Recursive maximum likelihood estimation for structural health monitoring: tangent filter implementations

Fabien Campillo, Laurent Mevel

Recursive maximum likelihood estimation for structural health monitoring: Kalman and particle filter implementations

Fabien Campillo & Laurent Mevel
INRIA/IRISA
Campus de Beaulieu
35042 Rennes cedex, France
Fabien.Campillo@inria.fr   Laurent.Mevel@inria.fr

***Abstract*— Flutter monitoring can be handled by tracking the real time variations of the modal parameters of a specified civil structure, be it a bridge or an aircraft. Previous algorithmic attempts encompass automated batch identification and damage detection through hypothesis testing. Both approaches appear impractical, the first one because of computational time considerations and the difficulty to select a windows length with the best trade off between bias and variance, the second because of the difficulty to obtain reference data set close to flutter regime. Here, we investigate the capabilities of a sample wise recursive Kalman filter–based gradient approach and compare it to its particle filter counterpart.**

I. Introduction

A critical problem for mechanical structures exposed to unmeasured non stationary natural excitation (turbulence) is an instability phenomenon also known as *flutter*. It is formulated as the monitoring of the time varying complex eigenvalues associated to the discretized linear system corresponding to the monitored mechanical system. It has already been investigated through batch identification modal analysis using only output-only in-flight data has already been investigated. See Mevel *et al* [1] for a case study of monitored aircraft using subspace identification methods.

For improving the estimation of the parameters of interest, the collection of frequency and damping coefficients, and moreover for achieving this in real-time during flight tests, one possible route is to resort to tracking algorithms.

Frequency and damping coefficients are monitored by a recursive maximum likelihood (RML) procedure. The considered tracking procedure is a special case of adaptive algorithms where the gain is kept constant. The approximation of the associated score function is evaluated by a joint particle approximation of the conditional law and its derivative w.r.t. to the parameters (the tangent filter). Particle filtering techniques [2] are simple to implement and have many advantages in practice like robustness. Doucet & Tadic [3], Guyader *et al* [4], and Caylus *et al* [5] already applied these techniques to RML estimation. The problem presented here was previously treated by Fichou *et al* [6] in a more simple framework.

Particle approximation for health monitoring was already proposed by Yoshida & Sato [7] in order to handle non–Gaussian noise. Modal characteristics monitoring is also considered by Ching *et al* [8]. In both cases authors use a state augmentation approach by including the unknown parameters in the state process. It seems preferable to directly identify these parameters by a likelihood approach.

In a first part the structural health monitoring problem is written in term of recursive maximum likelihood estimate (RMLE) in a state–space model. Then a Kalman filter expression of the score function is proposed together with an alternate particle filter approximation. Last part is devoted to a case study.

II. The problem

A. Dynamical model and structural parameters

Let us consider observations sampled at a rate $1/\delta$

$$\mathsf{y}_k = \mathbf{L}\, Z(k\,\delta) \tag{1}$$

of the state $Z(t)$ of a n–degrees of freedom mechanical system. These measurements are gathered through d sensors, i.e. y_k takes values in $\mathbb{R}^d$. The matrix $\mathbf{L}$ indicates which components of the state vector are actually measured, i.e. where the sensors are located. The behavior of the mechanical system is described by the following linear dynamical system

$$\mathbf{M}\,\ddot Z(t) + \mathbf{C}\,\dot Z(t) + \mathbf{K}\,Z(t) = \sigma\,\zeta(t) \tag{2}$$

where the (non measured) input force ζ is a non–stationary white Gaussian noise with time-varying covariance matrix $\mathcal{Q}^\zeta(t)$. $\mathbf{M}$, $\mathbf{C}$, $\mathbf{K}$ are respectively the matrices of mass, damping and stiffness.

Now let us describe the structural characteristics of the system (2). The modes or eigenfrequencies μ and the associated eigenvectors Φ_μ of the system (2) are solutions of

$$\begin{aligned} \det[\mu^2\,\mathbf{M} + \mu\,\mathbf{C} + \mathbf{K}] &= 0\,, \\ [\mu^2\,\mathbf{M} + \mu\,\mathbf{C} + \mathbf{K}]\,\Phi_\mu &= 0\,. \end{aligned} \tag{3}$$

Then the mode–shapes are $\Psi_\mu = \mathbf{L}\;\Phi_\mu$. The frequency and damping coefficients are

$$\mathbf{f} = \tfrac{\mathbf{b}}{2\pi}\ \text{(Hz)}\,, \qquad \mathbf{d} = \tfrac{|\mathbf{a}|}{\sqrt{\mathbf{a}^2+\mathbf{b}^2}} \in [0,1] \tag{4}$$

with $\mathbf{a} = \Re(\mu)$ and $\mathbf{b} = \Im(\mu)$.

The monitored structure is defined by its modal characteristics: the collection of frequencies, dampings and mode shapes, as well as the covariances of the noises. The problem is to follow the slow evolutions of the structural characteristics of the mechanical system (2) by a recursive tracking method, whose starting values will be defined as the output of the data driven subspace method as described in Van Overschee & De Moor [9, Fig. 3.13 p. 90].

The tracking algorithm will focus on the frequencies and dampings, the mode shapes are assumed not to change significantly during the monitoring in regard to the changes in the eigenvalues. A change in the mode shapes would most likely be a local change in the structure, thus will indicate the presence of damage, whereas a change in the eigenvalues can still occur without presence of damage and not affect significantly the mode shapes (as for example the effect of temperature on the stiffness of the structure).

B. State–space model and canonical parameterization

We rewrite the preceding system (1)–(2) as a linear state–space model. Define

$$X_k \stackrel{\text{def}}{=} \begin{bmatrix} Z(k\,\delta) \\ \dot Z(k\,\delta) \end{bmatrix}$$

and $F \stackrel{\text{def}}{=} e^{\delta A}$ with $A \stackrel{\text{def}}{=} \begin{bmatrix} 0 & I \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C} \end{bmatrix} \in \mathbb{R}^{2n\times 2n}$.

From (2) we get

$$X_{k+1} = F\,X_k + \sigma\,\zeta_k \tag{5}$$

where $\zeta_k \stackrel{\text{def}}{=} \int_{(k-1)\delta}^{k\delta} e^{(k\delta-u)\,A}\; d\begin{bmatrix} 0 \\ dB_u \end{bmatrix}$ and $B_t \stackrel{\text{def}}{=} \int_0^t \zeta(s)\,ds$ is a Brownian motion. Hence ζ_k is a (discrete–time) white Gaussian noise with covariance matrix

$$\int_{(k-1)\delta}^{k\delta} e^{(k\delta-u)\,A}\;\begin{bmatrix} 0 & 0 \\ 0 & \mathbf{M}^{-1}\,\mathcal{Q}^{\zeta}(u)\,(\mathbf{M}^{-1})^* \end{bmatrix}\; e^{(k\delta-u)\,A^*}\;du$$

which is approximated by $\delta\,\mathcal{Q}_k^{\zeta}$ with

$$\mathcal{Q}_k^{\zeta} \stackrel{\text{def}}{=} \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{M}^{-1}\,\mathcal{Q}^{\zeta}(k\delta)\,(\mathbf{M}^{-1})^* \end{bmatrix}.$$

From (1) we get

$$\mathsf{y}_k = [\mathbf{L}\;0]\,X_k + \nu\,\mathsf{v}_k \tag{6}$$

where $[\mathbf{L}\;0] \in \mathbb{R}^{d\times 2n}$ and v_k is a $N(0,\mathcal{Q}_k^{\mathsf{v}})$ white Gaussian noise which allows to take into account of the errors of modeling and the measurement noise. We suppose that the Hermitian matrix $\mathcal{Q}_k^{\mathsf{v}}$ is positive definite.

Let (λ,Φ_λ) be the eigenstructure of the state transition matrix F, namely

$$\det(F-\lambda\,I)=0\,,\quad (F-\lambda\,I)\,\Phi_\lambda = 0\,. \tag{7}$$

The parameters (μ,Φ_μ) in (3) can be deduced from the (λ,Φ_λ)'s using $e^{\delta\,\mu}=\lambda$ and $\Phi_\mu=\Phi_\lambda$. The frequency and damping coefficients (4) are recovered from a discrete eigenvalue λ through

$$\mathbf{a} = \tfrac{1}{\delta}\,\log|\lambda|\,,\qquad \mathbf{b} = \tfrac{1}{\delta}\,\arctan\left[\tfrac{\Im(\lambda)}{\Re(\lambda)}\right].$$

Hypothesis: *We suppose that F admits $2n$ pairwise complex conjugate distinct eigenvalues $\lambda_{1:n},\bar\lambda_{1:n}$ with associated orthonormal set of eigenvectors $\Phi_{1:n},\bar\Phi_{1:n}$ ([1]). We also suppose that these eigenvalues have modulus less than one.*

[1]Notations: x^T is the transpose of x, $\bar x$ is the complex conjugate, x^* is the transpose/conjugate, $|x|$ the modulus, j will denote $\sqrt{-1}$.

