Adaptive-gain Extended Kalman Filter: Extension to the Continuous-discrete Case

Nicolas Boizot
Faculty of Science,
Technology and Communication,
University of Luxemburg,
L-1359 Luxemburg
Email: nicolas.boizot@uni.lu

Eric Busvelle
IUT Dijon-Auxerre,
LE2I
Route des plaines de l'Yonne,
Auxerre, F-59000, France
Email: busvelle@u-bourgogne.fr

Jean-Paul Gauthier
IUT of Toulon,
University of Toulon,
UMR 6169,
F-83130 La Garde, France
Email:jean-paul.gauthier@univ-tln.fr

***Abstract*— In the present article we propose a nonlinear observer that merges the behaviors 1) of an extended Kalman filter, mainly designed to smooth off noise , and 2) of high-gain observers devoted to handle large perturbations in the state estimation. We specifically aim at continuous-discrete systems.**

The strategy consists in letting the high-gain self adapt according to the innovation.

We define *innovation computed over a time window* and justify its usage via an important lemma. We prove the general convergence of the resulting observer.

I. Introduction

As usual in control we consider a system of the form

$$\left\{ \begin{array}{rcl} \frac{dx}{dt} & = & f(x,u,t) \\ y & = & h(x,u,t) \end{array} \right. \tag{1}$$

The observation problem is that of estimating the state variables (x) based on the knowledge of the control variables (u) and the measured variables (y).

Solutions to this problem first came out in the linear case in the 1960's with the works of Bucy, Kalman [11] and Luenberger [12]. Those algorithms were then modified so as to cope with nonlinear systems.

One of them, the extended Kalman filter, consists in linearizing the system along the estimated trajectory (the actual one being unknown) and using the equations of the linear filter. Although this biased linearization of the system prevents from analytically proving the convergence of the observer for any initial error, such proofs exist when initial estimation errors are small enough: [2], [4].

However provided that the nonlinear system can be transformed into one of the special forms that express observability and configuring the observer with a structure denoted *high-gain*, then the convergence for any initial error can be proven: [3], [10]. This high-gain structure is a modification of the covariance matrices R and Q by the use of a fixed scalar parameter (denoted θ): convergence is effective when θ is large enough: [8]–[10].

Contrarily to the extended Kalman filter which has good noise smoothing properties [13], its high-gain counterpart is likely to amplify the effect of noise and therefore renders the estimated state unusable. In the following, we propose an extended Kalman filter that has a dynamically sized parameter θ. This adaptation is driven by a quality measurement of the estimation: *innovation computed over a time window*. The basic idea we developed for this observer is to let θ be around 1 when the estimation error is small and have θ increase when the estimation error is large. The present article is dedicated to continuous-discrete systems (see Sec II-A below).

On the front of discrete time systems, M. Boutayeb *and al.* (e.g. in [7]) proposed a criteria for asymptotic convergence of the extended Kalman filter. Their criteria can be linked to innovation thus rendering the observer adaptive (see equation (51) in [7]). In a different manner, we take into consideration systems that are under the observability normal form: it shortens the range of usable systems but gives us exponential convergence in arbitrarily small time instead of asymptotic convergence only.

In Section II we define both the system under consideration and the observer. The proof of convergence is given in Section III. the main result is stated in Sec III-A.

Remark:

We chose to dedicate this article to the full exposure of the proof of convergence of the observer. We tried to keep it as self contained as possible but for the part dealing with the Riccati equation.

As a consequence, no illustrative example comes to enlighten a rather technical proof. Such a practical implementation of the adaptative high-gain extended Kalman filter may be found, although in the continuous case, in [6] and in [3] (chapter 3.5). Details on the implementation of the continuous-discrete case will be available in [5].

II. Definitions

A. System definition

For the sake of clarity, we restrain the proof to single output systems. As there is no unique observability form for multiple output systems, then the observer has to be adapted to each situation. This consists only of minor modifications. The overall convergence result remains valid.

