A PROBABILISTIC SCHEME FOR JOINT PARAMETER ESTIMATION AND STATE FORECASTING IN A METEOROLOGICAL MODEL

Sara Pérez-Vieites⋆, Inés P. Mariño†, Joaquín Míguez‡

Department of Signal Theory and Communications, Universidad Carlos III de Madrid⋆‡

Department of Biology and Geology, Physics and Inorganic Chemistry, Universidad Rey Juan Carlos†

spvieites@tsc.uc3m.es⋆, ines.perez@urjc.es†, joaquin.miguez@uc3m.es‡

Abstract

We introduce a general methodology, *nested filtering*, that combines two layers of inference: a grid-based scheme to approximate the posterior distribution of the fixed parameters and filtering to track and predict the distribution of the state variables in a recursive way. We specifically explore the use of Monte Carlo and Gaussian filtering methods, but other approaches fit naturally within the new framework.

State-space Model

States: The value of the system state $\boldsymbol{x}(t)$ can be approximated at times t_k applying RK4 method as

$$\bar{\boldsymbol{x}}_k = \bar{\boldsymbol{x}}_{k-1} + \boldsymbol{F}\left(\bar{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}, h, \sigma \boldsymbol{w}_k\right).$$

- $\boldsymbol{F}(\bar{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}, h, \sigma \boldsymbol{w}_k)$: estimate of the time derivatives $\mathsf{d}\boldsymbol{x}(\mathsf{t_k})$.
- $\boldsymbol{w}_k$: zero-mean ind. state noise.
- $\sigma \geq 0$.
- $h > 0$: time-discretisation step.

Observations

Observations: States, $\bar{\boldsymbol{x}}_k \approx \boldsymbol{x}(t_k)$, and unknown parameters, $\boldsymbol{\theta}$, are estimated from a sequence of observation vectors, modelled as

$$\bar{\boldsymbol{y}}_{km} = \boldsymbol{g}(\bar{\boldsymbol{x}}_{km}, \boldsymbol{\theta}) + \sigma_o \boldsymbol{v}_{km}, \quad k = 1, 2, \dots, \quad m \geq 1$$

- $\boldsymbol{g} : \mathbb{R}^{d_x} \to \mathbb{R}^{d_y}$, being $d_y \leq d_x$.
- $\boldsymbol{v}_{km}$: zero-mean ind. observational noise.
- $\sigma_o > 0$.
- m: discrete observation period.

Dynamical Model

To put the states and the observations in the same time scale, we work with the pair of random sequences $\boldsymbol{x}_n = \bar{\boldsymbol{x}}_{nm}$ and $\boldsymbol{y}_n = \bar{\boldsymbol{y}}_{nm}$, as

$$\boldsymbol{y}_n = \boldsymbol{g}(\boldsymbol{x}_n, \boldsymbol{\theta}) + \sigma_o \boldsymbol{v}_n, \quad n = 1, 2, \dots \tag{1}$$

$$\boldsymbol{x}_n = \boldsymbol{F}^m(\boldsymbol{x}_{n-1}, \boldsymbol{\theta}, h, \sigma \boldsymbol{v}_n), \tag{2}$$

where $\boldsymbol{F}^m$ represents the transformation from $\boldsymbol{x}_{n-1} = \bar{\boldsymbol{x}}_{(n-1)m}$ to $\boldsymbol{x}_n = \bar{\boldsymbol{x}}_{nm}$ in m steps.

Nested Filters

We design a sequential Monte Carlo method based on the parameter likelihoods $u_n(\boldsymbol{\theta}) = p(\boldsymbol{y}_n | \boldsymbol{\theta}, \boldsymbol{y}_{1:n-1})$. We approximate first the predictive measure $p(\boldsymbol{x}_n | \boldsymbol{\theta}, \boldsymbol{y}_{1:n-1}) d\boldsymbol{x}_n$ and then use use the relationship

$$u_n(\boldsymbol{\theta}) = \int p(\boldsymbol{y}_n | \boldsymbol{x}_n, \boldsymbol{\theta}) p(\boldsymbol{x}_n | \boldsymbol{\theta}, \boldsymbol{y}_{1:n-1}) d\boldsymbol{x}_n.$$

General Nested Filter

Initialisation:Draw $\boldsymbol{\theta}_0^i, i = 1, \dots, N$, i.i.d. samples from $p(\boldsymbol{\theta})d\boldsymbol{\theta}$. **Recursive step**:

