, \dots, J$.

In the Monte Carlo study we simulate multiple time series of counts with $J = 5$. The

parameter values are set to those considered by Jung, Liesenfeld, and Richard (2011). The coefficients for the dynamic processes are given by $\phi_0 = \phi_j = 0.98$, $\sigma_{\eta,0} = 0.1$ and $\sigma_{\eta,j} = 0.05$ for $j = 1, \dots, J$. All constants are set to $c_j = 1.5$. The loading coefficients are given by $\gamma_1 = 1$, $\gamma_2 = 1.2$, $\gamma_3 = 0.8$, $\gamma_4 = 1.5$ and $\gamma_5 = 0.7$.

C.2 Likelihood estimation and design of Monte Carlo study

We examine the performances of log-likelihood estimation via the importance sampling methods that are listed in Table 1 of the main paper. The design of the Monte Carlo study is as follows. We consider 500 random time series of the state space models that are discussed above. The choice of 500 simulations is taken to avoid the dependence of our conclusions on particular trajectories of the simulated states and series. For each simulated time series, we estimate the log-likelihood function at the true parameters a hundred times using different common random numbers. Hence each cell in the presented tables below reflects 50,000 likelihood evaluations. We report the results for the different sample sizes $n = 1,000$ and $n = 5,000$ while we use $S = 200$ importance samples for each likelihood evaluation.

The reported statistics are computed by

$$\text{Bias} = 50000^{-1} \cdot \sum_{i=1}^{500} \sum_{j=1}^{100} \left(\log \widehat{L}^j(y^i; \psi) - \log L(y^i; \psi) \right),$$

$$\text{Variance} = 500^{-1} \cdot \sum_{i=1}^{500} 100^{-1} \cdot \sum_{j=1}^{100} \left(\log \widehat{L}^j(y^i; \psi) - \log \bar{L}(y^i; \psi) \right)^2,$$

where y^i is the i th simulated time series, $\log \widehat{L}^j(y^i; \psi)$ is the “true” log-likelihood value, $\log \widehat{L}^j(y^i; \psi)$ is the j th estimate of the log-likelihood function for a particular method and $\log \bar{L}(y^i; \psi) = 100^{-1} \sum_{j=1}^{100} \log L^j(y^i; \psi)$. The true log-likelihood value is unknown and we approximate it by taking the log of the average of likelihood estimates from the NAIS method. It is based on the NAIS likelihood estimate from $S = 200 \times 100 = 20,000$ importance samples. Hence the approximation error of the true likelihood is negligible. We compute the mean squared error (MSE) as the sum of the variance and the square of the bias estimate. The variance and the MSE are reported as a ratios with respect to the standard EIS method. Our key summary statistic is the time normalised variance

ratio of method a against the benchmark method b and it is given by

$$\text{Variance}_{a/b} \times \left(1 + \frac{\text{Time}_{\text{I+II}}^a - \text{Time}_{\text{I+II}}^b}{\text{Time}_{\text{I}}^a} \right)^{-1}$$

where $\text{Variance}_{a/b}$ is the ratio of the Monte Carlo variance for methods a and b and Time_j^m is the time length of task j , for $j = \text{I}, \text{II}, \text{I+II}$, by method m , for $m = a, b$. We have excluded Time_{I}^a from the denominator because it is a fixed cost and not relevant for drawing additional samples.

The number of nodes for the numerical integration calculations is $M = 20$. We base the initialisation of each importance sampling method on the local approximation of SPDK. To reduce the simulation variance for all likelihood evaluations, we use antithetic variables for location as in Durbin and Koopman (2000), except for the NAISc methods. In our study, for all models, we have found no evidence of importance sampling weights that constitute an infinite variance; see the discussions in Koopman, Shephard, and Creal (2009). Our diagnostic procedure includes the verification of how sensitive the importance sampling weights are to artificial outliers as in Richard and Zhang (2007). We have efficiently implemented all methods using MATLAB and C code. The computer code is available from the authors upon request.

C.3 Monte Carlo results for models with a scalar state

Tables 1 and 2 present the results for two dynamic stochastic model with a scalar state: stochastic conditional duration and stochastic copula models. The findings are similar across the stochastic volatility model of the main paper and these models.

C.4 Monte Carlo results for models with multiple states

Table 3 reports the findings for the dynamic factor model for multivariate Poisson counts. We have implemented the EIS-BF algorithm as proposed in Jung, Liesenfeld, and Richard (2011). Given that the computer code is more involved for multivariate models, we have not implemented the NAIS-BF and NAIS-BF-Ctrl method. The presented results are similar to the results that are reported for the SV model with $k = 2$.