It turns out that this collection of modes forms a very natural parameterization for structural analysis. It is invariant w.r.t. changes in the state basis of system (5)–(6). In other words, the (λ,Φ_λ)'s form a canonical parameterization of the eigenstructure (or equivalently the pole part) of that system.

1) Change of variables: Define

$$\mathbf{\Phi} \stackrel{\text{def}}{=} [\Phi_{1:n}]\,,\quad \mathbf{\Psi} \stackrel{\text{def}}{=} [\Psi_{1:n}]\,,\quad \mathbf{\Lambda} \stackrel{\text{def}}{=} \mathrm{diag}(\lambda_{1:n})\,.$$

We introduce the following linear transformation

$$T \stackrel{\text{def}}{=} [\mathbf{\Phi}\;\bar{\mathbf{\Phi}}] \in \mathbb{C}^{2n\times 2n}\,,$$

i.e. the matrix whose columns are the eigenvectors of F. It is a unitary matrix, i.e. $T^{-1}=T^*$. Then

$$\begin{bmatrix} \mathbf{\Lambda} & (0) \\ (0) & \bar{\mathbf{\Lambda}} \end{bmatrix} = T^*\,F\,T \in \mathbb{C}^{2n\times 2n}\,.$$

Define also

$$H \stackrel{\text{def}}{=} [\mathbf{L}\;0]\,T = [\mathbf{L}\;0]\,[\mathbf{\Phi}\;\bar{\mathbf{\Phi}}] = [\mathbf{\Psi}\;\bar{\mathbf{\Psi}}] \in \mathbb{C}^{d\times 2n}\,,$$

Then after the change of variables

$$\tilde X_k \stackrel{\text{def}}{=} T^*\,X_k\,,$$

the vector $\tilde X_k$ is of the form $\left[\begin{smallmatrix} \mathsf{x}_k \\ \bar{\mathsf{x}}_k \end{smallmatrix}\right]$ and (5) reduces to

$$\mathsf{x}_{k+1} = \mathbf{\Lambda}\,\mathsf{x}_k + \sigma\,\mathbf{\Phi}^*\,\zeta_k\,,\qquad \zeta_k \stackrel{\text{iid}}{\sim} N(0,\delta\,\mathcal{Q}_k^{\zeta})\,.$$

Note that in practice we just have access to the mode shapes matrix $\Psi_{1:n}$ and not to the eigenvectors matrix $\Phi_{1:n}$, so in order to fully specify the state equation we suppose that the covariance matrix $\mathcal{Q}_k^{\zeta}$ is of the form $[\mathbf{L}\;0]^*\,\mathcal{Q}_k\,[\mathbf{L}\;0]$ for a given covariance matrix $\mathcal{Q}_k$. Hence $\mathsf{w}_k \stackrel{\text{def}}{=} \mathbf{\Phi}^*\,\zeta_k$ is a white Gaussian noise with covariance matrix $\mathcal{Q}_k^{\mathsf{w}} \stackrel{\text{def}}{=} \delta\,\mathbf{\Psi}^*\,\mathcal{Q}_k\,\mathbf{\Psi}$.

The observation equation (6) becomes

$$\mathsf{y}_k = \mathbf{\Psi}\,\mathsf{x}_k + \bar{\mathbf{\Psi}}\,\bar{\mathsf{x}}_k + \nu\,\mathsf{v}_k\,,\qquad \mathsf{v}_k \stackrel{\text{iid}}{\sim} N(0,\mathcal{Q}_k^{\mathsf{v}})\,.$$

Note that $\mathbf{\Psi}\,x + \bar{\mathbf{\Psi}}\,\bar x = 2\,\Re\{\mathbf{\Psi}\,x\}$ is a linear operator.

2) The state/space system: One finally obtains the following system

$$\mathsf{x}_{k+1} = \mathbf{\Lambda}\,\mathsf{x}_k + \sigma\,\mathsf{w}_k\,,\qquad \mathsf{w}_k \stackrel{\text{iid}}{\sim} N(0,\mathcal{Q}_k^{\mathsf{w}})\,, \tag{8}$$

$$\mathsf{y}_k = 2\,\Re\{\mathbf{\Psi}\,\mathsf{x}_k\} + \nu\,\mathsf{v}_k\,,\qquad \mathsf{v}_k \stackrel{\text{iid}}{\sim} N(0,\mathcal{Q}_k^{\mathsf{v}})\,. \tag{9}$$

In this model all parameters are assumed known, or previously estimated, except the eigenvalues matrix $\mathbf{\Lambda} \stackrel{\text{def}}{=} \mathrm{diag}(\lambda_{1:n})$ and the noise intensities σ and ν. The mode shapes matrix $\mathbf{\Psi} = [\Psi_{1:n}]$, the sampling period δ, and the covariance matrices $\mathcal{Q}_k$ and $\mathcal{Q}_k^{\mathsf{v}}$ are given (then $\mathcal{Q}_k^{\mathsf{w}} = \delta\,\mathbf{\Psi}^*\,\mathcal{Q}_k\,\mathbf{\Psi}$). From now on we suppose that $\mathcal{Q}_k^{\mathsf{v}} = I$.

C. The RMLE procedure

Let $L_k(\theta)$ be the likelihood function of θ for the observations $\mathsf{y}_{1:k}$. We will see that the normalized log–likelihood function $\ell_k(\theta) \stackrel{\text{def}}{=} \frac{1}{k}\,\log L_k(\theta)$ admits an incremental formulation

$$\ell_k(\theta) = \frac{1}{k}\sum_{l=1}^{k} r_l(\theta)\,.$$

Then the score function is

$$\dot\ell_k(\theta) = \frac{1}{k} \sum_{l=1}^{k} \dot r_l(\theta)$$

and the RMLE procedure is

$$\theta_k \leftarrow \theta_{k-1} + \gamma_k \times \dot r_k(\theta_{k-1})$$

where γ_k is a non–increasing sequence of positive numbers.

In § III we present the Kalman filter–based approximation of the score increment $\dot r_k(\theta)$ and in § IV its particle filter–based counterpart. This last approximation, contrary to the first one, is valid in the nonlinear/non–Gaussian case. These procedures are applied to our problem in § V.

III. Kalman RMLE for a linear system

Consider the following linear system

$$\begin{aligned} \mathsf{x}_{k+1} &= F(\theta)\,\mathsf{x}_k + G(\theta)\,\mathsf{w}_k\,, & \mathsf{w}_k &\overset{\text{iid}}{\sim} N(0,\mathcal{Q}^{\mathsf{w}}_k)\,,\\ \mathsf{y}_k &= H(\theta)\,\mathsf{x}_k + \Sigma(\theta)\,\mathsf{v}_k\,, & \mathsf{v}_k &\overset{\text{iid}}{\sim} N(0,\mathcal{Q}^{\mathsf{v}}_k)\,, \end{aligned}$$

The state process x_k takes values in $\mathbb{C}^n$, the observation process y_k in $\mathbb{C}^d$. The state initial law is $\mathsf{x}_0 \sim N(\bar{\mathsf{x}}_0, \mathcal{R}_0)$. Initial condition x_0, state noise w_k and observation noise v_k are mutually independent.

Here $\theta \in \mathbb{R}$ is an unknown *real* parameter: the derivative w.r.t. this parameter will be denoted "∂_θ" or "$\cdot$" where there is no ambiguity.

Suppose that the matrices $F(\theta) \in \mathbb{C}^{n\times n}$, $G(\theta) \in \mathbb{C}^{n\times n'}$, $H(\theta) \in \mathbb{C}^{d\times n}$ and $\Sigma(\theta) \in \mathbb{C}^{d\times d'}$ are differentiable w.r.t. θ.

