Designing Likelihood Function in Block Particle Filter for Beating Curse of Dimensionality

TIANZHI LI, MARIO ROMERO and XI VINCENT WANG

ABSTRACT

Particle filter (PF) has proven effective for nonlinear identification scenarios; however, its performance in high-dimensional problems is often limited by the curse of dimensionality. To overcome this challenge, block particle filter (BPF) is proposed to reformulate a high-dimensional model into several blocks, so the identification of one high-dimensional system can be simplified into that for many lower-dimensional blocks. However, due to the coupling between blocks, the likelihood function for each state subgroup depends not only on its own state components but also on its neighboring subgroups - a dependency that the BPF does not address. In this work, we extend the BPF by developing likelihood functions that incorporate nuisance components, thereby enabling its application to coupled systems. The performance of the proposed algorithms is demonstrated through a numerical example of a forty-story Bouc-Wen frame structure subjected to ground motion.

INTRODUCTION

Bayesian estimation methods - including the extended Kalman filter, unscented Kalman filter, and particle filter (PF) - have been extensively explored in a variety of engineering fields. The estimation results can then serve specific purposes, including diagnostics [1, 2], prognostics [3, 4], target tracking [5, 6], and end-to-end learning [7, 8]. However, challenges related to high dimensionality and nonlinearity often emerge, when state variables and parameters, commonly referred to as state components, are integrated into the state vector and interacted in the process and measurement equations.

Within the field of structural health monitoring (SHM), PF has been well demonstrated for nonlinear applications, including structural system identification [9, 10], and fatigue damage growth prognostics in metal [11] and composites [1]. However, these ap-

Tianzhi Li, Postdoctoral Researcher, Email: tianzhil@kth.se. Department of Production Engineering, KTH Royal Institute of Technology, Stockholm, Sweden
Mario Romero, Associate Professor, Email: mario.romero@liu.se. Department of Science and Technology, Linkoping University, Norrkoping, Sweden
Xi Vincent Wang, Associate Professor, Email: wangxi@kth.se. Department of Production Engineering, KTH Royal Institute of Technology, Stockholm, Sweden

plications potentially involve the deployment of multiple sensor measurements and the estimation of a large number of state components, leading to challenges associated with high dimensionality. As the dimension of the state vector expands, PF often demands an exponentially larger number of particles. This requirement can lead to a significant, and sometimes prohibitive, increase in computational resources needed, which is commonly referred to as the curse of dimensionality [12]. To mitigate this issue, state partitioning has been applied to SHM practice, where one state vector can be divided into two [9] or more [1] state subgroups for estimation. qqq The curse of dimensionality also presents a significant challenge, when applying PF to other engineering applications, notably in data assimilation. Rebeschini and Van Handel [12] have claimed that the decay of correlations property can offer the possibility to tackle the curse of dimensionality. As a high-dimensional system is spatially distributed over large distances, the conditional distribution of one state subgroup should be significantly dependent on the observations collected at its nearby location. Then, the posterior distribution of each subgroup can be approximated by the nearby measurements through one PF. This leads to a simple but effective PF for beating the curse of dimensionality, i.e., block PF. Though this work is theoretical in nature and has not been anticipated to be applicable to real high-dimensional problems [12], the idea of block or partitioning has led to many successful applicable PFs, including iterated BPF [13] and adaptive BPF [14].

The above BPFs have shown to be effective in certain high-dimensional applications, where the likelihood for each state subgroup only includes the state components from this subgroup and its nearby measurements. This can not be directly applicable to the scenario showing more complex coupling among each state subgroup within the measurement equation. Let us take the structural system identification model [1] as one reference example. A whole state vector from this model can be partitioned into several subgroups, each consisting of components (e.g., velocity and displacement) from a single degree of freedom (DOF). For each subgroup, the likelihood includes components from within the subgroup and from others, alongside specific measurement components. As such, the transferability of BPF to other complex applications is likely limited.

