State space partitioning based on constrained spectral clustering for block particle filtering

Rui Min (1,2), Christelle Garnier (2), François Septier (3), John Klein (1)

(1) Univ. Lille, CNRS, UMR 9189 - CRIStAL (2) IMT Nord Europe, UMR 9189 - CRIStAL, F-59000 Lille
(3) Univ Bretagne Sud, UMR CNRS 6205, LMBA, F-56000 Vannes

Motivation

- To overcome the curse of dimensionality of the particle filter (PF), the block PF (BPF) proposed by Rebeschini and Van Handel [1] consists in partitioning the state space into several blocks of smaller dimension.
- But the authors provide no method to decide how to split the state space into blocks. This work is an attempt to fill this gap.

  ⇒ **We propose a partitioning method to provide a relevant partition to use in the BPF.**

Block particle filtering [1]

A blocking step is inserted between the prediction and correction steps of the usual bootstrap PF.

- **Partition into K blocks**:
$$\mathbf{x}_t^T = [\mathbf{x}_{t,1}^T \mathbf{x}_{t,2}^T \dots \mathbf{x}_{t,K}^T]$$
where $\mathbf{x}_{t,k} = \{x_t(i) : i \in B_k\}$. The subsets of indexes $\{B_k\}_{k=1}^K$ verify: $\cup_{k=1}^K B_k = \{1 : d_x\}$ and $B_k \cap B_{k'} = \emptyset, \forall k \neq k'$.
- **Approximation of the K marginal densities of each block using local weights**:
The correction and resampling steps are performed independently on each block.
- **Approximation of the joint filtering distribution by the product of the marginal densities**.

Bias and variance of the filtering distribution estimate:

- The variance is reduced using blocks of small dimension.
- But a bias is introduced by the approximation itself and by breaking inter-block correlation (especially at boundaries).

⇒ **How to partition the state space ?**

A new state space partitioning method for block particle filtering

Objective: To mitigate the correlation loss due to the blocking step, the aim is to bring together in the same blocks the most correlated state variables.

Principle of our method: The state space partitioning problem in the BPF is revisited as a clustering problem:

- State variables are seen as data points,
- Clusters are blocks,
- Correlation among variables is used to quantify similarity between data points.

Introduction of two new steps in the BPF:

1. **Estimation of the state vector correlation matrix**

   The N_p predicted particles provide an unbiased estimator of the covariance matrix of the predictive posterior pdf $p(\mathbf{x}_t|\mathbf{y}_{1:t-1})$:
$$\hat{\mathbf{\Sigma}}_{\mathbf{x},t} = \frac{1}{N_p - 1} \sum_{i=1}^{N_p} \left(\mathbf{x}_t^{(i)} - \bar{\mathbf{x}}_t\right) \left(\mathbf{x}_t^{(i)} - \bar{\mathbf{x}}_t\right)^\top,$$
   where:
$$\bar{\mathbf{x}}_t = \frac{1}{N_p} \sum_{i=1}^{N_p} \mathbf{x}_t^{(i)}.$$
   Then the correlation matrix estimate $\hat{\mathbf{C}}_{\mathbf{x},t}$ is directly derived from $\hat{\mathbf{\Sigma}}_{\mathbf{x},t}$.

2. **Constrained spectral clustering for state space partitioning**

   $|\hat{\mathbf{C}}_{\mathbf{x},t}|$ is used as similarity matrix for spectral clustering [2].
   A constraint is introduced on the maximal cluster size to avoid the resurgence of the curse of dimensionality on blocks.

[1] Patrick Rebeschini and Van Handel.
Can local particle filters beat the curse of dimensionality?
The Annals of Applied Probability, 25(5):2809–2866, 2015.

[2] Andrew Ng, Michael Jordan, and Yair Weiss.
On spectral clustering: Analysis and an algorithm.
Advances in neural information processing systems, 14:849–856, 2001.

Constrained spectral clustering

Spectral clustering (SC) [2] is an unsupervised learning approach which relies on two main steps: learning a new representation of the inputs, and then applying an usual clustering approach on the transformed inputs.
To prevent SC from creating large clusters, we set an upper bound ζ on the cluster sizes.

- Compute the normalized symmetric Laplacian $\mathbf{L}_{\mathbf{sym}}$ from the similarity matrix $|\hat{\mathbf{C}}_{\mathbf{x},t}|$.
- Compute the K eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_K$ of $\mathbf{L}_{\mathbf{sym}}$ associated to the K smallest eigenvalues.
- Concatenate these vectors as columns to create the matrix $\mathbf{U} \in \mathbb{R}^{d_x \times K}$ and normalize the rows of $\mathbf{U}$ to unit norm.
- Get the new points $\tilde{\mathbf{z}}_t(i) \in \mathbb{R}^K$ from the i^{th} rows of $\mathbf{U}$, $\forall i \in [d_x]$.
- Partition the new points $\{\tilde{\mathbf{z}}_t(i)\}_{i=1}^{d_x}$ into K clusters $B_1, \dots, B_K$ using the K-means algorithm under the constraint $|B_k| \leq \zeta$, $\forall k \in [K]$.