The multiple inputs/single output observability form of ([4]) is adapted to the continuous-discrete setting (on this topic, see also [1]) as:

$$\left\{ \begin{array}{rcl} \frac{dx}{dt} & = & A(u(t))x + b(x(t),u(t)) \\ y_k & = & C(u_k)x_k \end{array} \right. \tag{2}$$

where

- δ_t is the constant sampling time
- $x_k = x(k\delta_t) \in \mathcal{X} \subset \mathbb{R}^n$, $\mathcal{X}$ compact, $k \in \mathbb{N}$,
- $y_k \in \mathbb{R}$, $k \in \mathbb{N}$,
- $u_k = u(k\delta_t) \in \mathcal{U}_{\text{adm}} \subset \mathbb{R}^{n_u}$, bounded, $k \in \mathbb{N}$.

The matrices $A(u)$ and $C(u_k)$ are:

$$A(u) = \begin{pmatrix} 0 & a_2(u) & 0 & \cdots & 0 \\ & 0 & a_3(u) & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & 0 & a_n(u) \\ 0 & & \cdots & & 0 \end{pmatrix}$$

$$C(u_k) = \begin{pmatrix} a_1(u_k) & 0 & \cdots & 0 \end{pmatrix}$$

with $0 < a_m \leq a_i(u) \leq a_M$ for any u in $\mathcal{U}_{\text{adm}}$. The $\mathcal{C}^1$ vector field $b(x,u)$ is assumed to be compactly supported and to have the triangular structure:

$$b(x,u) = \begin{pmatrix} b_1(x_1,u) \\ b_2(x_1,x_2,u) \\ \vdots \\ b_n(x_1,\ldots,x_n,u) \end{pmatrix}$$

We denote L_b the bound of $b^*(x,u)$, the Jacobian matrix of $b(x,u)$ *w.r.t* x (i.e. $\|b^*(x,u)\| \leq L_b$). Since $b(x,u)$ is compactly supported and u is bounded, b is Lipschitz uniformly in x: $\|b(x_1,u) - b(x_2,u)\| \leq L_b \|x_1 - x_2\|$.

B. Observer definition

The observer is defined by two sets of equations:

1) between two consecutive measurements (i.e. for $t \in [(k-1)\delta_t; k\delta_t[$, $k \in \mathbb{N}^*$) a prediction of the state estimate is computed from continuous equations,
2) whenever a new sample is available (i.e. for $t = k\delta_t$, $k \in \mathbb{N}^*$) a correction is applied to the estimation. A new state estimation is now at hand.

In the following

- $z(t)$ is the estimated state for $t \in [(k-1)\delta_t; k\delta_t[$,
- $z_k(-)$ is the estimated state at the end of a prediction period, before any correction step,
- $z_k(+)$ is the estimated state after a correction step.

Prediction equations for $t \in [(k-1)\delta_t; k\delta_t[$:

- initial values are $z_{k-1}(+)$, $S_{k-1}(+)$, $\theta((k-1)\delta t)$,
- final values are $z_k(-)$, $S_k(-)$, $\theta(k\delta t)$,

$$\left\{ \begin{array}{rcl} dz/dt &=& A(u)z + b(z,u) \\ dS/dt &=& -\left(A(u) + b^*(z,u)\right)' S - S\left(A(u) + b^*(z,u)\right) \\ && -SQ_\theta S \\ d\theta/dt &=& \mathcal{F}(\theta, \mathcal{I}_{k,d}) \end{array} \right. \quad (3)$$

Correction equations at time $t = k\delta_t$:

- state before update: $z_k(-)$, $S_k(-)$, $\theta(k\delta t)$
- final state: $z_k(+)$, $S_k(+)$, $\theta(k\delta t)$:

$$\left\{ \begin{array}{rcl} z_k(+) &=& z_k(-) - S_k(+)^{-1} C' r_\theta^{-1} \delta_t (C z_k(-) - y_k) \\ S_k(+) &=& S_k(-) + C' r_\theta^{-1} C \delta_t \\ \mathcal{I}_{k,d} &=& \sum_{i=0}^{i=d} \|y_{k-i} - \hat{y}_{k-i}\|^2 \end{array} \right. \quad (4)$$

Moreover, let us assume that

- $\theta(0) = 1$, S_0 is a symmetric definite positive matrix taken inside a compact of the form $aId \leq S_0 \leq bId$.
- $\mathcal{I}_{0,d} = 0$,
- r and Q are symmetric definite positives matrices and:
  - $\Delta = diag\{1, \frac{1}{\theta}, ..., \frac{1}{\theta^{n-1}}\}$,
  - $Q_\theta = \theta \Delta^{-1} Q \Delta^{-1}$,
  - $r_\theta = \frac{1}{\theta} r$.

