Nested Smoothing Algorithms for Joint Parameter and State Estimation of Heterogeneous Multi-Scale State-Space Systems

IMT Nord Europe
École Mines-Télécom
IMT-Université de Lille

uc3m

Sara Pérez-Vieites*, Harold Molina-Bulla†, Joaquín Míguez‡

CERI Systèmes Numériques, IMT Nord Europe, Lille Douai.*
Department of Signal Theory and Communications, Universidad Carlos III de Madrid.† ‡
Emails: sara.perez-vieites@imt-nord-europe.fr*, hmolina@tsc.uc3m.es†, joaquin.miguez@uc3m.es‡

Abstract

We introduce a **recursive methodology** (based on [1]) for Bayesian inference of a class of multi-scale systems (with variables that work at **different time scales**). The proposed scheme combines three intertwined layers of filtering techniques that approximate recursively the **joint posterior probability distribution of the parameters and both sets of dynamic state variables** given a sequence of partial and noisy observations.

State-space Model

We consider a class of **multidimensional stochastic differential equations (SDEs)** that can be written as

$$d\boldsymbol{x} = f_{\boldsymbol{x}}(\boldsymbol{x},\boldsymbol{\theta})d\tau + g_{\boldsymbol{x}}(\boldsymbol{z},\boldsymbol{\theta})d\tau + \boldsymbol{Q}_x d\boldsymbol{v}, \quad (1)$$

$$d\boldsymbol{z} = f_{\boldsymbol{z}}(\boldsymbol{x},\boldsymbol{\theta})d\tau + g_{\boldsymbol{z}}(\boldsymbol{z},\boldsymbol{\theta})d\tau + \boldsymbol{Q}_z d\boldsymbol{w}, \quad (2)$$

- τ denotes continuous time,
- $\boldsymbol{x}(\tau) \in \mathbb{R}^{d_x}$ and $\boldsymbol{z}(\tau) \in \mathbb{R}^{d_z}$ are the **slow and fast states** of the system, respectively,
- $f_{\boldsymbol{x}}$, $g_{\boldsymbol{x}}$, $f_{\boldsymbol{z}}$ and $g_{\boldsymbol{z}}$ are drift functions parameterized by $\boldsymbol{\theta} \in \mathbb{R}^{d_\theta}$,
- Matrices $\boldsymbol{Q}_x$ and $\boldsymbol{Q}_z$ are diffusion coefficients,
- and $\boldsymbol{v}(\tau)$ and $\boldsymbol{w}(\tau)$ are vectors of independent standard Wiener processes.

Dynamical Model

We apply a **macro-micro solver** that runs an **Euler-Maruyama** scheme for each set of **state variables** with different integration steps ($\Delta_x \gg \Delta_z$):

$$\boldsymbol{x}_t = \boldsymbol{x}_{t-1} + \Delta_x(f_{\boldsymbol{x}}(\boldsymbol{x}_{t-1},\boldsymbol{\theta}) + g_{\boldsymbol{x}}(\bar{\boldsymbol{z}}_t,\boldsymbol{\theta})) + \sqrt{\Delta_x}\boldsymbol{Q}_x\boldsymbol{v}_t, \quad (3)$$

$$\boldsymbol{z}_n = \boldsymbol{z}_{n-1} + \Delta_z(f_{\boldsymbol{z}}(\boldsymbol{x}_{\lfloor \frac{n-1}{h} \rfloor},\boldsymbol{\theta}) + g_{\boldsymbol{z}}(\boldsymbol{z}_{n-1},\boldsymbol{\theta})) + \sqrt{\Delta_z}\boldsymbol{Q}_z\boldsymbol{w}_n, \quad (4)$$

where $t \in \mathbb{N}$ denotes discrete time in the time scale of the slow variables, $n \in \mathbb{N}$ denotes discrete time in the fast time scale and

$$\bar{\boldsymbol{z}}_t = \frac{1}{h}\sum_{i=h(t-1)+1}^{ht} \boldsymbol{z}_i. \quad (5)$$

The **observations** are available only in the (slow) time scale of $\boldsymbol{x}$:

$$\boldsymbol{y}_t = l(\boldsymbol{z}_{ht},\boldsymbol{x}_t,\boldsymbol{\theta}) + \boldsymbol{r}_t. \quad (6)$$

References

[1] Pérez-Vieites, S., Mariño, I. P., & Míguez, J. (2018). Probabilistic scheme for joint parameter estimation and state prediction in complex dynamical systems. Physical Review E, 98(6).