Table 1: LOG-LIKELIHOOD ERRORS FOR THE STOCHASTIC CONDITIONAL DURATION MODEL.

The table presents the numerical and computational performance of different IS methods for log-likelihood estimation. We simulate 500 different realisations from the model. For each of these realisations, we obtain log-likelihood estimates for 100 different sets of random numbers. We estimate the variance associated with each method as the average sample variance across the 500 realisations. We define the mean-square error (MSE) as the sum of the variance and the square of the average bias across the 500 realisations. We show these statistics as ratios with the standard implementation of the EIS method as the benchmark. The time for step 1 column gives the fixed time cost for obtaining the parameters of the importance density, while the time for step 2 refers to the computational cost of sampling from the importance density and calculating the likelihood estimate. The TNVAR column reports the time normalised variance ratio. We list the methods (with their acronyms) in Table 1 of the main paper. We specify the stochastic conditional duration model as: $y_t \sim \text{Weibull}(\lambda_t, \psi = 1.2)$ with $\lambda_t = \exp(\alpha_t)$ and $\alpha_{t+1} = 0.98\alpha_t + \eta_t$ where $\eta_t \sim N(0, \sigma_\eta^2 = 0.0225)$ for $t = 1, \dots, n$.

				Time Step 1	Time Step 2	
$n = 1000, S = 200$	Variance	MSE	($\times 10$)	($\times 10$)	TNVAR	
SPDK	14.888	15.034	0.052	0.212	7.039	
EIS-BF	1.000	1.000	0.278	0.221	1.000	
EIS-JSDK	1.008	1.008	0.284	0.210	0.980	
NAIS-BF	0.562	0.561	0.114	0.207	0.302	
NAIS-JSDK	0.563	0.562	0.110	0.209	0.302	
NAIS-BF-Ctrl	0.378	0.379	0.112	0.222	0.216	
NAIS-JSDK-Ctrl	0.379	0.380	0.114	0.214	0.210	

				Time Step 1	Time Step 2	
$n = 5000, S = 200$	Variance	MSE	($\times 10$)	($\times 10$)	TNVAR	
SPDK	12.980	21.861	0.200	1.228	5.515	
EIS-BF	1.000	1.000	1.766	1.325	1.000	
EIS-JSDK	1.012	1.011	1.762	1.208	0.919	
NAIS-BF	0.485	0.479	0.644	1.298	0.257	
NAIS-JSDK	0.499	0.492	0.629	1.243	0.252	
NAIS-BF-Ctrl	0.401	0.398	0.716	1.339	0.226	
NAIS-JSDK-Ctrl	0.405	0.402	0.694	1.257	0.212	

Table 2: LOG-LIKELIHOOD ERRORS FOR THE STOCHASTIC COPULA MODEL.

The table presents the numerical and computational performance of different IS methods for log-likelihood estimation. We simulate 500 different realisations from the model. For each of these realisations, we obtain log-likelihood estimates for 100 different sets of random numbers. We estimate the variance associated with each method as the average sample variance across the 500 realisations. We define the mean-square error (MSE) as the sum of the variance and the square of the average bias across the 500 realisations. We show these statistics as ratios with the standard implementation of the EIS method as the benchmark. The time for step 1 column gives the fixed time cost for obtaining the parameters of the importance density, while the time for step 2 refers to the computational cost of sampling from the importance density and calculating the likelihood estimate. The TNVAR column reports the time normalised variance ratio. We list the methods (with their acronyms) in Table 1 of the main paper. We specify the stochastic copula model as: $u_{1t}, u_{2t} \sim C_{\nu=5, P_t}(u_t)$, $\rho_t = (1 - \exp(-\alpha_t))(1 + \exp(-\alpha_t))^{-1}$, $\alpha_{t+1} = 0.98\alpha_t + \eta_t$, where $\eta_t \sim N(0, \sigma_\eta^2 = 0.01)$.

$n = 1000, S = 200$	Variance	MSE	Time Step 1	Time Step 2	TNVAR
			($\times 10$)	($\times 10$)	
SPDK	10.480	10.484	0.027	0.397	6.485
EIS-BF	1.000	1.000	0.257	0.411	1.000
EIS-JSDK	0.989	0.989	0.285	0.397	1.025
NAIS-BF	0.622	0.622	0.146	0.401	0.477
NAIS-JSDK	0.612	0.612	0.143	0.401	0.467
NAIS-BF-Ctrl	0.385	0.386	0.151	0.419	0.312
NAIS-JSDK-Ctrl	0.388	0.388	0.155	0.408	0.308

