Identification of Time-varying Parameters using Variational Bayes - Sequensible Monte Carlo Sampler

Adolphus Lye, Ander Gray

Institute for Risk and Uncertainty, University of Liverpool, United Kingdom.
E-mail: {adolphus.lye, ander.gray}@liverpool.ac.uk

Edoardo Patelli

Centre for Intelligent Infrastructure, University of Strathclyde, United Kingdom.
Institute for Risk and Uncertainty, University of Liverpool, United Kingdom.
E-mail: edoardo.patelli@strath.ac.uk

This work presents an extended sequential Monte Carlo sampling algorithm embedded with a Variational Bayes step. The algorithm is applied to estimate the distribution of time-varying parameters in a Bayesian filtering procedure. This algorithm seeks to address the case whereby the state-evolution model does not have an inverse function. In the proposed approach, a Gaussian mixture model is adopted whose covariance matrix is determined via principle component analysis.

As a form of verification, a numerical example involving the identification of inter-storey stiffness within a 2-DOF shear building model is presented whereby the stiffness parameters degrade according to a simple State-evolution model whose inverse function can be derived. The Variational Bayes - sequential ensemble Monte Carlo sampler is implemented alongside the Sequential Monte Carlo sampler and the results compared on the basis of the accuracy and precision of the estimates as well computational time. A non-linear time-series model whose state-evolution model does not yield an inverse function is also analysed to show the applicability of the proposed approach.

Keywords: Variational Bayes, Bayesian Model Updating, Sequential Monte Carlo, Uncertainty Quantification, Gaussian Mixture Model, Markov Model.

1. Introduction

In recent years, on-line Bayesian inference has garnered significant interest to perform real-time identification of epistemic structural parameters in the field of engineering, especially so for structural health monitoring (Chang et al. (2003), Rocchetta et al. (2018)). A popular tool adapted for such purpose is the Sequential Monte Carlo (SMC) sampler. The SMC sampler is endowed with 2 key benefits over other sampling techniques such as Markov-Chain Monte Carlo (MCMC) and Transitional Markov-Chain Monte Carlo (TM-CMC): 1) the SMC sampler is able to sample from time-varying posteriors; and 2) it can infer time-varying model parameters (Doucet et al. (2011)).

Current SMC algorithms face the following limitations: 1) it is computationally inefficient; 2) requires knowledge of the inverse Markov model to compute the predictive PDF which reflects the prediction of the time-varying parameter in the next time-step prior to observing the next data; and 3) it does not have any mechanism to control the sample acceptance-rates within acceptable bounds (Roberts and Rosenthal (2001)). To address these limitations, this paper proposes a new sampler named Variational Bayes-Sequential Ensemble Monte Carlo (VB-SEMC). The sampler presents 3 key features: 1) the use of the Affine-invariant Ensemble Sampler (AIES) in place of the Metropolis-Hastings (MH) algorithm; 2) a Variational Bayes step to approximate the required distribution via the use of a Gaussian Mixture Model (Ormerod and Wand (2012)); and 3) adopt an adaptive step-size tuning algorithm to moderate the acceptance-rate of the sampler automatically. The proposed VB-SEMC sampler seeks to be: 1) computationally efficient; and 2) provide reasonable approximations to the predictive PDF.

To highlight the advantages and strengths of the proposed VB-SEMC, the sampler is compared against the traditional SMC sampler by Chopin (2002) for structural parameter identification. In particular, a 2-DOF shear building model is analysed where the objective is to identify the time-invariant floor masses and the time-varying inter-storey stiffness parameters. Following which, the VB-SEMC sampler is adopted for solving a benchmark filtering problem by Cappe et al. (2007) in the form of a non-linear time-series set-up from where the Markov model has no closed-form inverse function.

2   *Adolphus Lye, Ander Gray, and Edoardo Patelli*

1.1. *On-line Bayesian Inference*

Bayesian inference is a statistical inferential technique whose mathematical formulation is based on the Bayes' rule (Beck and Katafygiotis (1998)):

$$P(\boldsymbol{\theta}|\boldsymbol{D}) = \frac{P(\boldsymbol{D}|\boldsymbol{\theta}) \cdot P(\boldsymbol{\theta})}{P(\boldsymbol{D})} \qquad (1)$$

for which $\boldsymbol{\theta}$ is the vector of model parameters to be inferred, while $\boldsymbol{D}$ is the vector of observations obtained from measurements or experiments. The main terms in Eq. (1) are:

- $P(\boldsymbol{\theta})$ is the prior distribution;
- $P(\boldsymbol{D}|\boldsymbol{\theta})$ is the likelihood function;
- $P(\boldsymbol{\theta}|\boldsymbol{D})$ is the posterior distribution;
- $P(\boldsymbol{D})$ is the normalizing constant.

Detailed descriptions to the above terms can be found in Lye et al. (2019) and Lye et al. (2021).