The term $\mathcal{I}_{k,d}$ ($d \in \mathbb{N}^*$) is the *innovation*:

- y_{k-i}, for $i = 0, ..., d$ are the measurements,
- $\hat{y}_{k-i}$ is the output of (2), computed at epoch $k-i$,with initial conditions z_{k-d}. It is a prediction of the trajectory over the time window $[(k-d)\delta_t; k\delta_t]$, with the d-delayed estimated state as initial conditions.

Innovation is a quality measurement of the state estimation and plays a crucial role in the observer structure, as explained in Sec III-B.

III. Convergence result and proof

A. Main result

Theorem 1: For any time $T^* > 0$ and any $\epsilon^* > 0$, there exists:

- two real positive constants, μ and θ_1,
- $d \geq n - 1 \in \mathbb{N}^*$
- and a function $\mathcal{F}(\theta, \mathcal{I})$,

such that for all small enough δ_t (i.e. $2\theta_1\delta_t < \mu$ and $0 < d\delta_t < T^*$) and any time $t \geq T^*$ and any $(z(0), x(0)) \in \chi^2$:

$$\|z(t) - x(t)\|^2 \leq \epsilon^* e^{-a(t-T^*)}$$

where $a > 0$ does not depend on θ. $\diamondsuit$

This theorem is based on two well known convergence results of the Kalman filter:

- local convergence when $\theta = 1$,
- global convergence when the high-gain parameter is taken large enough.

At time 0, θ is equal to 1. Let us suppose that the initial state is far away from the real state. This is detected at a time $k\delta_t$ or more precisely at epoch k. This makes θ increase in such a way that it attains a high enough value for *convergence for any estimation error* to happen. Ultimately, when the estimation error is back into a neighborhood of 0, the high-gain parameter decreases.

There are clearly 3 different situations:

1) the estimation error is not detected yet or the high-gain hasn't reached a high enough value,
2) the observer is in high-gain mode,
3) the estimation error is back inside a neighborhood of zero, the high-gain is no more useful.

The proof consists in upper bounding the estimation error in each one of those three cases. The three inequalities are glued together and we show that with the appropriate choice of parameters the inequality of the theorem is met. We also describe some possible adaptation function ($\mathcal{F}$).

The quality of the state estimation is reported by ($\mathcal{I}_{k,d}$), the innovation computed over a time window of length $d\delta_t$, as defined in (4). We show in lemma 2 that this quantity upper bounds the estimation error.

The proof of the theorem is divided into 5 parts:

- the key lemma for innovation is the object of Sec III-B,
- preliminary calculations are done in Sec III-C,
- Sec III-D deals with the bounds of the Riccati matrix,
- useful technical lemmas are mentioned in Sec III-E,
- and finally, all the arguments are articulated in Sec III-F.

Remark:

We will see in Sec III-D that a property of the Riccati equation ($2\theta_1\delta_t < \mu$, in the theorem above), forces us to develop the beginning of the proof independently from δ_t (see Sec III-F). Those considerations are, of course, absent in the continuous case.

B. Innovation

The following lemma justifies the use of innovation as a quality measurement.

Lemma 2: Let $x^0, \xi^0 \in \mathbb{R}^n$ and $u \in \mathcal{U}_{\text{adm}}$. Let us consider the outputs $y_j\left(0, x^0\right)$ and $y_j\left(0, \xi^0\right)$ of system (2) with initial conditions respectively x^0 and ξ^0. Then the following condition (called persistent observability) holds:

$$\text{For all } d \in \mathbb{N}^*,\ d \geq n-1,\ \exists \lambda_d^0 > 0 \text{ such that } \forall u \in L_b^1(\mathcal{U}_{adm})$$
$$\|x^0 - \xi^0\|^2 \leq \frac{1}{\lambda_d^0} \sum_{i=0}^{i=d} \|y_i(0, x^0) - y_i\left(0, \xi^0\right)\|^2. \ \diamondsuit$$