For the likelihood of each state subgroup, the state components from other subgroups are taken as nuisance components. Therefore, developing an applicable BPF under coupling scenarios requires the design of likelihood function under nuisance components. In this study, we incorporate three most popular likelihood functions - plug-in likelihood [15], profile likelihood [16], and marginal likelihood [17] - into the likelihood function for each state subgroup, and apply BPF for state and parameter estimation in structural identification problem.

The rest of this paper is organized as follows: First, the block particle filter and likelihood functions under nuisance components are introduced. Then, the proposed algorithm is validated through a numerical simulation of a forty-story Bouc-Wen frame structure. Finally, the paper and potential directions for future research are summarized.

BLOCK PARTICLE FILTER

The state space model for a given system can be described as

$$\boldsymbol{X}_k = f\left(\boldsymbol{X}_{k-1}, \boldsymbol{\omega}_{k-1}\right) \tag{1}$$

$$\boldsymbol{X}_k = h\left(\boldsymbol{X}_k, \boldsymbol{\eta}_k\right) \tag{2}$$

where Eq.(1) and Eq.(2) are the process and measurement equations, respectively, $\boldsymbol{X}$ and $\boldsymbol{Y}$ denote the state and measurement vectors, respectively, k means the k th time step, and $\boldsymbol{\omega}$ and $\boldsymbol{\eta}$ represent the process and measurement noises, respectively.

In high-dimensional models, e.g., Lorenz 96 [14] and measles transmission [13] models, the state and measurement vectors are partitioned into b subgroups:

$$\boldsymbol{X}_k = \begin{bmatrix} \boldsymbol{x}_k^1 \\ \boldsymbol{x}_k^2 \\ \vdots \\ \boldsymbol{x}_k^b \end{bmatrix}, \quad \boldsymbol{Y}_k = \begin{bmatrix} \boldsymbol{y}_k^1 \\ \boldsymbol{y}_k^2 \\ \vdots \\ \boldsymbol{y}_k^b \end{bmatrix}. \tag{3}$$

Note that each state subgroup has unique state components not shared with other subgroups, and each measurement subgroup likewise has unique measurement components. Then a high-dimensional model can be partitioned into b blocks, each having a state and measurement subgroup and one local measurement equation as

$$\boldsymbol{y}_k^j = h^j\left(\boldsymbol{x}_k^j, \boldsymbol{\eta}_k^j\right), \tag{4}$$

where the superscript j means the j th subgroup.

The dynamics and measurements at a spatial location depend only on the states at locations in a neighborhood, a phenomenon known as the decay of correlation property. Therefore, each state subgroup can evolve by itself and its nearby subgroups, and its likelihood can be calculated by its corresponding measurement subgroup only, as illustrated in Fig. 1.

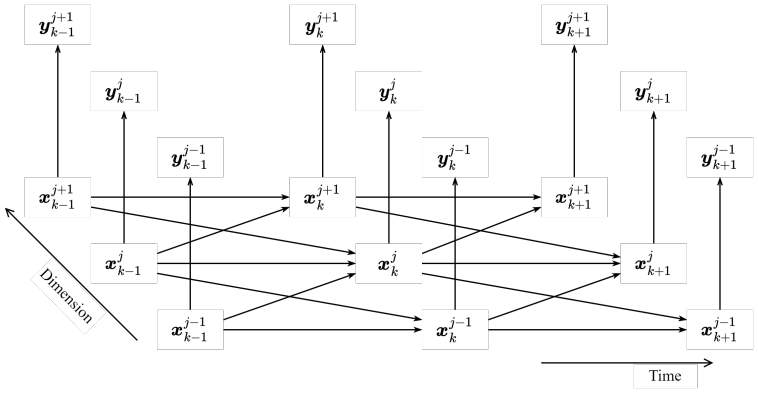

Figure 1. Example of one high-dimensional model in [12].

The above partitioning leads to the development of BPF algorithm. The one-step-ahead prediction is the same as PF. Then, each block includes one independent PF, i.e., the weight is calculated as

$$w_k^{j,i} \propto p\left(\boldsymbol{y}_k^j \mid \boldsymbol{x}_k^{j,i}\right), \tag{5}$$

and the resampling is implemented within each block. The BPF is summarized below.