For every fixed θ, the conditional laws $\operatorname{law}(\mathsf{x}_k|\mathsf{y}_{1:k-1}) = N(\hat{\mathsf{x}}^\theta_{k-}, \mathcal{R}^\theta_{k-})$ and $\operatorname{law}(\mathsf{x}_k|\mathsf{y}_{1:k}) = N(\hat{\mathsf{x}}^\theta_k, \mathcal{R}^\theta_k)$ are given recursively by the Kalman filter (see Part *b* in TABLE I).

A. Likelihood function

One expresses the law of the observations $\mathsf{y}_{1:k}$ via the innovation process

$$\hat\imath_k \overset{\text{def}}{=} \mathsf{y}_k - \mathbb{E}_\theta[\mathsf{y}_k|\mathsf{y}_{1:k-1}] = \mathsf{y}_k - H(\theta)\,\hat{\mathsf{x}}^\theta_{k-}\,.$$

Note that

$$\mathbb{P}_\theta(\mathsf{y}_{1:k} \in dy_{1:k}) = \textstyle\prod_{l=1}^k \mathbb{P}_\theta(\mathsf{y}_l \in dy_l|\mathsf{y}_{1:l-1} = y_{1:l-1})$$

and $\operatorname{law}(\mathsf{y}_k|\mathsf{y}_{1:k-1}) = N(H(\theta)\,\hat{\mathsf{x}}^\theta_{k-}, \mathcal{S}^\theta_k)$ where $\mathcal{S}^\theta_k$ is the covariance of the innovation process (see Part *b* in TABLE I). We get

$$\mathbb{P}_\theta(\mathsf{y}_{1:k} \in dy_{1:k}) = \textstyle\prod_{l=1}^k g^\theta_l(y_l)\,dy_l$$

where $g^\theta_k(y)$ is the p.d.f. of the $N(H(\theta)\,\hat{\mathsf{x}}^\theta_{k-}, \mathcal{S}^\theta_k)$ law. This means that

$$r_l(\theta) \overset{\text{def}}{=} \log g^\theta_l(\mathsf{y}_l)\,.$$

B. Score function

In order to calculate the score increment $\dot r_k(\theta)$, one sets an auxiliary result. Consider the p.d.f. $q^\theta(x)$ of the normal law $N(\mu(\theta),\mathcal{R}(\theta))$ on $\mathbb{C}^n$ whose mean $\mu(\theta)$ and covariance matrix $\mathcal{R}(\theta) > 0$ are differentiable w.r.t. a scalar parameter $\theta \in \mathbb{R}$. Then the two classical identities

$$\begin{aligned} \partial_\theta \log|\mathcal{R}(\theta)| &= \tfrac{\partial_\theta|\mathcal{R}(\theta)|}{|\mathcal{R}(\theta)|} = \operatorname{trace}\big\{[\mathcal{R}(\theta)]^{-1}\,\dot{\mathcal{R}}(\theta)\big\}\,,\\ \partial_\theta[\mathcal{R}(\theta)]^{-1} &= -[\mathcal{R}(\theta)]^{-1}\,\dot{\mathcal{R}}(\theta)\,[\mathcal{R}(\theta)]^{-1} \end{aligned}$$

a – initialization

$$\theta \text{ (initial guess)} \qquad \hat{\mathsf{x}}^\theta_0 = \bar{\mathsf{x}}_0 \qquad \mathcal{R}^\theta_0 = \mathcal{R}_0$$

b – Kalman filter

$$\begin{aligned} \hat{\mathsf{x}}^\theta_{k-} &= F(\theta)\,\hat{\mathsf{x}}^\theta_{k-1}\\ \mathcal{R}^\theta_{k-} &= F(\theta)\,\mathcal{R}^\theta_{k-1}\,F(\theta)^* + G(\theta)\,\mathcal{Q}^{\mathsf{w}}_{k-1}\,G(\theta)^*\\ \hat\imath^\theta_k &= \mathsf{y}_k - H(\theta)\,\hat{\mathsf{x}}^\theta_{k-}\\ \mathcal{S}^\theta_k &= H(\theta)\,\mathcal{R}^\theta_{k-}\,H(\theta)^* + \Sigma(\theta)\,\mathcal{Q}^{\mathsf{v}}_k\,\Sigma(\theta)^*\\ K^\theta_k &= \mathcal{R}^\theta_{k-}\,H(\theta)^*\,[S^\theta_k]^{-1}\\ \hat{\mathsf{x}}^\theta_k &= \hat{\mathsf{x}}^\theta_{k-} + K^\theta_k\,\hat\imath^\theta_k\\ \mathcal{R}^\theta_k &= \big\{I - K^\theta_k\,H(\theta)\big\}\,\mathcal{R}^\theta_{k-} \end{aligned}$$

c – tangent Kalman filter

$$\begin{aligned} \dot{\hat{\mathsf{x}}}^\theta_{k-} &= F(\theta)\,\dot{\hat{\mathsf{x}}}^\theta_{k-1} + \dot F(\theta)\,\hat{\mathsf{x}}^\theta_{k-1}\\ \dot{\mathcal{R}}^\theta_{k-} &= F(\theta)\,\dot{\mathcal{R}}^\theta_{k-1}\,F(\theta)^*\\ &\quad + \dot F(\theta)\,\mathcal{R}^\theta_{k-1}\,F(\theta)^* + F(\theta)\,\mathcal{R}^\theta_{k-1}\,\dot F(\theta)^*\\ &\quad + \dot G(\theta)\,\mathcal{Q}^{\mathsf{w}}_{k-1}\,G(\theta)^* + G(\theta)\,\mathcal{Q}^{\mathsf{w}}_{k-1}\,\dot G(\theta)^*\\ \dot{\hat\imath}^\theta_k &= -\dot H(\theta)\,\hat{\mathsf{x}}^\theta_{k-} - H(\theta)\,\dot{\hat{\mathsf{x}}}^\theta_{k-}\\ \dot{\mathcal{S}}^\theta_k &= H(\theta)\,\dot{\mathcal{R}}^\theta_{k-}\,H(\theta)^*\\ &\quad + \dot H(\theta)\,\mathcal{R}^\theta_{k-}\,H(\theta)^* + H(\theta)\,\mathcal{R}^\theta_{k-}\,\dot H(\theta)^*\\ &\quad + \dot\Sigma(\theta)\,\mathcal{Q}^{\mathsf{v}}_k\,\Sigma(\theta)^* + \Sigma(\theta)\,\mathcal{Q}^{\mathsf{v}}_k\,\dot\Sigma(\theta)^*\\ \dot K^\theta_k &= \dot{\mathcal{R}}^\theta_{k-}\,H(\theta)^*\,[S^\theta_k]^{-1} + \mathcal{R}^\theta_{k-}\,\dot H(\theta)^*\,[S^\theta_k]^{-1}\\ &\quad - \mathcal{R}^\theta_{k-}\,H(\theta)^*\,[S^\theta_k]^{-1}\,\dot S^\theta_k\,[S^\theta_k]^{-1}\\ \dot{\hat{\mathsf{x}}}^\theta_k &= \dot{\hat{\mathsf{x}}}^\theta_{k-} + \dot K^\theta_k\,\hat\imath^\theta_k + K^\theta_k\,\dot{\hat\imath}^\theta_k\\ \dot{\mathcal{R}}^\theta_k &= \big\{I - K^\theta_k\,H(\theta)\big\}\,\dot{\mathcal{R}}^\theta_{k-}\\ &\quad - \big\{\dot K^\theta_k\,H(\theta) + K^\theta_k\,\dot H(\theta)\big\}\,\mathcal{R}^\theta_{k-} \end{aligned}$$

d – score increment

$$\begin{aligned} \tilde\imath^\theta_k &= [\mathcal{S}^\theta_k]^{-1}\,\hat\imath^\theta_k\\ \dot r_k(\theta) &= -\tfrac{1}{2}\operatorname{trace}\big\{[\mathcal{S}^\theta_k]^{-1}\,\dot{\mathcal{S}}^\theta_k\big\} - \tfrac{1}{2}\,[\tilde\imath^\theta_k]^*\,\dot{\mathcal{S}}^\theta_k\,\tilde\imath^\theta_k\\ &\quad + \Re\Big\{[\tilde\imath^\theta_k]^*\,\big\{H(\theta)\,\dot{\hat{\mathsf{x}}}^\theta_{k-} + \dot H(\theta)\,\hat{\mathsf{x}}^\theta_{k-}\big\}\Big\} \end{aligned}$$

e – parameters update

$$\theta \leftarrow \theta + \gamma_k \times \dot r_k(\theta)$$

TABLE I

Kalman recursive maximum likelihood procedure.