Proof: Let $x(t)$ and $\xi(t)$ be the solutions of the first equation of (2) with x^0 and ξ^0 as initial values and the controls $u(t)$. For any $a \in [0,1]$:

$$\begin{aligned} & b(a\xi + (1-a)x, u) \\ & = b(x,u) + \int_0^a \frac{\partial b}{\partial x}(\alpha\xi + (1-\alpha)x, u) \frac{\partial(\alpha\xi + (1-\alpha)x)}{\partial \alpha} d\alpha \\ & = b(x,u) + \left[\int_0^a \frac{\partial b}{\partial x}(\alpha\xi + (1-\alpha)x, u) d\alpha\right](\xi - x) \end{aligned} \tag{5}$$

and so, for $a = 1$,

$$b(\xi, u) - b(x, u) = B(t)(\xi - x) \tag{6}$$

where $B(t) = (b_{i,j})_{(i,j)\in\{1,..,n\}}$ is a lower triangular matrix since $b(x,u) = \left[b(x_1,u), b(x_1,x_2,u), \ldots, b(x,u)\right]'$. Set $\varepsilon = x - \xi$,

$$\begin{aligned} \frac{d\varepsilon}{dt} &= A(u)x + b(x,u) - A(u)\xi - b(\xi,u) \\ &= [A(u) + B(t)]\,\varepsilon \end{aligned} \tag{7}$$

and with $C(u_k)\varepsilon_k = a_1(u_k)\varepsilon_{1,k}$ as output. Let us consider $\Psi(t)$, the resolvent of the system (7), and the Gramm observability matrix

$$G_d = \sum_{i=0}^{i=d} \Psi(i\delta_t)' C' C \Psi(i\delta_t)$$

From the lower triangular structure of $B(t)$, the upper triangular structure of $A(u)$ and the form of the matrix C, we can deduce that G_d is invertible when $d \geq n-1$ (we need at least n points to obtain a full rank matrix), and therefore is symmetric positive definite. As $\|B(t)\| \leq L_b$ each non-zero $b_{i,j}(t)$ can be seen as a bounded element of $L^\infty_{[0,d]}(\mathbb{R})$. We consider the function:

$$\begin{aligned} \Lambda:\ L^\infty_{[0,d]}\left(\mathbb{R}^{\frac{n(n+1)}{2}}\right) \times L^\infty_{[0,d]}\left(\mathbb{R}^{n_u}\right) &\longrightarrow \mathbb{R}^+ \\ (b_{i,j})_{(j\leq i)\in\{1,..,n\}}, u &\hookrightarrow \lambda_{\min}(G_d) \end{aligned}$$

where $\lambda_{\min}(G_d)$ is the smallest eigenvalues of G_d.

We endow $\mathbb{R}$ with the topology of uniform convergence and $E = L^\infty_{[0,d]}\left(\mathbb{R}^{\frac{n(n+1)}{2}}\right) \times L^\infty_{[0,d]}\left(\mathbb{R}^n\right)$ with the *-weak topology such that Λ is continuous as a composition of continuous functions.

Since b is Lipschitz and $\mathcal{U}_{\text{adm}}$ is bounded, B and u lies in a relatively compact subset A in E. Then $\Lambda(A)$ lies in a compact subset of $\mathbb{R}^+$ that does not contain 0 since G_d is positive definite. Hence, there exists λ_d^0 such that $G_d \geq \lambda_d^0 Id$ for any u and any matrix $B(t)$ having the structure depicted above. We obtain:

$$y_i\left(0, x^0\right) - y_i\left(0, \xi^0\right) = C\Psi(i\delta_t)x^0 - C\Psi(i\delta_t)\xi^0$$

then

$$\left\|y_i\left(0,x^0\right) - y_i\left(0,\xi^0\right)\right\|^2 = \left\|C\Psi(i\delta_t)x^0 - C\Psi(i\delta_t)\xi^0\right\|^2$$

and finally

$$\begin{aligned} \sum_{i=0}^{i=d} \left\|y_i\left(0,x^0\right) - y_i\left(0,\xi^0\right)\right\|^2 &= \left(x^0 - \xi^0\right)' G_d \left(x^0 - \xi^0\right) \\ &\geq \lambda_d^0 \left\|x^0 - \xi^0\right\|^2 \end{aligned}$$

■

C. Preparation for the proof

In this section a few preliminary relations are established. We first recall the *matrix inversion lemma* [10]:

Lemma 3: If M is symmetric positive definite, and if λ is small, then $(M + \lambda M C' C M)^{-1} = M^{-1} - C'(\lambda^{-1} + CMC')^{-1}C$.