Algorithm: Block Particle Filter

For $k = 0$:
  Initialization: draw N particles $\{\boldsymbol{X}_k^i : i = 1, 2, \ldots, N\}$.
For $k = 1, 2, \ldots$:
  Prediction: draw N particles $\{\boldsymbol{X}_k^i : i = 1, 2, \ldots, N\}$ by $p(\boldsymbol{X}_k \mid \boldsymbol{X}_{k-1})$.
  For each block: calculate the particle weight through $p\left(\boldsymbol{y}_k^j \mid \boldsymbol{x}_k^{j,i}\right)$ and resample.

LIKELIHOOD UNDER NUISANCE COMPONENTS

BPF and its extensions such as adaptive BPF [14] and iterated BPF [13] have been demonstrated in state and parameter estimation for high-dimensional models, e.g., linear Gaussian model [14], Lorenz 96 model [14], and measles transmission model [13]. These models do not involve the coupling of each subgroup within the measurement equation, as presented in Fig. 2 (a). However, many dynamic systems often exhibit the coupling phenomenon in Fig. 2 (b). For example, when formulating a state space model to describe the dynamic behavior of a multiple degree-of-freedom (DOF) structure, it is common to consider the stiffness, displacement, and velocity as the state components to be estimated, with acceleration serving as the measurement [1, 2].

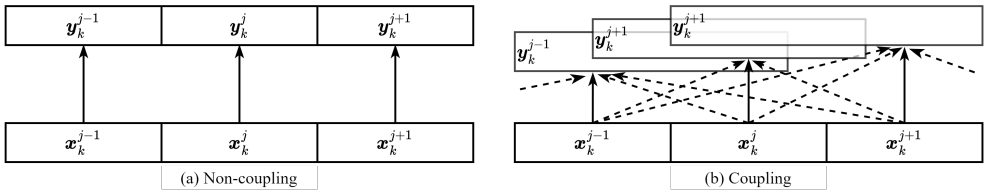

Figure 2. Coupling between Blocks

When each state subgroup still does not share same components with others, the likelihood for one subgroup requires the measurements and the state components from the other subgroups [1]. Then Eq.(4) is no longer valid, and the local measurement equation has to be re-built as

$$\boldsymbol{y}_k^j = h^j\left(\boldsymbol{x}_k^j, \boldsymbol{q}_k^j, \boldsymbol{\eta}_k^j\right), \tag{6}$$

where $\boldsymbol{q}_k^j$ means a vector of state components from the other subgroups that are necessary for the likelihood of j th subgroup. As the interest is to estimate $\boldsymbol{x}_k^j$ instead of $\boldsymbol{q}_k^j$, the later can be viewed as a vector of nuisance components. Then, it becomes clear that $p\left(\boldsymbol{y}_k^j \mid \boldsymbol{x}_k^{j,i}\right)$ has to be obtained under the existence of $\boldsymbol{q}_k^j$. To do so, this study adopts the three most common likelihoods, i.e., plug-in likelihood [15, 18], profile likelihood, and marginal likelihood.

The plug-in likelihood [15, 18] is often the simplest approach for handling nuisance components, where the estimates of nuisance components are used for the likelihood of

each subgroup. This study considers either a single sample selected from the prior distribution $\{\boldsymbol{q}_k^i : i = 1, 2, \ldots, N\}$ or the mean values $\hat{\boldsymbol{q}}_k^j$ of these samples as the approximate estimate, yielding two variants of the plug-in likelihoods.

$$p\left(\boldsymbol{y}_k^j \mid \boldsymbol{x}_k^{j,i}\right) = p\left(\boldsymbol{y}_k^j \mid \boldsymbol{x}_k^{j,i}, \boldsymbol{q}_k^{j,i}\right), \tag{7}$$

$$p\left(\boldsymbol{y}_k^j \mid \boldsymbol{x}_k^{j,i}\right) = p\left(\boldsymbol{y}_k^j \mid \boldsymbol{x}_k^{j,i}, \hat{\boldsymbol{q}}_k^j\right), \tag{8}$$