[2] Pérez-Vieites, S., Molina-Bulla, H., & Míguez, J. (2022). Nested smoothing algorithms for inference and tracking of heterogeneous multi-scale state-space systems. arXiv preprint arXiv:2204.07795.

Nested Smoother

We aim at performing joint Bayesian estimation of the parameters, $\boldsymbol{\theta}$, and all states, $\boldsymbol{x}$ and $\boldsymbol{z}$ by approximating the posterior pdf

$$p(\boldsymbol{z}_{ht},\boldsymbol{x}_{0:t},\boldsymbol{\theta}|\boldsymbol{y}_{1:t}) = p(\boldsymbol{z}_{ht}|\boldsymbol{x}_{0:t},\boldsymbol{y}_{1:t},\boldsymbol{\theta})p(\boldsymbol{x}_{0:t}|\boldsymbol{y}_{1:t},\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{y}_{1:t}) \quad (7)$$

1^{st} layer

$$p(\boldsymbol{\theta}|\boldsymbol{y}_{1:t}) \propto \underbrace{p(\boldsymbol{y}_t|\boldsymbol{\theta},\boldsymbol{y}_{1:t-1})}_{\text{likelihood of } \boldsymbol{\theta}} \underbrace{p(\boldsymbol{\theta}|\boldsymbol{y}_{1:t-1})}_{\text{posterior pdf at } t-1}$$

2^{nd} layer

$$p(\boldsymbol{y}_t|\boldsymbol{\theta},\boldsymbol{y}_{1:t-1}) = \int \overbrace{p(\boldsymbol{y}_t|\boldsymbol{x}_{0:t},\boldsymbol{y}_{1:t-1},\boldsymbol{\theta})}^{\text{likelihood of } \boldsymbol{\theta} \text{ and } \boldsymbol{x}_{0:t}} \underbrace{p(\boldsymbol{x}_{0:t}|\boldsymbol{y}_{1:t-1},\boldsymbol{\theta})}_{\text{joint predictive pdf of } \boldsymbol{x}_{0:t}} d\boldsymbol{x}_{0:t}$$

3^{rd} layer

$$p(\boldsymbol{y}_t|\boldsymbol{x}_{0:t},\boldsymbol{y}_{1:t-1},\boldsymbol{\theta}) = \int \overbrace{p(\boldsymbol{y}_t|\boldsymbol{z}_{ht},\boldsymbol{x}_t,\boldsymbol{\theta})}^{\text{likelihood of } \boldsymbol{\theta},\, \boldsymbol{x}_t \text{ and } \boldsymbol{z}_{ht}} \times \underbrace{p(\boldsymbol{z}_{ht}|\boldsymbol{x}_{0:t},\boldsymbol{y}_{1:t-1},\boldsymbol{\theta})}_{\text{predictive pdf of } \boldsymbol{z}_{ht}} d\boldsymbol{z}_{ht}$$

In the **second layer**, the joint predictive pdf of $\boldsymbol{x}_{0:t}$ is computed as

$$p(\boldsymbol{x}_{0:t}|\boldsymbol{y}_{1:t-1},\boldsymbol{\theta}) = p(\boldsymbol{x}_t|\boldsymbol{x}_{0:t-1},\boldsymbol{y}_{1:t-1},\boldsymbol{\theta})p(\boldsymbol{x}_{0:t-1}|\boldsymbol{y}_{1:t-1},\boldsymbol{\theta}), \quad (8)$$

and in the **third layer** we can compute $p(\boldsymbol{x}_t|\boldsymbol{x}_{0:t-1},\boldsymbol{y}_{1:t-1},\boldsymbol{\theta})$ as

$$p(\boldsymbol{x}_t|\boldsymbol{x}_{0:t-1},\boldsymbol{y}_{1:t-1},\boldsymbol{\theta}) = \int p(\boldsymbol{x}_t|\boldsymbol{z}_{h(t-1)+1:ht},\boldsymbol{x}_{0:t-1},\boldsymbol{y}_{1:t-1},\boldsymbol{\theta}) \times p(\boldsymbol{z}_{h(t-1)+1:ht}|\boldsymbol{x}_{0:t-1},\boldsymbol{y}_{1:t-1},\boldsymbol{\theta})d\boldsymbol{z}_{h(t-1)+1:ht}.$$