applied to $\log q^\theta(x)$ give

$$\partial_\theta \log q^\theta(x) = -\tfrac{1}{2}\, \text{trace}\big\{[\mathcal{R}(\theta)]^{-1}\, \dot{\mathcal{R}}(\theta)\big\}$$
$$+ \Re\Big\{ \big\{x - \mu(\theta)\big\}^* \, [\mathcal{R}(\theta)]^{-1}\, \dot{\mu}(\theta)\Big\}$$
$$- \tfrac{1}{2}\, \big\{x - \mu(\theta)\big\}^* \, [\mathcal{R}(\theta)]^{-1}\, \dot{\mathcal{R}}(\theta)\, [\mathcal{R}(\theta)]^{-1}\, \big\{x - \mu(\theta)\big\}\,.$$

From this result the score increment is

$$\dot{r}_k(\theta) = -\tfrac{1}{2}\, \text{trace}\big\{[\mathcal{S}_k^\theta]^{-1}\, \dot{\mathcal{S}}_k^\theta\big\}$$
$$+ \Re\Big\{ \big\{\mathsf{y}_k - H(\theta)\, \hat{\mathsf{x}}_{k^-}^\theta\big\}^* \, [\mathcal{S}_k^\theta]^{-1}$$
$$\times \big\{H(\theta)\, \dot{\hat{\mathsf{x}}}_{k^-}^\theta + \dot{H}(\theta)\, \hat{\mathsf{x}}_{k^-}^\theta\big\}\Big\}$$
$$- \tfrac{1}{2}\, \big\{\mathsf{y}_k - H(\theta)\, \hat{\mathsf{x}}_{k^-}^\theta\big\}^* \, [\mathcal{S}_k^\theta]^{-1}\, \dot{\mathcal{S}}_k^\theta\, [\mathcal{S}_k^\theta]^{-1}$$
$$\times \big\{\mathsf{y}_k - H(\theta)\, \hat{\mathsf{x}}_{k^-}^\theta\big\}\,.$$

TABLE I page 3 presents the complete algorithm.

IV. Particle filter RMLE for a nonlinear system

A. The problem

Consider a state/observation process whose law depends on an unknown parameter $\theta \in \mathbb{R}$. The state process $\mathsf{x} = \{\mathsf{x}_k\}_{k\geq 0}$ takes values in $\mathbb{R}^n$, it is Markovian with transition kernel Q_k^θ and initial probability law μ_0,

$$Q_k^\theta(dx'|x) \stackrel{\text{def}}{=} \mathbb{P}_\theta(\mathsf{x}_{k+1} \in dx' | \mathsf{x}_k = x)\,, \tag{10}$$
$$\mu_0(dx) \stackrel{\text{def}}{=} \mathbb{P}_\theta(\mathsf{x}_0 \in dx)\,. \tag{11}$$

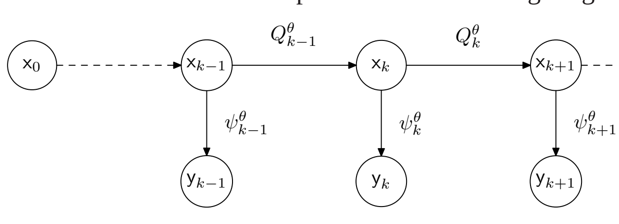
This process describes the evolution of a non observed system. The observation process $\mathsf{y} = \{\mathsf{y}_k\}_{k\geq 1}$ takes values in $\mathbb{R}^d$. We suppose that (*i*) conditionally to the state process, the observations y_k are independent, and (*ii*) the observation y_k depends only on x_k (y_k is the observation of x_k), i.e.

$$\mathbb{P}_\theta(\mathsf{y}_{1:k} \in dy_{1:k} | \mathsf{x}_{0:k} = x_{0:k}) = \prod_{l=1}^k \mathbb{P}_\theta(\mathsf{y}_l \in dy_l | \mathsf{x}_l = x_l)\,. \tag{12}$$

The law of the state/observation process (x, y) is now completely specified. We assume moreover that the conditional law of y_k given x_k admits a density w.r.t. the Lebesgue measure:

$$\psi_k^\theta(y|x)\, dy \stackrel{\text{def}}{=} \mathbb{P}_\theta(\mathsf{y}_k \in dy | \mathsf{x}_k = x)\,. \tag{13}$$

Then the law of the process (x, y) can be expressed explicitly according to the three basic terms (10), (11) and (13), see § IV-C. This situation corresponds to the following diagram



The system depends on the parameter θ through the kernel Q_k^θ and the local likelihood function ψ_k^θ. For simplicity we suppose that the initial law does not depend on θ.

Here, like in Doucet & Tadic [3], we consider the case where the Markov kernel Q_k^θ admits a density w.r.t. the Lebesgue measure

$$Q_k^\theta(dx'|x) = q_k^\theta(x'|x)\, dx\,.$$

The case of a Markov kernel without density was treated in Guyader *et al* [4] and Fichou *et al* [6]. This hypothesis is not required to establish the equations of the nonlinear optimal filter but it simplifies the derivation of the tangent filter.

Example: Consider a state–space model

$$\mathsf{x}_{k+1} = f_k^\theta(\mathsf{x}_k) + \sigma_{\mathsf{w}}^\theta\, \mathsf{w}_k\,, \qquad \mathsf{w}_k \sim N(0, \mathcal{Q}_k^{\mathsf{w}})\,,$$
$$\mathsf{y}_k = h_k^\theta(\mathsf{x}_k) + \sigma_{\mathsf{v}}^\theta\, \mathsf{v}_k\,, \qquad \mathsf{v}_k \sim N(0, \mathcal{Q}_k^{\mathsf{v}})\,,$$

where x_k and y_k take with values in $\mathbb{R}^n$ and $\mathbb{R}^d$; w_k, v_k, x_0 are independent; $\mathsf{x}_0 \sim \mu_0$; $\sigma_{\mathsf{w}}^\theta$ and $\sigma_{\mathsf{v}}^\theta$ are scalar positive numbers. Here $Q_k^\theta(dx'|x) = q_k^\theta(x'|x)\, dx'$ and

$$q_k^\theta(x'|x) = \big\{(2\,\pi)^n\, (\sigma_{\mathsf{w}}^\theta)^{2n}\, |\mathcal{Q}_k^{\mathsf{w}}|\big\}^{-1/2}$$
$$\times \exp\Big\{ -\tfrac{1}{2\,(\sigma_{\mathsf{w}}^\theta)^2} [x' - f_k^\theta(x)]^*\, [\mathcal{Q}_k^{\mathsf{w}}]^{-1}\, [x' - f_k^\theta(x)]\Big\}\,,$$
$$\psi_k^\theta(y|x) = \big\{(2\,\pi)^d\, (\sigma_{\mathsf{v}}^\theta)^{2d}\, |\mathcal{Q}_k^{\mathsf{v}}|\big\}^{-\frac{1}{2}}$$
$$\times \exp\Big\{ -\tfrac{1}{2\,(\sigma_{\mathsf{v}}^\theta)^2} [y - h_k^\theta(x)]^*\, [\mathcal{Q}_k^{\mathsf{v}}]^{-1}\, [y - h_k^\theta(x)]\Big\}\,.$$

B. Nonlinear filter

Define the nonlinear filter and the predicted nonlinear filter

$$\pi_k^\theta(dx|y_{1:k}) \stackrel{\text{def}}{=} \mathbb{P}_\theta(\mathsf{x}_k \in dx | \mathsf{y}_{1:k} = y_{1:k})\,,$$
$$\pi_{k^-}^\theta(dx|y_{1:k-1}) \stackrel{\text{def}}{=} \mathbb{P}_\theta(\mathsf{x}_k \in dx | \mathsf{y}_{1:k-1} = y_{1:k-1})\,.$$