For on-line Bayesian inference, a Markov model $\tilde{M}(\boldsymbol{\theta}^{t+1}|\boldsymbol{\theta}^t)$ is used to model the time-evolution of $\boldsymbol{\theta}$ between successive time-steps t and $t + 1$. Hence, the Markov model relating $\boldsymbol{\theta}^t$ and $\boldsymbol{\theta}^{t+1}$ can be expressed as:

$$\boldsymbol{\theta}^{t+1} = \Gamma(\boldsymbol{\theta}^t) + \nu_{\boldsymbol{\theta}} \qquad (2)$$

where $\Gamma(\bullet)$ is the nominal evolution function, and $\nu_{\boldsymbol{\theta}} \sim N(0, \sigma_{\nu})$ is the process "noise" with σ_{ν} assumed to be a constant. Defining $\boldsymbol{D}^{1:t} = \{\boldsymbol{D^1}, \dots, \boldsymbol{D^t}\}$, and assuming independence between observations obtained between difference time-steps t, the on-line Bayesian procedure is summarized as such: At time-step $t = 1$, the posterior $P(\boldsymbol{\theta}^t|\boldsymbol{D}^{1:t}, M)$ is defined (see Eq. (1)). Following which, the predictive distribution $P(\boldsymbol{\theta}^{t+1}|\boldsymbol{D}^{1:t}, M)$ is computed:

$$P(\boldsymbol{\theta}^{t+1}|\boldsymbol{D}^{1:t}, M) = \int \tilde{M}(\boldsymbol{\theta}^{t+1}|\boldsymbol{\theta}^t) \cdot$$
$$P(\boldsymbol{\theta}^t|\boldsymbol{D}^{1:t}, M) \cdot d\boldsymbol{\theta}^t \quad (3)$$

The predictive distribution $P(\boldsymbol{\theta}^{t+1}|\boldsymbol{D}^{1:t}, M)$ is then set as the new prior to be updated and the process is repeated for time-step $t = t + 1$ until the final time-step $t = t_{end}$. A useful sampling technique for such set-up would be the SMC sampler whose details can be found in Chopin (2002).

2. Proposed sampler: VB-SEMC

The workings of the VB-SEMC sampler to generate N samples from $P(\boldsymbol{\theta}^t|\boldsymbol{D}^{1:t}, M)$ follows: At iteration $j = 0$, the sampler is initiated by generating samples $\boldsymbol{\theta}_i^{j+1}$ (for $i = 1, \dots, N$). Next,

at iteration $j = 1$, the normalized weights of the samples $\hat{w}_i^j$ are computed:

$$\hat{w}_i^j = \frac{P(\boldsymbol{D}^j|\boldsymbol{\theta}_i^j)}{\sum_{i=1}^N P(\boldsymbol{D}^j|\boldsymbol{\theta}_i^j)} \qquad (4)$$

From which the Effective sample size N_{eff} is computed:

$$N_{eff} = \frac{1}{\sum_{i=1}^N \left(\hat{w}_i^j\right)^2} \qquad (5)$$

If $N_{eff} > \frac{N}{2}$, the algorithm proceeds to the MCMC step. Otherwise, resampling is executed where the samples $\boldsymbol{\theta}_i^t$ are sampled according to their weights $\hat{w}_i^j$ and then resets $\hat{w}_i^j = \frac{1}{N}$. After which, the algorithm proceeds to the MCMC step.

In the MCMC step, N single-step Markov Chains are initiated from starting samples $\boldsymbol{\theta}_i^j$ which are randomly sampled according to $\hat{w}_i^j$. These N starting samples form the starting ensemble $\vec{\boldsymbol{\theta}}_{start} = \{\boldsymbol{\theta}_1^j, \dots, \boldsymbol{\theta}_N^j\}$ which will be updated with a 1 iteration of the AIES sampler to yield the updated ensemble $\vec{\boldsymbol{\theta}}_{updated}$ whose samples correspond to the Posterior $P(\boldsymbol{\theta}^j|\boldsymbol{D}^{1:j}, M)$. Details to the AIES algorithm can be found in Goodman and Weare (2010).