1. For $i = 1, \dots, N$:
   (a) Draw $\bar{\boldsymbol{\theta}}_n^i$ from a Markov kernel $\kappa_N(d\boldsymbol{\theta} | \boldsymbol{\theta}_{n-1}^i)$.
   (b) Approximate $\hat{p}(\boldsymbol{x}_n | \bar{\boldsymbol{\theta}}_n^i, \boldsymbol{y}_{1:n-1}) d\boldsymbol{x}_n \approx p(\boldsymbol{x}_n | \bar{\boldsymbol{\theta}}_n^i, \boldsymbol{y}_{1:n-1}) d\boldsymbol{x}_n$.
   (c) Use this approximation to compute $\hat{u}_n(\bar{\boldsymbol{\theta}}_n^i)$ and let $w_n^i \propto \hat{u}_n(\bar{\boldsymbol{\theta}}_n^i)$ be the normalised weight of $\bar{\boldsymbol{\theta}}_n^i$.
2. Resample to obtain the set $\{\boldsymbol{\theta}_n^i\}_{i=1}^N$ and the approximation $p(\boldsymbol{\theta} | \boldsymbol{y}_{1:n}) d\boldsymbol{\theta} \approx \frac{1}{N} \sum_{i=1}^N \delta_{\boldsymbol{\theta}_n^i}(d\boldsymbol{\theta})$.

A variety of techniques can be used in both layers of the filter. In the first one we focus on Monte Carlo schemes (SMC and SQMC), while in the second one we opt for the use of Gaussian filters (EKF and EnKF).

References

[1] Crisan, D., & Miguez, J. (2018). Nested particle filters for online parameter estimation in discrete-time state-space Markov models. Bernoulli, 24(4A), 3039-3086.

[2] Pérez-Vieites, S., Mariño, I. P., & Míguez, J. (2017). A probabilistic scheme for joint parameter estimation and state prediction in complex dynamical systems. arXiv preprint arXiv:1708.03730.

A Stochastic Lorenz 96 Model

Ground truth. The model consists of two sets of dynamic variables, $\boldsymbol{x}$ and $\boldsymbol{z}$, that displays some key features of atmosphere dynamics.

$$\mathsf{d}\boldsymbol{x} = \boldsymbol{f}_1(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{\alpha})\mathsf{d}t + \sigma \mathsf{d}\boldsymbol{w}_1$$
$$\mathsf{d}\boldsymbol{z} = \boldsymbol{f}_2(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{\alpha})\mathsf{d}t + \bar{\sigma} \mathsf{d}\boldsymbol{w}_2$$

Let us assume there are d_x slow variables and L fast variables per slow variable. The components of the maps, $\boldsymbol{f}_1$ and $\boldsymbol{f}_2$ functions (with dimensions d_x and L respectively), can be written as

$$f_{1j}(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{\alpha}) = -x_{j-1}(x_{j-2} - x_{j+1}) - x_j + F - \frac{HC}{B} \sum_{l=(j-1)L}^{Lj-1} z_l,$$

$$f_{2l}(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{\alpha}) = -CBz_{l+1}(z_{l+2} - z_{l-1}) - Cz_l + \frac{CF}{B} + \frac{HC}{B} x_{\lfloor \frac{l-1}{L} \rfloor}.$$

Observations are linear but can only be collected from this system once every m discrete time steps and only one out of K slow variables can be observed.

Forecast model. We use the differential equation:

$$\mathsf{dx_j} = \mathsf{f_j}(\boldsymbol{x}, \boldsymbol{\theta})\mathsf{d}t + \sigma \mathsf{dw_j} = [-\mathsf{x_{j-1}}(\mathsf{x_{j-2}} - \mathsf{x_{j+1}}) - \mathsf{x_j} + \mathsf{F} - \ell(\mathsf{x_j}, \mathsf{a})]\mathsf{d}t + \sigma \mathsf{dw_j} \tag{3}$$

- $\mathsf{a} = [\mathsf{a}_1, \mathsf{a}_2]^\top$: (constant) parameter vector.
- $\boldsymbol{\theta} = [F, \mathsf{a}^\top]^\top$: contains all the parameters.
- $\ell(x_j, \mathsf{a}) = \mathsf{a}_1 x_j^2 + \mathsf{a}_2 x_j$: ansatz, a polynomial in x_j of degree 2, for the coupling term $\frac{HC}{B} \sum_{l=(j-1)L}^{Lj-1} z_l$.

Computer simulations

Integration step	$h = 5 \times 10^{-3}$
Variables parameters	$K = 2$ and $L = 10$
Fixed model parameters	$F = 8$, $H = 0.75$, $C = 10$ and $B = 15$
Noise scaling factors	$\sigma = \frac{h}{4} = 0.25 \times 10^{-3}$ and $\sigma_o = 4$