A Stochastic two-scale Lorenz 96 Model

- The system is described, in continuous-time τ, by the **SDEs**

$$dx_j = \left[-x_{j-1}(x_{j-2}-x_{j+1}) - x_j + F - \frac{HC}{B}\sum_{l=(j-1)R}^{Rj-1} z_l\right]d\tau + \sigma_x dv_j,$$

$$dz_l = \left[-CBz_{l+1}(z_{l+2}-z_{l-1}) - Cz_l + \frac{CF}{B} + \frac{HC}{B}x_{\lfloor (l-1)/R \rfloor}\right]d\tau + \sigma_z du$$

Let us assume there are d_x slow variables and R fast variables per slow variable, and $\boldsymbol{\theta} = (F,H,C,B)^\top \in \mathbb{R}$ are **static model parameters**.

- The **discrete-time state equations** can be written as

$$x_{t+1,j} = x_{t,j} + \Delta_x(f_{\boldsymbol{x},j}(\boldsymbol{x}_t,\boldsymbol{\theta}) + g_{\boldsymbol{x},j}(\bar{\boldsymbol{z}}_{t+1},\boldsymbol{\theta})) + \sqrt{\Delta_x}\sigma_x v_{t+1,j}, \quad (9)$$

$$z_{n+1,l} = z_{n,l} + \Delta_z(f_{\boldsymbol{z},l}(\boldsymbol{x}_{\lfloor \frac{n}{h} \rfloor},\boldsymbol{\theta}) + g_{\boldsymbol{z},l}(\boldsymbol{z}_n,\boldsymbol{\theta})) + \sqrt{\Delta_z}\sigma_z w_{n+1,l}, \quad (10)$$

where

$$\boldsymbol{x}_t = (x_{t,0},\dots,x_{t,d_x-1})^\top \quad \text{and} \quad \boldsymbol{z}_n = (z_{n,0},\dots,z_{n,d_z-1})^\top.$$

- We assume that the **observations** are **linear and Gaussian**, namely,

$$\boldsymbol{y}_t = \boldsymbol{A}_t \begin{bmatrix}\boldsymbol{x}_t \\ \boldsymbol{z}_{ht}\end{bmatrix} + \boldsymbol{r}_t, \quad (11)$$

where $\boldsymbol{A}_t$ is a known $d_y \times (d_x + d_z)$ matrix and $\boldsymbol{r}_t$ is a d_y-dimensional Gaussian random vector with known covariance matrix.

Computer simulations

A variety of techniques (Monte Carlo or Gaussian filters such as EnKF, EKF and UKF) can be used in **any layer** of the filter. For this experiment we have implemented a **SMC-EnKF-EKF**.

Integration step	$\Delta_x = 10^{-3}$ and $\Delta_z = 10^{-4}$
Variables parameters	$d_x = 10$, $R = 5$ and $d_z = 50$
Fixed model parameters	$F = 8$, $H = 0.75$, $C = 10$ and $B = 15$
Noise scaling factors	$\sigma_x^2 = \frac{1}{2}$ and $\sigma_z^2 = \frac{1}{16}$