As to the profile likelihood, it effectively profiles out the nuisance vector by maximizing the likelihood over this vector, allowing the focus to remain on the components of interest. Formally, the profile likelihood can be written as

$$p\left(\boldsymbol{y}_k^j \mid \boldsymbol{x}_k^{j,i}\right) = \max_{\boldsymbol{q}_k^j} p\left(\boldsymbol{y}_k^j \mid \boldsymbol{x}_k^{j,i}, \boldsymbol{q}_k^j\right), \tag{9}$$

Within the scope of PF algorithm, either deterministic or stochastic sampling method can be a choice for obtaining the profile likelihood. T representative samples $\{\boldsymbol{q}_k^{j,t} : t = 1, 2, \ldots, T\}$ can be drawn or selected from the given prior distribution, yielding T likelihoods. Then the maximum likelihood is taken as the profile likelihood. This study employs two strategies to generate representative samples for the nuisance components: the sigma-point (SP) method [5] and the sub-sampling (SS) method.

The marginal likelihood aims to marginalize out the nuisance components as

$$p\left(\boldsymbol{y}_k^j \mid \boldsymbol{x}_k^{j,i}\right) = \frac{1}{T} \sum_{t=1}^{T} p\big(\boldsymbol{y}_k^j \mid \boldsymbol{x}_k^{j,i}, \boldsymbol{q}_k^{j,t}\big) \quad , \tag{10}$$

In order to perform a reasonable comparison, this study still employs the SP and SS methods to generate samples representing the nuisance components.

NUMERICAL EXAMPLE

A Forty-story Bouc-Wen frame structure under the ground motion serves as the numerical example for evaluating the BPF performances under different likelihoods. The stiffness, velocity, and hysteretic displacement of each story are estimated, and the accelerations of certain stories serves as the measurement. The state space model can be built as [1]

$$\boldsymbol{X}_k = \begin{bmatrix} s_{1,k} \\ \vdots \\ v_{1,k} \\ \vdots \\ z_{1,k} \\ z_{2,k} \\ \vdots \end{bmatrix} = \begin{bmatrix} s_{1,k-1} \\ \vdots \\ v_{1,k-1} \\ \vdots \\ z_{1,k-1} \\ z_{2,k-1} \\ \vdots \end{bmatrix} + \begin{bmatrix} \omega_{1,k-1} \\ \vdots \\ a_{1,k-1}\Delta t + \omega_{41,k-1} \\ \vdots \\ \Delta t\Big(v_{1,k-1} - \beta\left|v_{1,k-1}\right|\left|z_{1,k-1}\right|^{n-1} z_{1,k-1} \\ -\gamma v_{1,k-1}\left|z_{1,k-1}\right|^{n}\Big) + \omega_{81,k} \\ \Delta t\Big(\left(v_{2,k-1} - v_{1,k-1}\right) - \beta\left|v_{2,k-1} - v_{1,k-1}\right|\left|z_{2,k-1}\right|^{n-1} z_{2,k-1} \\ -\gamma\left(v_{2,k-1} - v_{1,k-1}\right)\left|z_{2,k-1}\right|^{n}\Big) + \omega_{82,k} \\ \vdots \end{bmatrix}, \tag{11}$$

$$\boldsymbol{Y}_k = \begin{bmatrix} a_{1,k} \\ \vdots \\ a_{40,k} \end{bmatrix} = -\begin{bmatrix} a_{g,k} \\ \vdots \\ a_{g,k} \end{bmatrix} - \begin{bmatrix} \frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_{40}} \end{bmatrix} \left(\begin{bmatrix} c_1 + c_2 & -c_2 & \\ & \ddots & \\ & -c_{40} & c_{40} \end{bmatrix} \begin{bmatrix} v_{1,k} \\ \vdots \\ v_{40,k} \end{bmatrix} + \begin{bmatrix} s_{1,k} + s_{2,k} & -s_{2,k} & \\ & \ddots & \\ & -s_{40,k} & s_{40,k} \end{bmatrix} \begin{bmatrix} z_{1,k} \\ \vdots \\ z_{40,k} \end{bmatrix} \right) + \begin{bmatrix} \eta_{1,k} \\ \vdots \\ \eta_{40,k} \end{bmatrix}, \tag{12}$$