Next, the algorithm proceeds to generate $\boldsymbol{\theta}_i^{j+1}$ by passing the samples of $P(\boldsymbol{\theta}^j|\boldsymbol{D}^{1:j}, M)$ through the $\tilde{M}(\boldsymbol{\theta}^{t+1}|\boldsymbol{\theta}^t)$. Simultaneously, the algorithm executes the VB step to approximate the predictive PDF. This is done using a Gaussian Mixture Model whose covariance matrix Σ_{GMM} is computed via the following steps (Patelli et al. (2014)):

- Step 1: The covariance matrix of $\boldsymbol{\theta}_i^{j+1}$ is computed and defined as $\boldsymbol{\Sigma}_{pred}$;
- Step 2: Spectral decomposition is performed on $\boldsymbol{\Sigma}_{pred}$ to obtain the matrix of eigen-vectors $\boldsymbol{R}$;
- Step 3: Transform the sample vector $\boldsymbol{\theta}^{j+1}$ in the basis of $\boldsymbol{\Sigma}_{pred}$ following: $\boldsymbol{\Theta}^{j+1} = \boldsymbol{\theta}^{j+1}\boldsymbol{R}$;
- Step 4: For each column entry of $\boldsymbol{R}$ (i.e. $\boldsymbol{R}(:,k)$), obtain a bandwidth estimate v_k (for $k = 1, \dots, N_{dim}$, where N_{dim} is the dimensionality of $\boldsymbol{\theta}^{j+1}$) via MATLAB $ksdensity$ function;
- Step 5: Define an empirical scaling factor: $\Omega = 1 + 3 \cdot (1 - 0.05^{\frac{1}{N}})$, and compute the column entry of matrix $\boldsymbol{P}$: $\boldsymbol{P}(:,k) = \boldsymbol{R}(:,k) \cdot v_k \cdot \Omega$;
- Step 6: Compute $\boldsymbol{\Sigma}_{pred} = \boldsymbol{P}^T\boldsymbol{P}$.

From there, the approximated predictive PDF is set as the new prior PDF, while the samples $\boldsymbol{\theta}_i^{j+1}$ are set as the new prior samples.

Following which, the algorithm proceeds to tune the step-size u^{j+1} for the AIES sampler based on the current acceptance rate of the samples α^j. With reference to the target acceptance rate α_{tr} defined as (Roberts and Rosenthal (2001)):

$$\alpha_{tr} = \frac{0.21}{N_{dim}} + 0.23 \qquad (6)$$

The nominal step-size is computed:

$$u_{nom} = u^j \cdot exp\left(\alpha^j - \alpha_{tr}\right) \qquad (7)$$

If $u_{nom} < 1$, u^{j+1} is set to 1.01, otherwise, $u^{j+1} = u_{nom}$. $u^{j=1}$ is set at 2.

Finally, the algorithm proceeds to iteration $j = j + 1$ and repeats the procedure from the re-computation of sample weights.

2.1. *Advantages of Proposed Work*

For the SMC samplers, one way to approximate the predictive PDF would be through the convolution of $P(\boldsymbol{\theta}^j|\boldsymbol{D}^{1:j}, M)$ and the inverse Markov model $\tilde{M}^{-1}$:

$$P(\boldsymbol{\theta}^{j+1}|\boldsymbol{D}^{1:j}, M) \approx G \circ \tilde{M}^{-1}(\boldsymbol{\theta}^{j+1}) \qquad (8)$$

whereby $G(\bullet)$ is the analytical function of the Posterior PDF while $\tilde{M}^{-1}$ is defined as:

$$\tilde{M}^{-1} : \boldsymbol{\theta}^t = \Gamma^{-1}\left(\boldsymbol{\theta}^{t+1} - \nu_{\boldsymbol{\theta}}\right) \qquad (9)$$

However, not all $\Gamma^{-1}(\bullet)$ have a defined inverse function which would make such approximate not feasible. The VB step in the proposed algorithm solves this problem using Gaussian Mixture Models as approximations based on information from the predictive samples $\boldsymbol{\theta}^{j+1}$.

Another benefit of the VB-SEMC sampler is the adaptive tuning algorithm which automatizes the tuning of the step-size of the AIES sampler. This helps to control the acceptance rates of the sampler in an automatised fashion, ensuring the acceptance rates fall within $[0.15, 0.50]$.

Finally, through the use of the AIES sampler, there is no need to define a proposal distribution unlike MH sampler. This removes the need of the algorithm to compute the covariance matrix of the proposal distribution at each iteration which can be computationally costly under high-dimensions. This also helps to improve the computational efficiency of the VB-SEMC sampler given it only needs to compute the step-size u^j for each iteration which is independent of the dimensionality of the problem.

3. Numerical Example: 2-DOF shear building structure
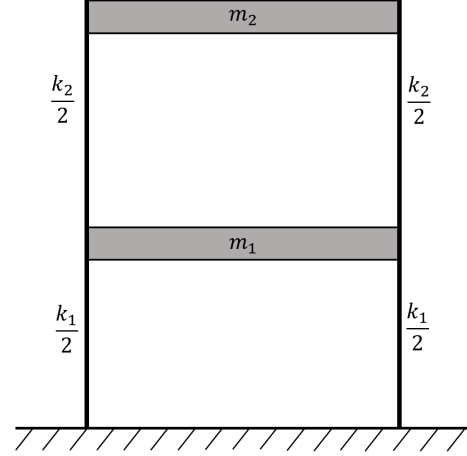

Fig. 1.   Schematic diagram of 2-DOF Shear Building. Image adapted from Betz et al. (2016).