We also use the notation $\pi_k^\theta(dx) \stackrel{\text{def}}{=} \pi_k^\theta(dx|\mathsf{y}_{1:k})$, $\pi_{k^-}^\theta(dx) \stackrel{\text{def}}{=} \pi_{k^-}^\theta(dx|\mathsf{y}_{1:k-1})$ and $\psi_k^\theta(x) \stackrel{\text{def}}{=} \psi_k^\theta(\mathsf{y}_k|x)$. These conditional densities can be recursively obtained through the classical two steps procedure:

$$\pi_{k-1}^\theta \xrightarrow{\text{prediction}} \pi_{k^-}^\theta = \pi_{k-1}^\theta\, Q_{k-1}^\theta \xrightarrow{\text{correction}} \pi_k^\theta = \Psi_k^\theta[\pi_{k^-}^\theta] \tag{14}$$

where the prediction (linear) operator Q_{k-1}^θ and the correction (nonlinear) operator Ψ_k^θ, which act on the space of probability measures, are defined by

$$\pi Q_{k-1}^\theta(dx') \stackrel{\text{def}}{=} \textstyle\int Q_{k-1}^\theta(dx'|x)\; \pi(dx)\,, \tag{15}$$
$$\Psi_k^\theta[\pi](dx) \stackrel{\text{def}}{=} \frac{\psi_k^\theta(x)\; \pi(dx)}{\langle \pi, \psi_k^\theta\rangle} \tag{16}$$

where

$$\langle \pi\,,\, \psi\rangle \stackrel{\text{def}}{=} \textstyle\int_{\mathbb{R}^n} \psi(x)\, \pi(dx)\,.$$

Q_k^θ is the transition kernel of the Markov chain x_k and the first step in (14) is the Chapman–Kolmogorov equation. The second step in (14) is a Bayes formula. The initial condition in (14) is $\pi_0^\theta = \mu_0$.

C. Likelihood and score functions

According to the previous section, the joint law of the state and observation processes is

$$\begin{aligned}&\mathbb{P}_\theta(\mathsf{x}_{0:k} \in dx_{0:k}, \mathsf{y}_{1:k} \in dy_{1:k})\\ &\quad = \mu_0(dx_0) \prod_{l=1}^k \left\{\psi_\ell^\theta(y_l|x_l)\; Q_{l-1}^\theta(dx_l|x_{l-1})\right\} dy_{1:k}\,.\end{aligned}$$

This proves that this statistical model is dominated and

$$L_k(\theta) = \int_{x_{0:k}} \mu_0(dx_0) \prod_{l=1}^k \left\{\psi_k^\theta(x_l)\; Q_{l-1}^\theta(dx_l|x_{l-1})\right\}.$$

Note that

$$\begin{aligned}&\mathbb{P}^\theta(\mathsf{y}_{1:k} \in dy_{1:k})\\ &\quad = \mathbb{P}^\theta(\mathsf{y}_1 \in dy_1)\; \prod_{l=2}^k \mathbb{P}^\theta(\mathsf{y}_l \in dy_l|\mathsf{y}_{1:l-1} = y_{1:l-1})\\ &= \prod_{l=1}^k \int_{x_l} \mathbb{P}^\theta(\mathsf{y}_l \in dy_l|\mathsf{x}_l = x_l)\; \pi_{l^-}^\theta(dx_l|y_{1:l-1})\\ &= \prod_{l=1}^k \int_{x_l} \psi_l^\theta(y_l|x_l)\; \pi_{l^-}^\theta(dx_l|y_{1:l-1})\; dy_{1:k}\end{aligned}$$

which leads to the other formulation

$$L_k(\theta) = \prod_{l=1}^k \langle \pi_{l^-}^\theta\,,\, \psi_l^\theta\rangle$$

so

$$r_l(\theta) \stackrel{\text{def}}{=} \log\langle \pi_{l^-}^\theta\,,\, \psi_l^\theta\rangle$$

and the score increment is

$$\begin{aligned}\dot r_k(\theta) &= \partial_\theta \log\langle \pi_{k^-}^\theta\,,\, \psi_k^\theta\rangle\\ &= \frac{\langle \dot\pi_{k^-}^\theta\,,\, \psi_k^\theta\rangle + \langle \pi_{k^-}^\theta\,,\, [\partial_\theta \log\psi_k^\theta]\,\psi_k^\theta\rangle}{\langle \pi_{k^-}^\theta\,,\, \psi_k^\theta\rangle}\,. \qquad (17)\end{aligned}$$

D. Tangent filter

The definition of derivative $\dot\pi_{k^-}^\theta$ of the nonlinear (prediction) filter requires some attention. Let $\mathcal{M}(\mathbb{R}^n)$ the set of finite (signed) measures on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, $\mathcal{M}_1^+(\mathbb{R}^n)$ that of probability measures on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, and $\mathcal{M}^0(\mathbb{R}^n)$ that of null mass. A measure μ on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ is said finite if $\mu^+(\mathbb{R}^n) + \mu^-(\mathbb{R}^n) < \infty$ where $\mu^\pm$ is the Hahn–Jordan decomposition of μ. Here $\pi_{k^-}^\theta \in \mathcal{M}_1(\mathbb{R}^n)$ then $\dot\pi_{k^-}^\theta \in \mathcal{M}^0(\mathbb{R}^n)$ and $\dot\pi_{k^-}^\theta \ll \pi_{k^-}^\theta$. See Heidergott & Vázquez–Abad [10] for details.

We now establish a recursive formulation for $\dot\pi_{k^-}^\theta$ and $\dot\pi_k^\theta$.

1) Prediction step: The derivative $\dot Q_{k-1}^\theta$ of the Markov kernel Q_{k-1}^θ w.r.t. the parameter θ is

$$\begin{aligned}\dot Q_{k-1}^\theta(dx'|x) &= \partial_\theta q_{k-1}^\theta(x'|x)\, dx'\\ &= \left\{\partial_\theta \log q_{k-1}^\theta(x'|x)\right\} q_{k-1}^\theta(x'|x)\, dx'\,.\\ &= \left[\partial_\theta \log q_{k-1}^\theta(x'|x)\right] Q_{k-1}^\theta(dx'|x)\,. \qquad (18)\end{aligned}$$

$\dot Q_{k-1}^\theta$ is a transition kernel on $\mathcal{M}^0(\mathbb{R}^n)$. Then, the derivative of the nonlinear filter w.r.t. the parameter θ is

$$\dot\pi_{k^-}^\theta = \partial_\theta\{\pi_{k-1}^\theta\, Q_{k-1}^\theta\} = \dot\pi_{k-1}^\theta\, Q_{k-1}^\theta + \pi_{k-1}^\theta\, \dot Q_{k-1}^\theta\,. \quad (19)$$

2) Correction step: One introduces $\mathsf{D}\Psi_k^\theta[\pi]\,\nu \in \mathcal{M}^0(\mathbb{R}^n)$ the derivative of the operator $\pi \mapsto \Psi_k^\theta[\pi]$ at the point $\pi \in \mathcal{M}_1^+(\mathbb{R}^n)$ in the direction $\nu \in \mathcal{M}^0(\mathbb{R}^n)$

$$\mathsf{D}\Psi_k^\theta[\pi]\,\nu \stackrel{\text{def}}{=} \frac{\psi_k^\theta\,\nu}{\langle \pi, \psi_k^\theta\rangle} - \frac{\langle \nu, \psi_k^\theta\rangle}{\langle \pi, \psi_k^\theta\rangle}\,\frac{\psi_k^\theta\,\pi}{\langle \pi, \psi_k^\theta\rangle}\,.$$

If $\nu \ll \pi$ and $\nu(dx) = \varrho(x)\,\pi(dx)$ then

$$\mathsf{D}\Psi_k^\theta[\pi]\,\nu = \left\{\varrho - \langle \Psi_k^\theta[\pi], \varrho\rangle\right\} \Psi_k^\theta[\pi]\,.$$

Moreover, one has

$$\begin{aligned}\partial_\theta \Psi_k^\theta[\pi] &= \left\{\partial_\theta \log\psi_k^\theta - \langle \Psi_k^\theta[\pi], \partial_\theta \log\psi_k^\theta\rangle\right\} \Psi_k^\theta[\pi]\\ &= \mathsf{D}\Psi_k^\theta[\pi]([\partial_\theta \log\psi_k^\theta]\,\pi)\,.\end{aligned}$$

Finally

$$\begin{aligned}\dot\pi_k^\theta &= \partial_\theta\left\{\Psi_k^\theta[\pi_{k^-}^\theta]\right\}\\ &= \mathsf{D}\Psi_k^\theta[\pi_{k^-}^\theta](\dot\pi_{k^-}^\theta) + \left\{\partial_\theta \Psi_k^\theta[\pi]\right\}\Big|_{\pi=\pi_{k^-}^\theta}\\ &= \mathsf{D}\Psi_k^\theta[\pi_{k^-}^\theta](\dot\pi_{k^-}^\theta + [\partial_\theta \log\psi_k^\theta]\,\pi_{k^-}^\theta)\,.\end{aligned}$$