where the subscript l denotes the l th story, a_g means the ground motion, a, v, z, c, m, and s mean the acceleration, velocity, hysteretic displacement, damping, mass, and stiffness, respectively, and β, γ, and n are the Bouc-Wen parameters. The goal is to estimate the stiffness, velocity, and displacement of each story via the settings below:

- Plug-in likelihood function: Eq.(7) & Eq.(8).
- Profile & Marginal likelihood functions: sigma-point method and sub-sampling method (10, 20, or 40 samples).

Fig. 3 resents the stiffness estimations for the third story using PF and BPF under different likelihoods, and Fig. 4 presents the displacement estimations. PF cannot produce accurate estimation, while BPF with each likelihood shows satisfactory performances, demonstrating its capacity in high-dimensional state and parameter estimation.

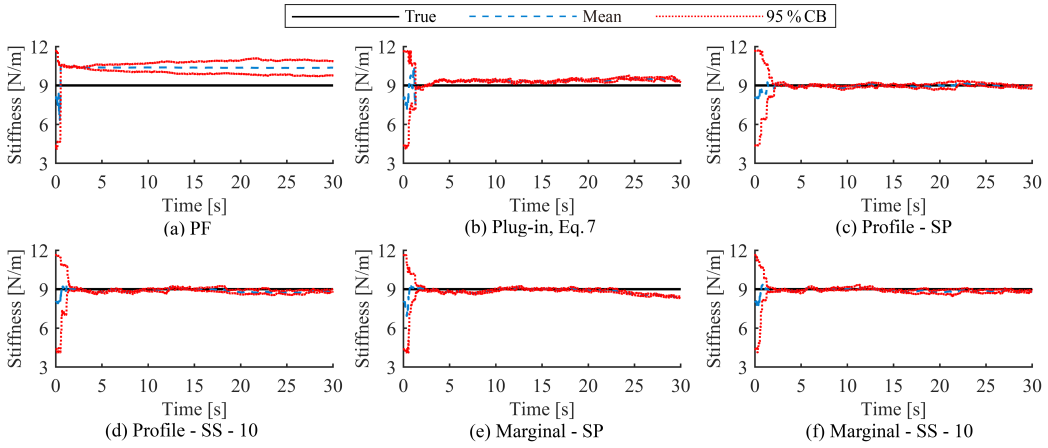

Figure 3. Stiffness estimations for third story using PF and BPF.

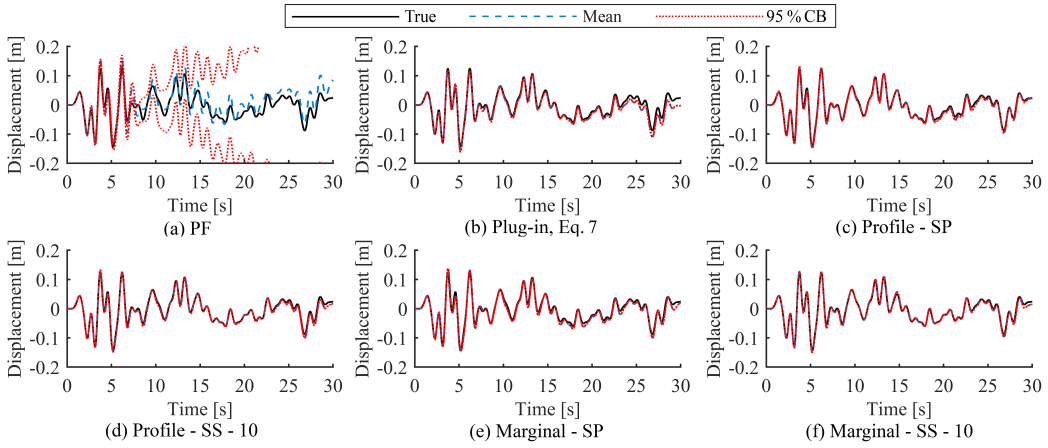

Figure 4. Displacement estimations for third story using PF and BPF.