Figure 1 presents the schematic diagram of the 2-DOF shear building set-up based on Beck and Au (2002). The masses of the first and second floors are m_1 and m_2 respectively while the total inter-storey stiffness between the ground and first storey is k_1 and that between the first and second storey is k_2. The dynamics of the structure can be described by the following:

$$\mathbb{M}\ddot{\boldsymbol{x}} + \mathbb{K}\boldsymbol{x} = 0 \qquad (10)$$

where $\mathbb{M}$ and $\mathbb{K}$ are the mass and stiffness matrix which are defined as:

$$\mathbb{M} = \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \qquad (11)$$

$$\mathbb{K} = \begin{bmatrix} k_1 + k_2 & -k_2 \\ -k_2 & k_2 \end{bmatrix} \qquad (12)$$

The floor masses take fixed values such that: $\{m_1, m_2\} = \{1.65, 1.64\} \times 10^4 \ kg$. The stiffness parameters k_1 and k_2 are time-varying such that they degrade with time according to the following Markov model respectively:

$$k_1^{t+1} = exp\left(-1.3\right) \cdot \left[k_1^t - (2.97 \times 10^4)\right] + \\ (2.97 \times 10^4) + \nu_{k_1} \qquad (13)$$

$$k_2^{t+1} = exp\left(-0.6\right) \cdot \left[k_2^t - (2.85 \times 10^4)\right] + \\ (2.85 \times 10^4) + \nu_{k_2} \qquad (14)$$

4   *Adolphus Lye, Ander Gray, and Edoardo Patelli*

where $\nu_{k_1} \sim N(0, \sigma_{k_1})$, $\nu_{k_2} \sim N(0, \sigma_{k_2})$, and $t \in [1, 6]$. Here, $\{\sigma_{k_1}, \sigma_{k_2}\} = \{800, 500\}$ N/m.

The true values of $k_1{}^t$ and $k_2{}^t$ for $t \in [1, 6]$ are presented in Table 1 while the graphical plots of their true values, with their respective nominal evolution models, are presented in Figures 2 and 3.

Table 1.   True values of $k_1{}^t$ and $k_2{}^t$ across different selected times.

t	$k_1{}^t$ (N/m)	$k_2{}^t$ (N/m)
1	4.397×10^4	3.464×10^4
2	3.390×10^4	3.246×10^4
3	2.944×10^4	2.971×10^4
4	3.021×10^4	3.034×10^4
5	2.942×10^4	2.933×10^4
6	2.975×10^4	2.952×10^4
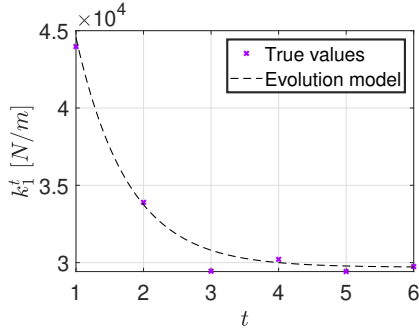
Fig. 2.   True values of $k_1{}^t$ and the corresponding nominal evolution model.

From the structure, measurements are obtained in the form of response frequencies under resonance. Let the 2 response frequencies corresponding to the eigenfrequencies for a given $k_1{}^t$ and $k_2{}^t$ be: $\{f_1{}^t, f_2{}^t\}$. At each t, 15 sets of synthetic data for $\{f_1{}^t, f_2{}^t\}$ are generated following:

$$f_l{}^t = \tilde{f}_l{}^t + \epsilon_l \qquad (15)$$

where $l = \{1, 2\}$, $\tilde{f}_l{}^t$ is the model used to predict $f_l{}^{t^t}$ which is obtained by solving for the eigenvalues of Eq. (10), $\epsilon_l \sim N(0, \sigma_l)$ is the measurement "noise". Here, $\{\sigma_1, \sigma_2\} = \{0.925, 0.153\}$ Hz. The graphical plots of the obtained synthetic data across all t are presented in Figure 4.
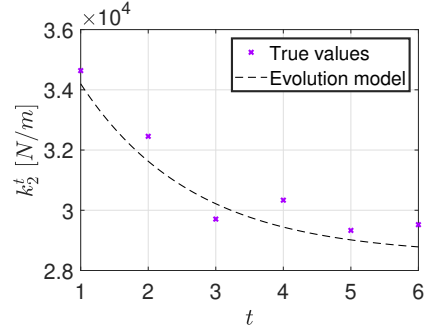

Fig. 3.   True values of $k_2{}^t$ and the corresponding nominal evolution model.