We denote the estimation error by $\epsilon(t) = z(t) - x(t)$. Let us consider the change of variables $\tilde{x} = \Delta x$, $\tilde{z} = \Delta z$, $\tilde{\epsilon} = \Delta\epsilon$, $\tilde{S} = \Delta^{-1}S\Delta^{-1}$, $\tilde{b}(.,u) = \Delta b(\Delta^{-1}.,u)$ and $\tilde{b}^*(.,u) = \Delta b^*(\Delta^{-1}.,u)\Delta^{-1}$. The Lipschitz constant of the vector field $b(x,u)$ is the same in the new system of coordinates.

1) Continuous Part: The error dynamics are given by

$$\frac{d\epsilon}{dt} = A(u)\epsilon + (b(z,u) - b(x,u)) \tag{8}$$

We have, with $\mathcal{N} = diag\{0, 1, \ldots, n-1\}$:

$$\begin{aligned} \frac{d}{dt}(\Delta) &= -\frac{\mathcal{F}(\theta,\mathcal{I})}{\theta}\mathcal{N}\Delta & \Delta A(u) &= \theta A(u)\Delta \\ \frac{d}{dt}(\Delta^{-1}) &= \frac{\mathcal{F}(\theta,\mathcal{I})}{\theta}\mathcal{N}\Delta^{-1} & \Delta^{-1}A'(u) &= \theta A'(u)\Delta^{-1}. \end{aligned} \tag{9}$$

For $t \in [k\delta_t; (k+1)\delta_t[$,

$$\frac{d\tilde{\epsilon}}{dt} = \theta\left[\left(-\frac{\mathcal{F}(\theta,\mathcal{I})}{\theta^2}\mathcal{N} + A(u)\right)\tilde{\epsilon} + \frac{1}{\theta}\left(\tilde{b}(\tilde{z},u) - b(\tilde{x},u)\right)\right] \tag{10}$$

and,

$$\frac{d\tilde{S}}{dt} = \theta\Big[-(A(u)+\tfrac{1}{\theta}\tilde{b}^*(z,u)-\tfrac{\mathcal{F}(\theta,\mathcal{I})}{\theta^2}\mathcal{N})'\tilde{S} \\ -\tilde{S}(A(u)+\tfrac{1}{\theta}\tilde{b}^*(z,u)-\tfrac{\mathcal{F}(\theta,\mathcal{I})}{\theta^2}\mathcal{N})-\tilde{S}Q\tilde{S}\Big] \quad (11)$$

Consider now the Lyapunov function $\tilde{\epsilon}'\tilde{S}\tilde{\epsilon}$ and use identities (10, 11) in order to compute its time derivative:

$$\frac{d\left(\tilde{\epsilon}'\tilde{S}\tilde{\epsilon}\right)}{dt} = \theta\left[\tfrac{2}{\theta}\tilde{\epsilon}'\tilde{S}(\tilde{b}(\tilde{z},u)-\tilde{b}(\tilde{x},u)-\tilde{b}^*(\tilde{z},u)\tilde{\epsilon}) \\ -\tilde{\epsilon}'\tilde{S}Q\tilde{S}\tilde{\epsilon}\right] \quad (12)$$

2) Discrete part: Estimation error at time $k\delta t$ is:

$$\tilde{\epsilon}_k(+) = \left(Id - \theta\delta_t\tilde{S}_k^{-1}(+)C'R^{-1}C\right)\tilde{\epsilon}_k(-) \quad (13)$$

and,

$$\tilde{S}_k(+) = \tilde{S}_k(-) + \theta\delta_t C'R^{-1}C \quad (14)$$

As for the continuous part, we use (13) and (14) to compute the Lyapunov function at time $k\delta_t$,

$$\left(\tilde{\epsilon}'\tilde{S}\tilde{\epsilon}\right)_k(+) = \tilde{\epsilon}_k'(-)\left[\tilde{S}_k(+) - 2\theta\delta_t C'R^{-1}C \\ +(\theta\delta_t)^2C'R^{-1}C\tilde{S}_k(+)^{-1}C'R^{-1}C\right]\tilde{\epsilon}_k(-) \quad (15)$$