Experiment results 1

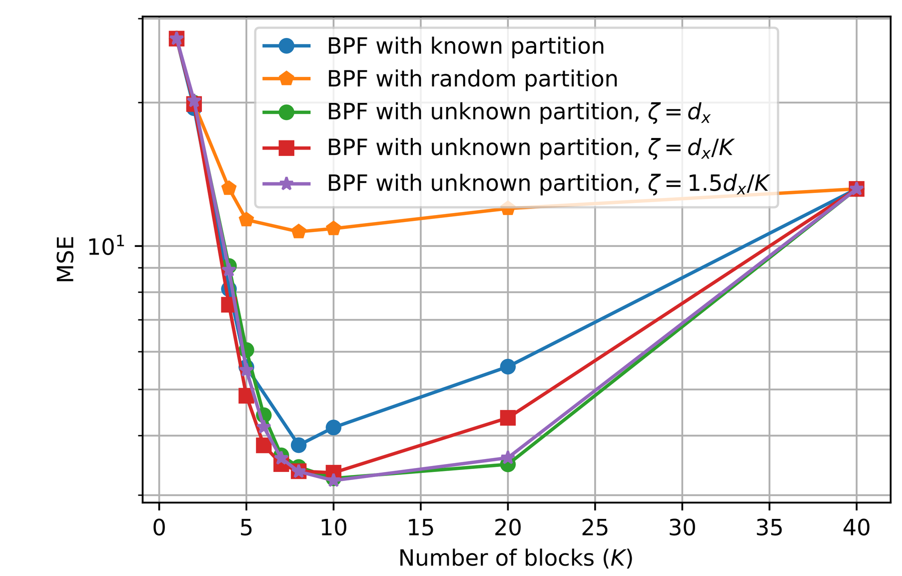
We consider a linear Gaussian model:
$$\mathbf{x}_t = \mathbf{F}\mathbf{x}_{t-1} + \mathbf{w}_t$$
$$\mathbf{y}_t = \mathbf{H}\mathbf{x}_t + \mathbf{v}_t$$
where $d_x = d_y = 100$, $\mathbf{F} = \mathbf{H} = \mathbf{I}$, $\mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$, $\mathbf{v}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and $\mathbf{X}_0 \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. The matrix $\mathbf{Q}$ has 10 blocks with different sizes (from 5 to 15) in the main diagonal and is time-varying. Inside each block, $Q(i,j) = \exp\left(-(i-j)^2/100\right)$.

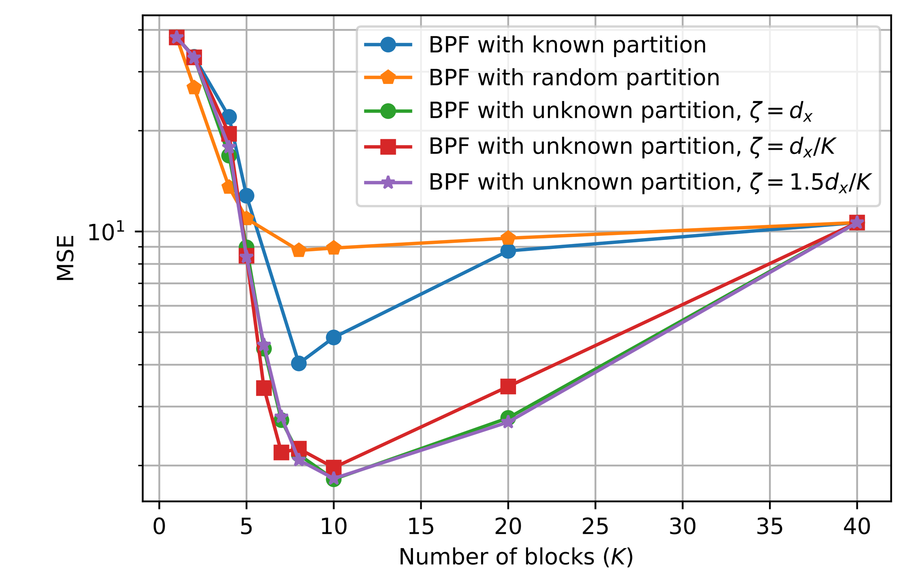
BPF	MSE	ARI
Known partition	0.8185	/
Random partition (same size blocks)	1.1466	/
Unknown partition ($\zeta = d_x$ / unconstrained SC)	0.8190	0.9938
Unknown partition ($\zeta = 10$ / same size blocks)	0.7067	0.7010
Unknown partition ($\zeta = 12$)	0.7473	0.8701
Unknown partition ($\zeta = 15$)	0.8070	0.9942

Table: Performance of BPFs with 10 blocks and 100 particles.

Experiment results 2

We consider a Lorenz 96 non-linear model:
$$\frac{\mathrm{d}x_t(n)}{\mathrm{d}t} = \left(x_t(n+1) - x_t(n-2)\right) x_t(n-1) - x_t(n) + F$$
$$y_t(n) = x_t(2n-1) + v_t(n), \forall n \in 1 \dots d_x/2$$
where $x_t(n)$ is the n^{th} state variable of $\mathbf{x}_t$ and indices follow periodic boundary conditions: $x_t(-1) = x_t(d_x - 1)$, $x_t(0) = x_t(d_x)$ and $x_t(d_x+1) = x_t(1)$.
$d_x = 40$, $F = 8$, $v_t(n) \sim \mathcal{N}(0,1)$ and $\mathbf{x}_0 \sim \mathcal{N}(\mathbf{0}, 0.01 \times \mathbf{I})$.





Contributions

- Online and automatic partitioning of high dimensional state spaces for BPF.
- Revisiting the state space partitioning problem as a clustering problem.
- Using the state correlation matrix estimated from predicted particles as similarity matrix for SC.
- Adding a constraint to prevent SC from creating too large blocks.