For a more comprehensive evaluation, PF and BPF are independently run for twenty times. For each of the 120 components, the root-mean-square error (RMSE) is computed, producing twenty RMSE values per component, and then we calculate the mean and standard deviation (STD) for the same type of components. Table I summarizes the average computation time per run, and the means and STDs of RMSEs.

TABLE I. Comparison of BPF Methods with Different Likelihood Functions

Algorithm		Time [s]	RMSE (Mean $\pm$ STD)		
			s [N/m]	v [10^{-2}m/s]	d [10^{-2}m]
PF	–	4.6	2.139 ± 1.344	19.742 ± 28.708	6.471 ± 4.714
BPF - Plug-in	(7)	9.7	1.640 ± 0.747	16.177 ± 18.920	4.351 ± 3.029
	(8)	14.0	1.350 ± 0.678	13.753 ± 16.032	3.551 ± 2.708
BPF - Profile	SP	164.5	0.908 ± 0.442	9.334 ± 8.910	2.179 ± 1.561
	SS-10	87.0	0.991 ± 0.518	10.043 ± 10.537	2.398 ± 1.883
	SS-20	161.8	0.859 ± 0.430	8.816 ± 8.114	2.033 ± 1.410
	SS-40	285.4	$\mathbf{0.819}^{*} \pm 0.394$	$\mathbf{8.753}^{*} \pm 8.090$	$\mathbf{2.017}^{*} \pm 1.428$
BPF - Marginal	SP	163.0	1.754 ± 0.766	16.224 ± 19.395	4.349 ± 2.998
	SS-10	87.4	1.826 ± 0.776	17.329 ± 20.360	4.639 ± 3.117
	SS-20	162.9	1.772 ± 0.771	16.744 ± 19.834	4.474 ± 3.042
	SS-40	286.0	1.774 ± 0.774	16.596 ± 19.946	4.495 ± 3.041

First, PF has the shortest computation time, but produces the least accurate estimation results, due to the curse of dimensionality. By partitioning a high-dimensional system into many blocks, all BPFs can provide more accurate results. Among these likelihood functions, the plug-in likelihoods Eq.(7) and Eq.(8) have the shortest computation time, whereas the BPF using profile or marginal likelihood takes longer computation time. As for the RMSEs, the profile likelihood frequently attains the lowest mean and standard deviation, indicating its superior accuracy and stability. Moreover, the use of more samples in sampling-based likelihood calculation generally leads to better performance.

As to the selection of likelihood selection in BPF, one can use profile likelihood to achieve more accurate estimation compared with plug-in and marginal likelihoods. On the other hand, the performance of plug-in likelihood is sufficiently satisfactory, as such it is a good option to keep a balance between accuracy and computation time. Marginal likelihood is not recommended in this example, as it takes the time as much as profile likelihood, while producing the results even worse than plug-in likelihood.

CONCLUSIONS

Given the complex interaction between each subgroup, the block particle filter (BPF) cannot be directly used for state and parameter estimation. This work has integrated three likelihood functions into BPF, enabling its feasibility in the coupling dynamic system. Results from the numerical example have shown that the proposed solutions can alleviate the curse of dimensionality to a certain level.

REFERENCES

1. Li, T., C. Sbarufatti, and F. Cadini. 2024. “Multiple local particle filter for high-dimensional system identification,” *Mechanical Systems and Signal Processing*, 209:111060.