3.1. *Bayesian Inference Set-up*

The objective is to infer $\boldsymbol{\theta}^t = \{m_1, m_2, k_1{}^t, k_2{}^t\}$ based on the data obtained from the measured response frequencies $\{f_1{}^t, f_2{}^t\}$ for $t \in [1, 6]$. The time-invariant mass parameters m_1 and m_2 are included in the Bayesian inference procedure to demonstrate the capability of the VB-SEMC sampler in estimating both time-invariant and time-varying parameters simultaneously.

The prior distribution for all inferred parameters are set to be non-informative uniform distribution whose respective bounds are stated in Table 2.

Table 2.   Bounds of the uniform prior distribution for the respective inferred parameters.

Parameter	Bounds	Units
m_1	$[1.0, 1.0 \times 10^5]$	kg
m_1	$[1.0, 1.0 \times 10^5]$	kg
k_1	$[1.0, 1.0 \times 10^5]$	N/m
k_2	$[1.0, 1.0 \times 10^5]$	N/m

The likelihood function takes on an approximate Normal distribution following (Beck and Au (2002)):

$$P(\boldsymbol{D}^t | \boldsymbol{\theta}^t) = exp \left\{ -\sum_{l=1}^{2} \sum_{m=1}^{15} \frac{\left[\left(\frac{\tilde{f}_l{}^t}{f_{l,m}{}^t} \right)^2 - 1 \right]^2}{2 \cdot \sigma_l{}^2} \right\}$$

(16)

where the argument is the Modal-of-Fit function defined in Beck and Au (2002).
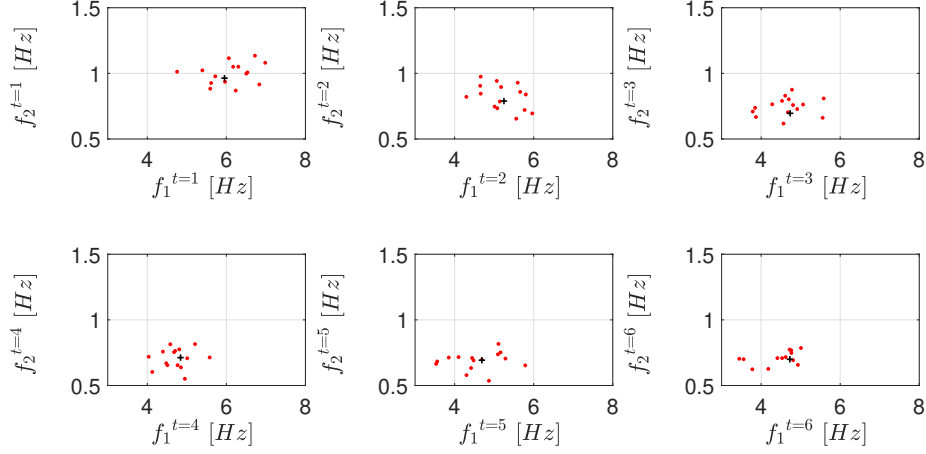
Fig. 4. Graphical plots of the synthetic data for ${f_1}^t$ and ${f_1}^t$ across all t. The black cross denote the model prediction.

3.2. *Results and Discussions*

A total of $N = 1000$ samples are obtained from the posterior at each iteration $j \in [1,6]$ by the SMC and VB-SEMC samplers. The SMC sampler took $46.35\ s$ while the VB-SEMC sampler took $29.78\ s$. This indicates that the proposed VB-SEMC sampler is computationally efficient compared to the SMC sampler.

The resulting estimates by the respective samplers for the inferred parameters are illustrated in the form of graphical plots shown in Figures 5 to 8. The error bars denote the $90\ \%$ credible interval of the estimates where the lower and upper bounds are the 5th and 95th-percentile of the Posterior estimates respectively. As seen in the figures, the resulting estimates of the inferred parameters show a good degree of agreement between the SMC and VB-SEMC samplers which indicates that the proposed sampler is well-verified against the SMC sampler. In addition, it can be seen in Figure 6 that the $90\ \%$ credible interval is significantly smaller for the case of VB-SEMC sampler compared to SMC sampler which indicates the higher-degree of precision of the estimate by the proposed sampler.

It can be noticed however, that not only is the uncertainty relatively large for m_1 and m_2, but that the Posterior mean also deviates significantly from the true values. This could be attributed to the poor sensitivity of the parameters m_1 and m_2 to the data of $\{{f_1}^t, {f_2}^t\}$ given the model used for the Bayesian inference procedure. This results in the Posteriors of m_1 and m_2 to not deviate much from their prior in terms of the credible intervals defined by those distributions.
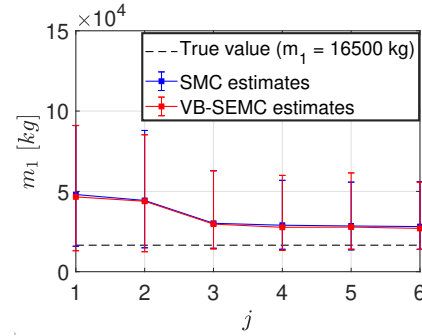