Hence we can prove recursively that the tangent filter is absolutely continuous w.r.t. the nonlinear filter, i.e. $\dot\pi_{k^-}^\theta \ll \pi_{k^-}^\theta$ and $\dot\pi_k^\theta \ll \pi_k^\theta$. Let

$$\varrho_{k^-}^\theta(x) \stackrel{\text{def}}{=} \frac{d\dot\pi_{k^-}^\theta}{d\pi_{k^-}^\theta}(x)\,, \qquad \varrho_k^\theta(x) \stackrel{\text{def}}{=} \frac{d\dot\pi_k^\theta}{d\pi_k^\theta}(x)\,.$$

This leads to

$$\dot\pi_k^\theta = \left\{\varrho_{k^-}^\theta + \partial_\theta \log\psi_k^\theta - \langle \pi_k^\theta, \varrho_{k^-}^\theta + \partial_\theta \log\psi_k^\theta\rangle\right\} \pi_k^\theta\,. \qquad (20)$$

3) Score increment: Expression (17) becomes

$$\dot r_k(\theta) = \langle \pi_k^\theta, \varrho_{k^-}^\theta + \partial_\theta \log\psi_k^\theta\rangle \qquad (21)$$

which is exactly the centering term in (20).

The joint nonlinear/tangent filters is summarized in TABLE II.

E. Particle approximation

We describe the simple "bootstrap" particle approximation. Suppose that at time $k-1$ we have particle approximation of the nonlinear and tangent filters

$$\begin{aligned}\pi_{k-1}^\theta &\simeq \pi_{k-1}^N \stackrel{\text{def}}{=} \tfrac{1}{N}\textstyle\sum_{i=1}^N \delta_{\xi_{k-1}^i}\,,\\ \dot\pi_{k-1}^\theta &\simeq \dot\pi_{k-1}^N \stackrel{\text{def}}{=} \tfrac{1}{N}\textstyle\sum_{i=1}^N \rho_{k-1}^i\, \delta_{\xi_{k-1}^i}\,.\end{aligned}$$

Note that $\sum_{i=1}^N \rho_{k-1}^i = 0$, i.e. $\dot\pi_{k-1}^N \in \mathcal{M}^0(\mathbb{R}^n)$. The idea of this approximation is to assure that $\dot\pi_{k-1}^N \ll \pi_{k-1}^N$, indeed

$$\varrho_{k-1}^N(\xi_{k-1}^i) \stackrel{\text{def}}{=} \frac{d\dot\pi_k^\theta}{d\pi_{k-1}^N}(\xi_{k-1}^i) = \frac{\sum_{i':\xi_{k-1}^{i'}=\xi_{k-1}^i} \rho_{k-1}^i}{\sum_{i':\xi_{k-1}^{i'}=\xi_{k-1}^i} 1} \quad (22)$$

for $i = 1\cdots N$ and $\varrho_{k-1}^N(x) = 0$ if $x \notin \{\xi_{k-1}^i; i = 1\cdots N\}$.

a – initialization
$\pi_0^\theta = \mu_0 \qquad \dot\pi_0^\theta = 0$
b – prediction [see (15) and (18)]
$\pi_{k-}^\theta = \pi_{k-1}^\theta\, Q_{k-1}^\theta$ $\dot\pi_{k-}^\theta = \dot\pi_{k-1}^\theta\, Q_{k-1}^\theta + \pi_{k-1}^\theta\, \dot Q_{k-1}^\theta$
c – correction [see (16) and (20)]
$\pi_k^\theta = \Psi_k^\theta[\pi_{k-}^\theta]$ $\dot\pi_k^\theta = \Big\{\varrho_{k-}^\theta + \partial_\theta \log\psi_k^\theta - \langle \pi_k^\theta, \varrho_{k-}^\theta + \partial_\theta\log\psi_k^\theta\rangle\Big\}\,\pi_k^\theta$ here $\varrho_{k-}^\theta = d\dot\pi_{k-}^\theta / d\pi_{k-}^\theta$
d – score increment [see (21)]
$\dot r_k(\theta) = \langle \pi_k^\theta, \varrho_{k-}^\theta + \partial_\theta \log\psi_k^\theta\rangle$

TABLE II

The joint nonlinear/tangent filters.

1) Prediction/sampling step: From (15),

$$\pi_{k-1}^N\, Q_{k-1}^\theta(dx') = \tfrac{1}{N}\textstyle\sum_{i=1}^N Q_{k-1}^\theta(dx'|\xi_{k-1}^i)\,.$$

We use the approximation

$$Q_{k-1}^\theta(dx'|\xi_{k-1}^i) \simeq \delta_{\xi_{k-}^i}(dx')$$

where $\xi_{k-}^i \sim Q_{k-1}^\theta(dx'|\xi_{k-1}^i)$ (independently). Hence

$$\pi_{k-}^N = \tfrac{1}{N}\textstyle\sum_{i=1}^N \delta_{\xi_{k-}^i} \text{ where } \xi_{k-}^i \sim Q_{k-1}^\theta(dx'|\xi_{k-1}^i)\,. \quad (23)$$

From the tangent filter prediction (19) and (18)

$$\begin{aligned}
&\dot\pi_{k-1}^N\, Q_{k-1}^\theta(dx') + \pi_{k-1}^N\, \dot Q_{k-1}^\theta(dx') \\
&\quad = \textstyle\int Q_{k-1}^\theta(dx'|x)\,\dot\pi_{k-1}^N(dx) \\
&\qquad\qquad + \textstyle\int Q_{k-1}^\theta(dx'|x)\,[\partial_\theta \log q_{k-1}^\theta(x'|x)]\;\pi_{k-1}^N(dx) \\
&\quad = \tfrac{1}{N}\textstyle\sum_{i=1}^N \big\{\rho_{k-1}^i + \partial_\theta \log q_{k-1}^\theta(x'|\xi_{k-1}^i)\big\} \\
&\qquad\qquad \times Q_{k-1}^\theta(dx'|\xi_{k-1}^i)\,.
\end{aligned}$$

Again let $Q_{k-1}^\theta(dx'|\xi_{k-1}^i) \simeq \delta_{\xi_{k-}^i}(dx')$ so that

$$\dot\pi_{k-}^N = \tfrac{1}{N}\textstyle\sum_{i=1}^N \rho_{k-}^i\,\delta_{\xi_{k-}^i}$$

with

$$\rho_{k-}^i \stackrel{\text{def}}{=} \rho_{k-1}^i + \partial_\theta \log q_{k-1}^\theta(\xi_{k-}^i|\xi_{k-1}^i)\,.$$

This approximation $\dot\pi_{k-}^N$ is not of null mass, it will be "centered" in the correction step.

Again $\varrho_{k-}^N(x)$ can be computed like in (22), but almost surely the particle positions ξ_{k-}^i are all distinct, so we have

$$\varrho_{k-}^N(\xi_{k-}^i) = \rho_{k-}^i\,, \quad i = 1\cdots N\,.$$

2) Correction/resampling step: Plugging the approximation (23) in (16) gives exactly

$$\Psi_k^\theta[\pi_{k-}^N] = \textstyle\sum_{i=1}^N \omega_k^i\,\delta_{\xi_{k-}^i}$$

where

$$\omega_k^i \stackrel{\text{def}}{=} \Psi_k^\theta(\xi_{k-}^i) / \textstyle\sum_{i'=1}^N \Psi_k^\theta(\xi_{k-}^{i'})\,.$$

The resampling step is the following: we multiply/discard particles $\{\xi_{k-}^i\}_{i=1:N}$ according to the high/low weights $\{\omega_k^i\}_{i=1:N}$, i.e. $\xi_k^i = \xi_{k-}^{\mathbf{s}[i]}$ where $\mathbf{s}[i]$ is the resampling mechanism associated with the weights $\{\omega_k^i\}_{i=1:N}$. The updated particle approximation is then

$$\pi_k^N = \tfrac{1}{N}\textstyle\sum_{i=1}^N \delta_{\xi_k^i} \quad \text{with} \quad \xi_k^i = \xi_{k-}^{\mathbf{s}[i]} \qquad (24)$$

and $\mathbf{s}$ is the resampling scheme associated with $\{\omega_k^i\}_{i=1:N}$.