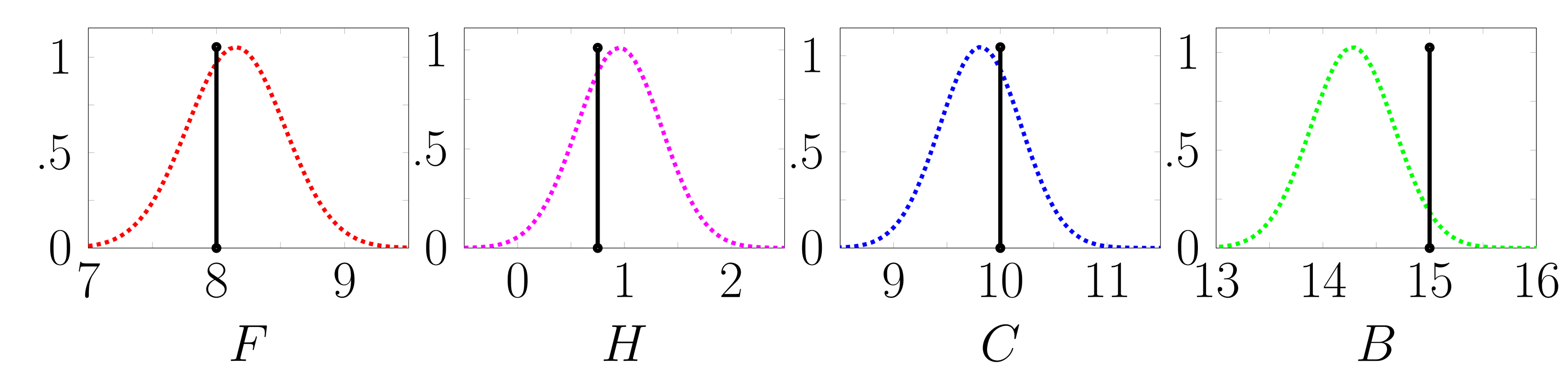

Figure: Posterior density of the parameters (dashed lines) at time $\tau = 20$. The true values are indicated by a black vertical line.

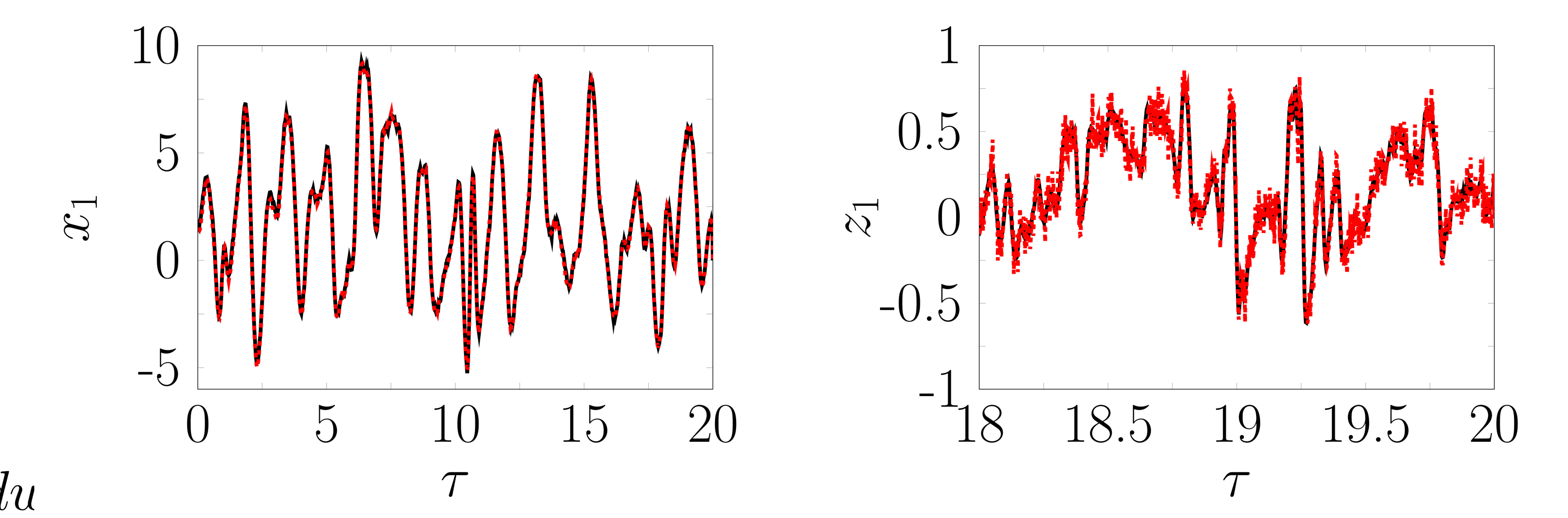

Figure: Sequences of state values (black line) and estimates (dashed red line) in x_1 and z_1 over time.

Summary of contributions

- We have introduced a recursive and multi-layer methodology that estimates the static parameters and the dynamical variables of a class of multi-scale state-space models.
- The inference techniques used in each layer can vary from Monte Carlo to Gaussian techniques, leading to different computational costs and degrees of accuracy.
- We have implemented a SMC-SMC-UKF and a SMC-EnKF-EKF that have obtained good results in terms of accuracy.