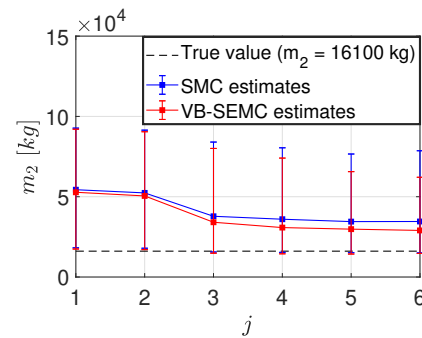Fig. 5. Estimates for m_1 by the SMC and VB-SEMC samplers across iterations j.

Fig. 6. Estimates for m_2 by the SMC and VB-SEMC samplers across iterations j.

Finally, the statistics of the acceptance-rates
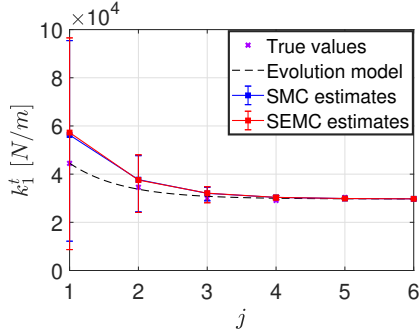
6   *Adolphus Lye, Ander Gray, and Edoardo Patelli*

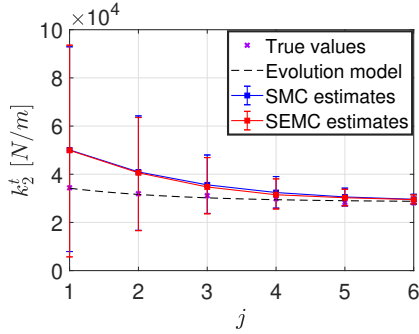Fig. 7.   Estimates for $k_1{}^t$ by the SMC and VB-SEMC samplers across iterations j.

Fig. 8.   Estimates for $k_2{}^t$ by the SMC and VB-SEMC samplers across iterations j.

of the respective samplers across iterations j are presented as graphical plots in Figure 9. As seen in the figure, both samplers are able to achieve acceptable acceptance-rates such that most of which fall within the optimum bounds as suggested by Roberts and Rosenthal (2001). This illustrates the effectiveness of the proposed adaptive tuning algorithm in controlling the acceptance-rates of the VB-SEMC sampler.

4. Application: Non-linear Time-series Model

This section presents a non-linear time-series set-up adapted from the benchmark filtering problem in Cappe et al. (2007). Consider a hypothetical sub-system whose dynamical response signal δ^t for $t \in [0, 100]$ is modelled by:

$$\delta^t = \frac{(v^t)^2}{20} + \epsilon_\delta \qquad (17)$$

where $\epsilon_\delta \sim N(0, \sigma_\delta)$ for $\sigma_\delta = 1$, while v^t is the time-varying parameter whose Markov model is defined as:
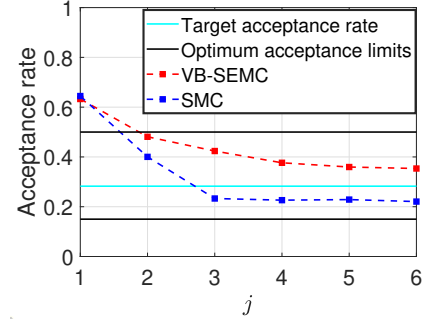

Fig. 9.   Evolution of acceptance-rates across iterations j for the SMC and VB-SEMC samplers. Target acceptance rate is 0.283.

$$v^{t+1} = \frac{v^t}{2} + 25 \cdot \frac{v^t}{1 + v^{t^2}} + 8 \cdot cos\left[1.2(t+1)\right] + \nu_v \quad (18)$$

where $\nu_v \sim N(0, \sigma_v)$ for $\sigma_v = \sqrt{10}$, and $v^{t=0} \sim N(0, \sigma_v)$. A graphical plot of v^t and δ^t for $t \in [0, 100]$ is presented in Figures 10 and 11, respectively.
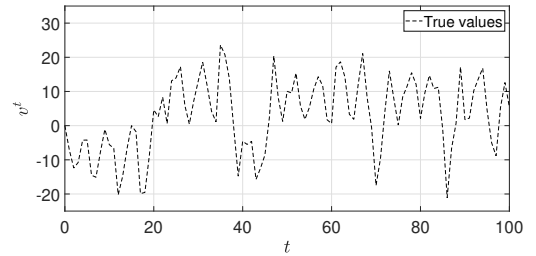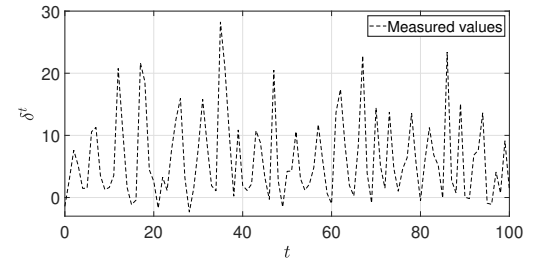

Fig. 10.   Plots of v^t across all t.