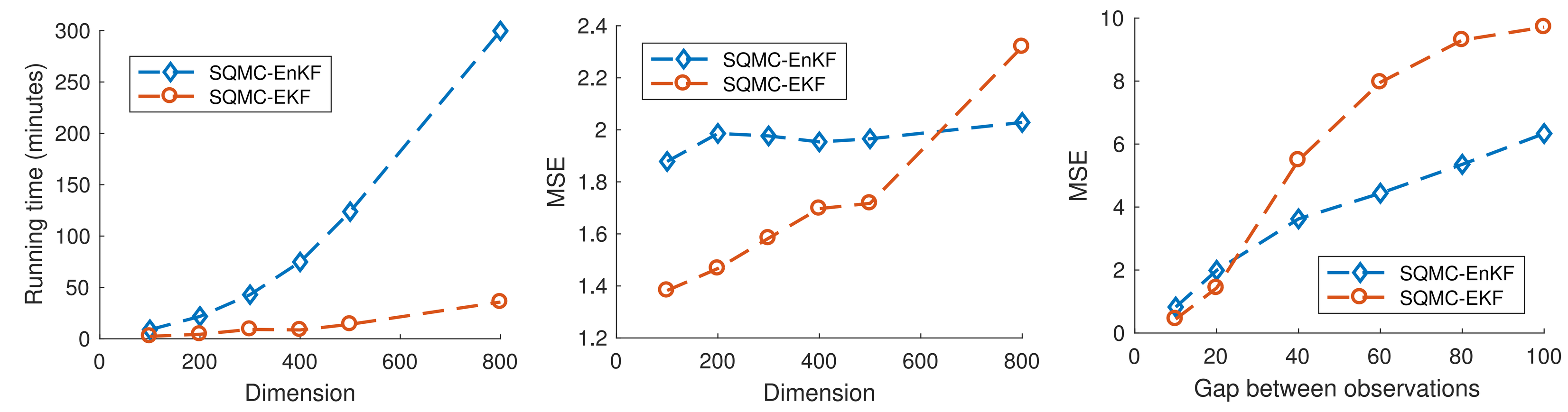

Figure 2: Comparison of the SQMC-EKF (red lines) and SQMC-EnKF (blue lines) in terms of their running time and their MSE as the state dimension d_x increases ($m = 20$ discrete time steps) and as the gap between observations m increases ($d_x = 100$).

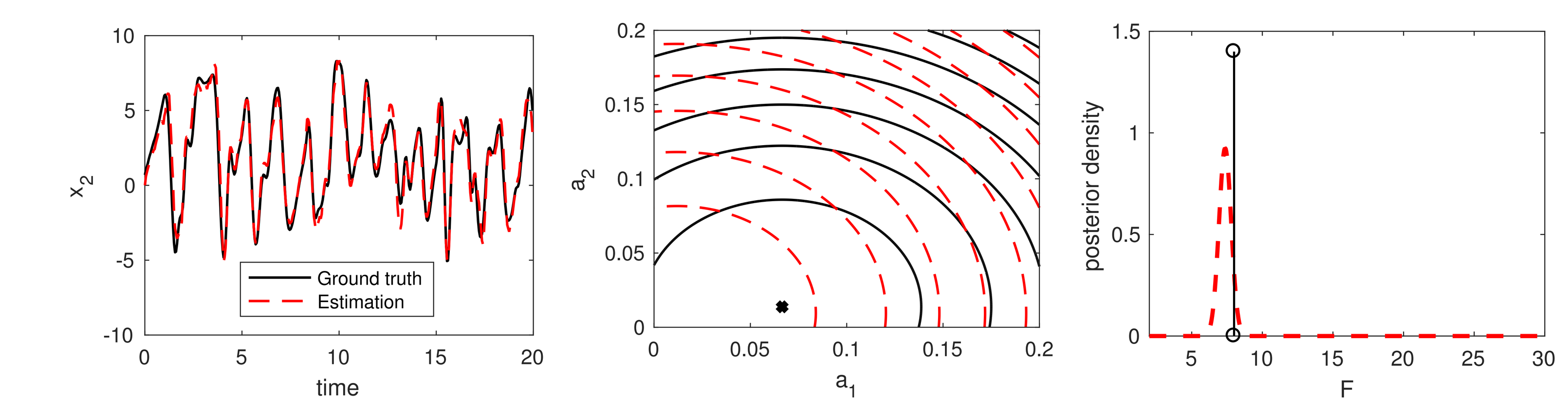

Figure 1: Sequences of state estimates in x_2 (unobserved variable) over time in a 5,000-dimensional Lorenz 96 model using SQMC-EnKF method. Posterior density of the parameters $\mathsf{a} = [\mathsf{a}_1, \mathsf{a}_2]^\top$ and F in $t = 5$. The reference values are represented in black lines.

Summary of contributions

- A nested filtering methodology to recursively estimate the static parameters and the dynamic variables is introduced.
- We combine Monte Carlo and quasi Monte Carlo schemes for the unknown static parameters with either EKF and EnKF for the time-varying states, showing specific results for a 5,000-dimensional system.
- We have proved, under very general assumptions, that the proposed method converges (with optimal Monte Carlo rates) to a possibly biased version of the posterior distribution of the parameters.