Substituting π_k^θ in (20) by its approximation (24) gives exactly

$$\dot\pi_k^N = \tfrac{1}{N}\textstyle\sum_{i=1}^N \rho_k^i\,\delta_{\xi_k^i}$$

where ([2])

$$\begin{aligned}
\rho_k^i &= \varrho_{k-}^\theta(\xi_k^i) + \partial_\theta \log\psi_k^\theta(\xi_k^i) - \text{centering term} \\
&= \rho_{k-}^{\mathbf{s}[i]} + \partial_\theta \log\psi_k^\theta(\xi_{k-}^{\mathbf{s}[i]}) - \text{centering term.} \qquad (25)
\end{aligned}$$

This last centering operation ensures that $\dot\pi_k^N \in \mathcal{M}^0(\mathbb{R}^n)$.

3) Score increment: Approximation (24) in (21) leads to

$$\begin{aligned}
r_k^N(\theta) &= \langle \pi_k^N, \varrho_{k-}^\theta + \partial_\theta \log\psi_k^\theta\rangle \\
&= \tfrac{1}{N}\textstyle\sum_{i=1}^N \big\{\rho_{k-}^{\mathbf{s}[i]} + \partial_\theta \log\psi_k^\theta(\xi_{k-}^{\mathbf{s}[i]})\big\}
\end{aligned}$$

which is exactly, like noticed in § IV-D.3, the centering term of (25).

The joint particle approximation of the nonlinear/tangent filters is summarized in TABLE III.

V. APPLICATION

In (8)–(9) there are two alternate parameterizations. The first one is in terms of real/imaginary part of the λ's (7)

$$\theta \stackrel{\text{def}}{=} (\alpha_{1:n}, \beta_{1:n}, \sigma, \nu) \in \mathbb{R}^{2n} \times \mathbb{R}_+^2 \qquad (26)$$

where $\alpha_p \stackrel{\text{def}}{=} \Re(\lambda_p)$ and $\beta_p \stackrel{\text{def}}{=} \Im(\lambda_p)$ for $p = 1\cdots n$. The second one is terms of frequency/damping coefficients (4)

$$\theta \stackrel{\text{def}}{=} (\mathbf{f}_{1:n}, \mathbf{d}_{1:n}, \sigma, \nu) \in \mathbb{R}_+^n \times (0,1)^n \times \mathbb{R}_+^2\,. \qquad (27)$$

If the behavior of the filter is quite equivalent in both parameterizations, the second is much simpler to use for the tuning of the parameters of the RMLE procedure.

Kalman filter formulation

Practical implementation of the algorithm describes in § III requires some adaptations. To prevent the degeneracy of the innovation covariance matrix it is necessary to reinforce the diagonal terms if $\mathcal{S}_k^\theta$ in Part *b* of TABLE I.

[2] "$\alpha_i = \beta_i$ – centering term" means that $\alpha_i = \beta_i - \sum_{i'=1}^N \beta_{i'}$

a – initialization

$$\begin{aligned}&\theta \qquad \text{(initial guess)}\\ &\xi_0^i \sim \mu_0(dx) \quad \text{(independently)}\\ &\rho_0^i \leftarrow 0\end{aligned}$$

b – mutation

$$\xi_{k-}^i \sim Q_{k-1}^\theta(dx'|\xi_{k-1}^i) \quad \text{(independently)}$$

b – weights evaluation

$$\begin{aligned}&\omega_k^i \leftarrow \psi_k^\theta(\xi_{k-}^i) / \textstyle\sum_{i'=1}^N \psi_k^\theta(\xi_{k-}^{i'})\\ &\tilde\rho_k^i \leftarrow \rho_{k-1}^i + \partial_\theta \log q_{k-1}^\theta(\xi_{k-}^i|\xi_{k-1}^i) + \partial_\theta \log \psi_k^\theta(\xi_{k-}^i)\end{aligned}$$

b – selection

$$\begin{aligned}&\mathbf{s} \quad \text{sampling mechanism based on } \{\omega_k^i\}_{i=1:N}\\ &\xi_k^i \leftarrow \xi_{k-}^{\mathbf{s}[i]} \quad \text{and} \quad \tilde\rho_k^i \leftarrow \tilde\rho_k^{\mathbf{s}[i]}\end{aligned}$$

c – score increment

$$\dot r_k^N(\theta) \leftarrow \tfrac{1}{N} \textstyle\sum_{i=1}^N \tilde\rho_k^i \quad \text{and} \quad \rho_k^i \leftarrow \tilde\rho_k^i + \dot r_k^N(\theta)$$

d – RMLE iteration

$$\theta \leftarrow \theta + \gamma_k \times \dot r_k^N(\theta)$$

TABLE III

The joint particle approximation of the nonlinear/tangent filters and the RMLE iteration.

Particle filter formulation

Introduce the function

$$\Upsilon_k^\theta(x,x') \stackrel{\text{def}}{=} \log q_k^\theta(x'|x) + \log \psi_k^\theta(x)$$

which appears in Part *b* of TABLE III. We compute the derivative of Υ_k^θ w.r.t. each component θ_p of the parameter for $p = 1 \cdots 2n+2$.

$$\begin{aligned}\Upsilon_k^\theta(x,x') = \text{Const} &- \tfrac{n}{2}\,\log\sigma\\ &- \tfrac{1}{2\,\sigma^2}\,[x'-\mathbf{\Lambda}\,x]^*\,[\mathcal{Q}_k^{\mathsf{w}}]^{-1}\,[x'-\mathbf{\Lambda}\,x]\\ &- \tfrac{d}{2}\,\log\nu - \tfrac{1}{2\,\nu^2}\,|\mathsf{y}_k - \Re(\mathbf{\Psi}\,x)|^2\,.\end{aligned}$$

Hence, for $p = 1:2n$

$$\begin{aligned}\partial_{\theta_p}\Upsilon_k^\theta(x,x') = {}&\tfrac{1}{2\,\sigma^2}\,[\partial_{\theta_p}\mathbf{\Lambda}\,x]^*\,[\mathcal{Q}_k^{\mathsf{w}}]^{-1}\,[x'-\mathbf{\Lambda}\,x]\\ &+ \tfrac{1}{2\,\sigma^2}\,[x'-\mathbf{\Lambda}\,x]^*\,[\mathcal{Q}_k^{\mathsf{w}}]^{-1}\,[\partial_{\theta_p}\mathbf{\Lambda}\,x]\end{aligned}$$

so that

$$\begin{aligned}\partial_{\theta_p}\Upsilon_k^\theta(x,x') &= \tfrac{1}{\sigma^2}\,\Re\Big\{[\partial_{\theta_p}\mathbf{\Lambda}\,x]^*\,[\mathcal{Q}_k^{\mathsf{w}}]^{-1}\,[x'-\mathbf{\Lambda}\,x]\Big\}\\ \partial_\sigma\Upsilon_k^\theta(x,x') &= -\tfrac{n}{2\,\sigma} + \tfrac{1}{\sigma^3}\,[x'-\mathbf{\Lambda}\,x]^*\,[\mathcal{Q}_k^{\mathsf{w}}]^{-1}\,[x'-\mathbf{\Lambda}\,x]\,,\\ \partial_\nu\Upsilon_k^\theta(x,x') &= -\tfrac{d}{2\,\nu} + \tfrac{1}{\nu^3}\,|\mathsf{y}_k - \Re(\mathbf{\Psi}\,x)|^2\,.\end{aligned}$$

RMLE implementation

For each component θ_p of the parameter, the RMLE iteration used in practice is

$$\theta_p \leftarrow \theta_p + \Big\{\frac{\gamma}{k} + \gamma^{\min}\Big\} \times \partial_{\theta_p} r_k(\theta)\big|_{[-r_p^{\max}, r_p^{\max}]}$$

where

$$\partial_{\theta_p} r_k(\theta)\big|_{[-r_p^{\max}, r_p^{\max}]} \stackrel{\text{def}}{=} \times\{\partial_{\theta_p} r_k(\theta) \wedge r_p^{\max}\} \vee (-r_p^{\max})\,.$$

The gain decreases toward a minimal positive value in order to track the possible evolutions of the parameters. In addition, the size of the gradient steps is limited.

A case study

The results presented in this paper are based on some simulated data. The numerical values are representative of the first two modes of a real civil structure, and more, the parameter values were estimated on the structure using a batch subspace identification procedure.