Fig. 11.   Plots of δ^t across all t.

4.1. *Bayesian Inference Set-up*

The objective is to infer $\boldsymbol{\theta}^t = \{v^t, \sigma_\delta\}$ based on the data of δ^t obtained for $t \in [0, 100]$.

The prior distribution for all inferred parameters are set to be non-informative Uniform distribution whose respective bounds are: $v^t \in [-30, 30]$ and $\sigma_\delta \in [0.01, 10]$. The likelihood function takes on a Normal distribution defined as follows:

$$P(\boldsymbol{D}^t|\boldsymbol{\theta}^t) = \frac{1}{\sigma_\delta \cdot \sqrt{2\pi}} \cdot exp \left[-\frac{\left(\delta^t - \tilde{\delta}^t\right)^2}{2\sigma_\delta{}^2} \right]$$
(19)

where $\tilde{\delta}^t = \frac{(v^t)^2}{20}$ denotes the model prediction of the observed δ^t.

4.2. *Results and Discussions*

A total of $N = 1000$ samples are obtained from the psosterior at each iteration $j \in [1, 101]$ by the VB-SEMC sampler over a duration of $752.97 \ s$.

The resulting estimates of v^t and σ_δ by the VB-SEMC sampler is presented in Figures 12 and 13 respectively. As seen from the figure, the credible interval of the estimates for v^t generally encloses its true values and its mean estimates are in good agreement with the true values as well. This highlights the effectiveness of the proposed VB-SEMC sampler in inferring time-varying parameters with approximations to the predictive PDF using the Gaussian Mixture Model in the VB step.

On the other hand, the estimates for σ_δ did not converge to the true value of 1. Instead, the convergence occurred about the value of 2.151 which corresponds to the Mean Absolute Deviation (MAD) between the observed δ^t and the most probable observed measurements $\hat{\delta}^t$ which can be computed following:

$$MAD = \frac{1}{101} \cdot \sum_{t=0}^{100} |\hat{\delta}^t - \delta^t|$$
(20)

The reason for this deviation is due to the fact that the observed values of δ^t incorporates the process "noise" carried by v^t at each t which adds to the existing measurement "noise" in δ^t. This accumulation of "noise" results in a larger deviation from nominal values predicted by the updated model $\tilde{\delta}^t$ resulting in the error being closer to the MAD value.

The model update for $\tilde{\delta}^t$ with the inferred values of v^t is generated to verify the results against the observed values of δ^t. The resulting plot is presented in Figure 14. As seen in the figure,
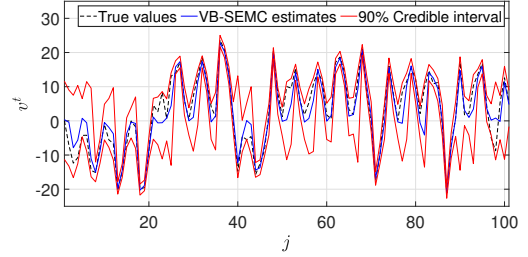

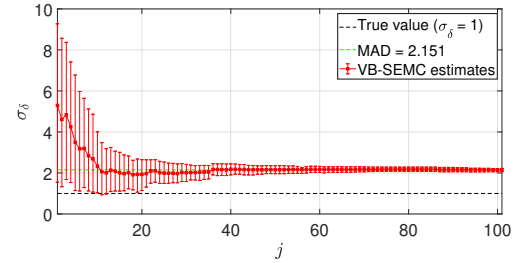Fig. 12.   Estimates of v^t by the VB-SEMC sampler across iterations j.

Fig. 13.   Estimates of σ_δ by the VB-SEMC sampler across iterations j.

the resulting updated model has credible interval which encloses the actual measured vales. This provides verification of the model updating results by the VB-SEMC sampler.
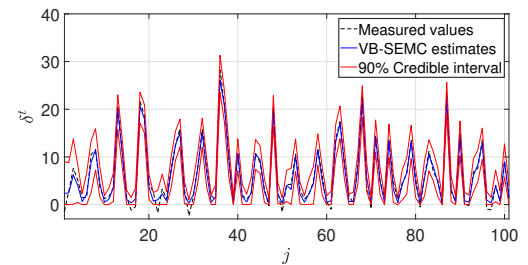

Fig. 14.   Model update for $\tilde{\delta}^t$ by the VB-SEMC sampler across iterations j which is verified against the measured δ^t.