$n = 5000, S = 200$	Variance	MSE	Time Step 1	Time Step 2	TNVAR
			($\times 10$)	($\times 10$)	
SPDK	16.229	16.294	0.110	2.024	9.390
EIS-BF	1.000	1.000	1.516	2.093	1.000
EIS-JSDK	1.001	1.001	1.581	2.046	1.010
NAIS-BF	0.543	0.543	0.856	2.054	0.405
NAIS-JSDK	0.549	0.548	0.808	2.043	0.400
NAIS-BF-Ctrl	0.329	0.330	0.885	2.132	0.258
NAIS-JSDK-Ctrl	0.328	0.329	0.869	2.061	0.247

Table 3: LOG-LIKELIHOOD ERRORS FOR THE DYNAMIC FACTOR MODEL FOR MULTIVARIATE POISSON COUNTS.

The table presents the numerical and computational performance of different IS methods for log-likelihood estimation. We simulate 500 different realisations from the model. For each of these realisations, we obtain log-likelihood estimates for 100 different sets of random numbers. We estimate the variance associated with each method as the average sample variance across the 500 realisations. We define the mean-square error (MSE) as the sum of the variance and the square of the average bias across the 500 realisations. We show these statistics as ratios with the standard implementation of the EIS method as the benchmark. The time for step 1 column gives the fixed time cost for obtaining the parameters of the importance density (in seconds), while the time for step 2 refers to the computational cost of sampling from the importance density and calculating the likelihood estimate. The TNVAR column reports the time normalised variance ratio. The last column reports the normalised variance when we ignore the fixed time cost of step 1. We list the methods (with their acronyms) in Table 1 of the main paper.

$n = 1000, S = 200$	Variance	MSE	Time	Time	TNVAR
			Step 1	Step 2	
SPDK	22.268	22.653	0.065	0.118	3.056
EIS-BF	1.000	1.000	0.283	0.195	1.00
EIS-JSDK	0.975	0.975	0.235	0.116	0.182
NAIS-JSDK	0.496	0.496	0.099	0.118	0.072
NAIS-JSDK-Ctrl	0.302	0.302	0.103	0.161	0.061

$n = 5000, S = 200$	Variance	MSE	Time	Time	TNVAR
			Step 1	Step 2	
SPDK	16.612	34.246	0.385	0.587	1.915
EIS-BF	1.000	1.000	1.616	1.088	1.000
EIS-JSDK	1.014	1.012	1.210	0.592	0.166
NAIS-JSDK	0.448	0.441	0.755	0.612	0.053
NAIS-JSDK-Ctrl	0.361	0.358	0.757	0.831	0.061

References

- Bauwens, L. and F. Galli (2009). Efficient importance sampling for ml estimation of scd models. *Computational Statistics and Data Analysis* 53(6), 1974–1992.
- Bauwens, L. and D. Veredas (2004). The stochastic conditional duration model: A latent factor model for the analysis of financial durations. *Journal of Econometrics* 119(2), 381–412.
- de Jong, P. and N. Shephard (1995). The simulation smoother for time series models. *Biometrika* 82, 339–350.
- Durbin, J. and S. J. Koopman (1997). Monte Carlo maximum likelihood estimation for non-Gaussian state space models. *Biometrika* (84), 669–684.
- Durbin, J. and S. J. Koopman (2000). Time series analysis of non-Gaussian observations based on state space models from both classical and Bayesian perspectives. *Journal of the Royal Statistical Society, Series B* (62), 3–56.
- Hafner, C. and H. Manner (2011). Dynamic stochastic copula models: Estimation, inference and applications. *Journal of Applied Econometrics*. forthcoming.
- Jung, R., R. Liesenfeld, and J. F. Richard (2011). Dynamic Factor Models for Multivariate Count Data: An Application to Stock-Market Trading Activity. *Journal of Business and Economic Statistics* 29(1), 73–85.
- Koopman, S. J. and J. Durbin (2000). Fast filtering and smoothing for multivariate state space models. *Journal of Time Series Analysis* 21(3), 281–296.
- Koopman, S. J. and A. Lucas (2008). A non-Gaussian panel time series model for estimating and decomposing default risk. *Journal of Business and Economic Statistics* 26(4), 510–525.
- Koopman, S. J., N. Shephard, and D. Creal (2009). Testing the assumptions behind importance sampling. *Journal of Econometrics* 149(1), 2–11.
- Monahan, J. F. (2001). *Numerical methods of statistics*. Cambridge: Cambridge University Press.
- Richard, J. F. and W. Zhang (2007). Efficient high-dimensional importance sampling. *Journal of Econometrics* 141, 1385–1411.

Shephard, N. and M. Pitt (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika* 84, 653–667.