Looking at two modes allows us to study parameter variations, which are characteristic of the flutter problem, which drives the application we are interested in. The parameter variations include frequencies crossing and abrupt changes in the damping. Those scenarios are illustrated in Fig. 1. Notice that, whereas we know what change scenarios we can expect from the frequency and damping in term of trend and amplitude, the associated eigenvalues variations have no real physical meaning.

The algorithm was preliminary initialized with some guessed starting values, then the filter was computed for a few hundred samples to initialize the tracking algorithm with correct estimates for the filter, then the tracking algorithm was processed on the time varying data.

The data samples were simulated with a sampling rate of 128Hz. The estimation plots are displayed with time (in sec.) on the x–coordinate. The simulated changes include for the first mode a slow increase in the frequency as well as a slow decrease in its damping value and for the second mode a slow decrease of the frequency and a abrupt increase in the damping.

Let $n = 2$ and $d = 4$.

- mode 1 : $\lambda_1 = 0.9832823 + j\,0.1520823$, $\mathbf{d}_1 = 0.032818$, $\mathbf{f}_1 = 3.1261001$

$$\psi_1 = \begin{bmatrix}-0.110149857\\ 0.003170271\\ -0.238437343\\ 0.011789335\end{bmatrix} + j\begin{bmatrix}-0.001391672\\ -0.000642400\\ 0.002764028\\ -0.000028845\end{bmatrix}$$

- mode 2 : $\lambda_2 = 0.9765406 + j\,0.1905859$, $\mathbf{d}_2 = 0.0261820$, $\mathbf{f}_2 = 3.9265001$

$$\psi_2 = \begin{bmatrix}-0.005535022\\ -0.116521290\\ -0.010837860\\ -0.219088797\end{bmatrix} + j\begin{bmatrix}-0.000479459\\ -0.000719393\\ -0.000364371\\ 0.005224397\end{bmatrix}$$

Looking at Fig. 1, we plot both estimated and true variations for both frequencies and dampings. The two frequencies are crossing each other. Nonetheless both frequency estimates stay very close to their expected value, whereas the damping estimates do exhibit worse behavior, but still react to the small changes in their nominal values. As expected, the algorithm has more problems to react to an abrupt change (see damping d_2) than a progressive change (see damping d_1). Considering the variations in the damping, it would be wise to associate a detection procedure to the tracking algorithm to decide whether the damping has changed or not.

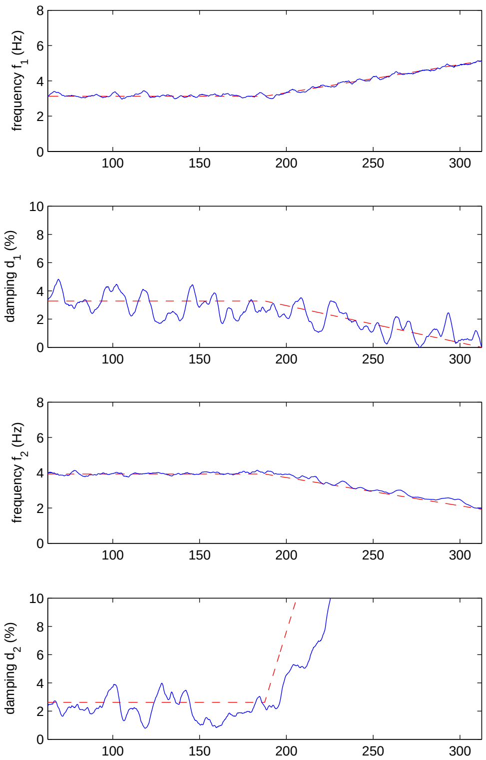


Fig. 1. Kalman RMLE procedure: monitoring of the frequency and damping parameters $(\mathbf{f}_i; \mathbf{d}_i)$, $i = 1, 2$ (true value: dashed line).

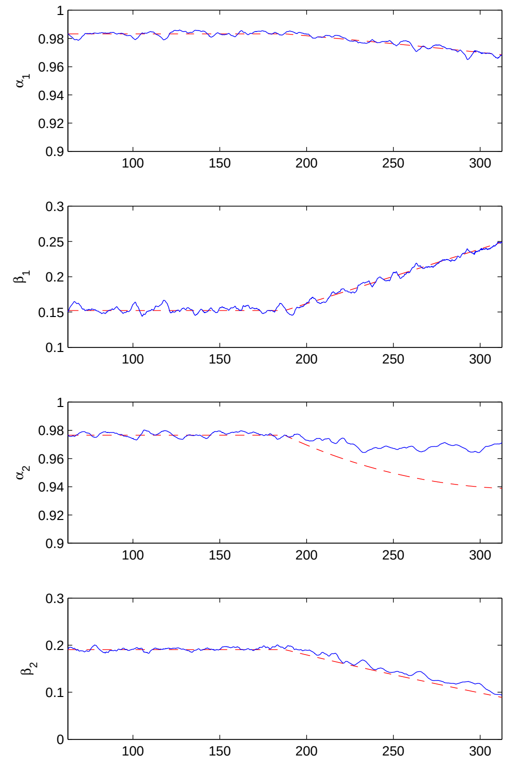


Fig. 2. Kalman RMLE procedure: monitoring of the eigenvalues $\alpha_i = \Re(\lambda_i)$ and $\beta_i = \Im(\lambda_i)$, $i = 1, 2$ (true value: dashed line).

Looking at Fig. 2, one can see that the large variations in damping d_2 do reflect in a bad estimation for α_2, whereas the slightly large variations in damping d_1 in Fig. 1 can not be inferred from the estimation of α_1 and β_1 in Fig. 2. This pleads in favor of parametrization (27).

VI. Perspectives

We have investigated the merits of both Kalman and particle filtering for structural health monitoring. The current case study is a simulation experiment, where it is expected that both methods will give similar results as seen in Fichou *et al* [6] on a simpler example.

It appears that frequency/damping parameterization (27) yields to an algorithm much simpler to tune than using the alternate parameterization (26). Moreover focusing on eigenvalues hide the uncertainties on the damping, which may be badly estimated whereas the associated eigenvalue estimation does not exhibit large variations.

The final paper will investigate the merits of both methods on a real non stationary aircraft structure where the simple “nonlinear” formulation of the particle filter may be more robust with respect to non stationary changes in the geometry of the excitation, where modes may be non excited for a short period of time.

References

[1] L. Mevel, M. Goursat, and A. Sam, “Automated on–line monitoring during a flight,” in *22nd International Modal Analysis Conference (IMAC)*, Dearborn, Mi., USA, 2004.

[2] A. Doucet, N. de Freitas, and N. J. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*, ser. Statistics for Engineering and Information Science. New York: Springer–Verlag, 2001.

[3] A. Doucet and V. B. Tadić, “Parameter estimation in general state–space models using particle methods,” *Annals of the Institute of Statistical Mathematics*, vol. 55, no. 2, pp. 409–422, 2003.

[4] A. Guyader, F. L. Gland, and N. Oudjane, “A particle implementation of the recursive MLE for partially observed diffusions,” in *Proceedings of the 13th IFAC/IFORS Symposium on System Identification*, Rotterdam, August 27-29 2003, pp. 1305–1310.

[5] N. Caylus, A. Guyader, F. L. Gland, and N. Oudjane, “Application du filtrage particulaire à l’inférence statistique des HMM,” in *Actes des XXXVIèmes Journées de Statistique (SFDS’04)*, 2004.

[6] J. Fichou, F. L. Gland, and L. Mevel, “Particle–based methods for parameter estimation and tracking : Numerical experiments,” IRISA, Rennes, Publication Interne 1604, Feb. 2004.

[7] I. Yoshida and T. Sata, “Health monitoring algorithm by the Monte Carlo filter based on non–Gaussian noise,” *Journal of Natural Disaster Science*, vol. 24, no. 2, pp. 101–107, 2002.

[8] J. Ching, J. L. Beck, K. A. Porter, and R. Shaikhutdinov, “Real–time bayesian state estimation of uncertain dynamical systems,” California Institute of Technology, Tech. Rep. EERL-2004-01, 2004.

[9] P. V. Overschee and B. D. Moor, *Subspace Identification for Linear Systems: Theory – Implementation – Methods*. Kluwer, 1996.

[10] B. Heidergott and F. J. Vázquez-Abad, “Measure valued differentiation for stochastic processes: The finite horizon case,” Submitted to Advances in Applied Probability.