Finally, the statistics of the acceptance-rates by the VB-SEMC sampler is presented in Figure 15. As seen in the figure, the acceptance-rates manage to converge to the target value of 0.335 with increased iterations j which demonstrates the effectiveness of the proposed adaptive tuning algorithm.

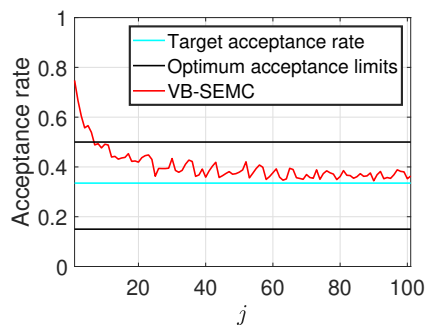8   *Adolphus Lye, Ander Gray, and Edoardo Patelli*

Fig. 15.   Evolution of acceptance-rates across iterations j for the VB-SEMC sampler. Target acceptance rate is 0.335.

5. Conclusion

This paper has presented an efficient Monte Carlo sampler based on the Variational Bayes and implements the Affine-invariant Ensemble Sampler in place of the Metropolis-Hastings sampler as well as an adaptive tuning algorithm to control the acceptance rates. To illustrate the strength and to verify the results obtained by the proposed algorithm, the algorithm has been adopted to estimate the time-variant parameters of a 2-DOF shear building structure where the nominal evolution functions of the respective time-varying parameters have closed-form inverse functions. The results show a very good degree of agreement between the proposed approach and traditional Sequential Monte Carlo sampler. However, the Variational Bayes - Sequential Ensemble Monte Carlo sampler is able to perform the computation in nearly half the time required by the Sequential Monte Carlo sampler. In addition, the proposed sampler is able to control the acceptance-rates such that they converge within the optimal bounds of $[0.15, 0.50]$. Finally, the Variational Bayes - Sequential Ensemble Monte Carlo sampler is then adopted for parameter identification of a non-linear time-series problem where the time-varying parameter has a nominal evolution function which does not yield a closed-form inverse function. Again, the proposed sampler has provided an estimation of the value of the parameter close to the true values for the considered time of interest.

In summary, the proposed Variational Bayes - Sequential Ensemble Monte Carlo sampler provides a computationally-efficient approach with acceptable approximations of the predictive distribution via Variational Bayes step and has a well-controlled acceptance-rates.

References

Beck, J. L. and S. K. Au (2002).   Bayesian Updating of Structural Models and Reliability using Markov Chain Monte Carlo Simulation. *Journal of Engineering Mechanics 128*.

Beck, J. L. and L. S. Katafygiotis (1998). Updating Models and Their Uncertainties. I: Bayesian Statistical Framework. *Journal of Engineering Mechanics 124*, 455–461.

Betz, W., I. Papaioannou, and D. Straub (2016). Transitional Markov Chain Monte Carlo: Observations and Improvements. *Journal of Engineering Mechanics 142*.

Cappe, O., S. J. Godsill, and E. Moulines (2007). An Overview of Existing Methods and Recent Advances in Sequential Monte Carlo. *Proceedings of the IEEE 95*, 899–924.

Chang, P. C., A. Flatau, and S. C. Liu (2003). Review Paper: Health Monitoring of Civil Infrastructure. *Structural Health Monitoring: An International Journal 2*, 257–267.

Chopin, N. (2002).   A sequential particle filter method for static models. *Biometrika 89*, 539–551.

Doucet, A., N. D. Freitas, and N. Gordon (2011). *Sequential Monte Carlo methods in practice*. New York: Springer.

Goodman, J. and J. Weare (2010). Ensemble samplers with affine invariance. *Communications in Applied Mathematics and Computational Science 5*, 65–80.

Lye, A., A. Cicirello, and E. Patelli (2019).   A Review of Stochastic Sampling Methods for Bayesian Inference Problems. *In Proceedings of the 29th European Safety and Reliability Conference*, 1866–1873.

Lye, A., A. Cicirello, and E. Patelli (2021). Sampling methods for solving Bayesian model updating problems: A tutorial. *Mechanical Systems and Signal Processing 159*, 107760.

Ormerod, J. T. and M. P. Wand (2012). Explaining Variational Approximations. *The American Statistician 64*, 140–153.

Patelli, E., M. Broggi, M. D. Angelis, and M. Beer (2014). OpenCossan: An efficient open tool for dealing with Epistemic and Aleatory Uncertainties. *Vulnerability, Uncertainty, and Risk 1*.

Roberts, G. O. and J. Rosenthal (2001). Optimal scaling for various Metropolis-Hastings algorithms. *Statistical Science 16*, 351–367.

Rocchetta, R., M. Broggi, Q. Huchet, and E. Patelli (2018). On-line Bayesian model updating for structural health monitoring. *Mechanical Systems and Signal Processing 103*, 174–195.