2. Huang, J., Y. Lei, X. Yang, X. Li, K. Xu, F. Wang, and X. Chen. 2025. "Robust identification of distributed dynamic loads acting on multi-degree-of-freedom structures based on sparse measurement," *Mechanical Systems and Signal Processing*, 228:112417, ISSN 0888-3270, doi:10.1016/J.YMSSP.2025.112417.
3. Li, T. 2024. "Particle filter-based fatigue damage prognosis using prognostic-aided model updating," *Mechanical Systems and Signal Processing*, 211:111244.
4. Liang, H., Y. Liu, J. Chen, S. Yuan, M. Giglio, and C. Sbarufatti. 2025. "An advanced nonlinear framework for early detection and prognosis of fatigue cracks in plate-like structures," *Mechanical Systems and Signal Processing*, 230:112632.
5. Úbeda Medina, L., Ángel F. García-Fernández, and J. Grajal. 2019. "Sigma-point multiple particle filtering," *Signal Processing*, 160:271–283, ISSN 0165-1684, doi:10.1016/J.SIGPRO.2019.02.019.
6. Li, T. and F. Hlawatsch. 2021. "A distributed particle-PHD filter using arithmetic-average fusion of Gaussian mixture parameters," *Information Fusion*, 73:111–124.
7. Ma, X., P. Karkus, D. Hsu, and W. S. Lee. 2020. "Particle filter recurrent neural networks," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 34, pp. 5101–5108.
8. Younis, A. and E. Sudderth. 2024. "Learning to be Smooth: An End-to-End Differentiable Particle Smoother," in A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang, eds., *Advances in Neural Information Processing Systems*, Curran Associates, Inc., vol. 37, pp. 7125–7155.
9. Olivier, A. and A. W. Smyth. 2017. "Particle filtering and marginalization for parameter identification in structural systems," *Structural Control and Health Monitoring*, 24(3):e1874.
10. Cadini, F., L. Lomazzi, M. F. Roca, C. Sbarufatti, and M. Giglio. 2022. "Neutralization of temperature effects in damage diagnosis of MDOF systems by combinations of autoencoders and particle filters," *Mechanical Systems and Signal Processing*, 162:108048.
11. Chen, J., S. Yuan, C. Sbarufatti, and X. Jin. 2021. "Dual crack growth prognosis by using a mixture proposal particle filter and on-line crack monitoring," *Reliability Engineering System Safety*, 215:107758, ISSN 0951-8320, doi:https://doi.org/10.1016/j.ress.2021.107758.
12. Rebeschini, P. and R. van Handel. 2015. "Can local particle filters beat the curse of dimensionality?" *The Annals of Applied Probability*, 25(5):2809 – 2866, doi:10.1214/14-AAP1061.
13. Ning, N. and E. L. Ionides. 2023. "Iterated Block Particle Filter for High-dimensional Parameter Learning: Beating the Curse of Dimensionality," *Journal of Machine Learning Research*, 24(82):1–76.
14. Min, R., C. Garnier, F. Septier, and J. Klein. 2022. "State space partitioning based on constrained spectral clustering for block particle filtering," *Signal Processing*, 201:108727, ISSN 0165-1684, doi:https://doi.org/10.1016/j.sigpro.2022.108727.
15. Popovic, G. C., F. K. Hui, and D. I. Warton. 2018. "A general algorithm for covariance modeling of discrete data," *Journal of Multivariate Analysis*, 165:86–100, ISSN 0047-259X, doi:10.1016/J.JMVA.2017.12.002.
16. Simpson, M. J., S. A. Walker, E. N. Studerus, S. W. McCue, R. J. Murphy, and O. J. Maclaren. 2023. "Profile likelihood-based parameter and predictive interval analysis guides model choice for ecological population dynamics," *Mathematical Biosciences*, 355:108950.
17. Battey, H. and D. Cox. 2020. "High dimensional nuisance parameters: an example from parametric survival analysis," *Information Geometry*, 3(2):119–148.
18. Wu, B., Q. Yao, and S. Zhu. 2013. "Estimation in the presence of many nuisance parameters: Composite likelihood and plug-in likelihood," *Stochastic Processes and their Applications*, 123:2877–2898, ISSN 0304-4149, doi:10.1016/J.SPA.2013.03.017.