from (14), we replace $\left(\theta\delta t C'r^{-1}C\right)$ by $\left(\tilde{S}_k(+)-\tilde{S}_k(-)\right)$,

$$\left(\tilde{\epsilon}'\tilde{S}\tilde{\epsilon}\right)_k(+) = \tilde{\epsilon}_k'(-)\left[\tilde{S}_k(-)\tilde{S}_k(+)^{-1}\tilde{S}_k(-)\right]\tilde{\epsilon}_k(-) \\ = \tilde{\epsilon}_k'(-)\left[\tilde{S}_k(-)^{-1}\tilde{S}_k(+)\tilde{S}_k(-)^{-1}\right]^{-1}\tilde{\epsilon}_k(-) \quad (16)$$

We use equation (14) and Lemma 3 with $\lambda = \frac{\theta\delta_t}{r}$ and $M = \tilde{S}_k^{-1}(-)$ to compute $[S_k^{-1}(-)S_k(+)S_k^{-1}(-)]^{-1}$ and then

$$\left(\tilde{\epsilon}'\tilde{S}\tilde{\epsilon}\right)_k(+) = \left(\tilde{\epsilon}'\tilde{S}\tilde{\epsilon}\right)_k(-) \\ -\tilde{\epsilon}_k'(-)\left[C'(\tfrac{r}{\theta\delta_t}+C\tilde{S}_k^{-1}(-)C')^{-1}C\right]\tilde{\epsilon}_k(-) \quad (17)$$

At the end of this subsection, we now have at our disposal the two important identities (12) and (17). They are the starting point of the proof of Sec. III-F.

D. On the Riccati equation

In order to upper bound estimation error in Sec. III-F, we need to upper and lower bound the Riccati matrix $\tilde{S}$.

Lemma 4: Consider the prediction-correction Riccati equations (11,14) where:

- the functions $a_i(u(t))$, $\left|\tilde{b}^*_{i,j}(\overline{z},\overline{u})\right|$, $\left|\frac{\mathcal{F}(\overline{\theta},\overline{\mathcal{I}})}{\overline{\theta}^2}\right|$ are smaller than $a_M > 0$
- $a_i(u(t)) \geq a_m > 0$, $\forall\ i = 1, ..., n$,
- $\theta(0) = 1$,
- $S(0) = S_0$ lives in a compact of the form $aId \leq S_0 \leq bId$.

then there exists a constant μ such that if the sample time δ_t is small enough (i.e. $\theta(t)\delta_t < \mu$, $\forall\ t > 0$) there exists two constants $0 < \alpha < \beta$ such that for all $k \in \mathbb{N}$ and for all $t \in [k\delta_t;(k+1)\delta_t]$

$$\alpha Id < \tilde{S} < \beta Id \quad (18)$$

where both α and β depend neither on θ nor on δ_t. $\diamondsuit$

We are lacking space to write the full proof of this lemma here. We only provide the reader with the main ideas.

In both (11) and (14) we spot the presence of a θ factor in the equations. Since we don't know yet the maximum value θ shall reach, we cannot use those equations to obtain the desired bounds. A change in the time scale, defined as $d\tau = \theta(t)dt$ makes those factors disappear (the $1/\theta$ factors that remain in (11) don't cause any problems).

In this time scale, we can prove that for any given $\tau_0 > 0$, an inequality of the form (18) is true for all $\tau \geq \tau_0$. This is done following the methodology of [4], [10].

Usually, at this point, we cannot deduce much more because $\tilde{S}(0)$ depends heavily on $\theta(0)$ which is unknown. As a consequence, the proof of the theorem has to be handled partly in the τ time scale. However in the present situation, $\theta(0) = 1$. Therefore $\tilde{S}(0) = S(0)$ and since we have $aId \leq S(0) \leq bId$ we can easily bound $\tilde{S}$, in the time scale τ, for $0 \leq \tau < \tau_0$, with the use Gronwall's lemma.

The double inequalities obtained for $0 \leq \tau < \tau_0$ and for $\tau_0 \leq \tau$ are merged into the double inequality (18) valid for all times τ. Thus it is also true for all times t. The two bounds are independent from both $\theta(t)$ and δ_t.

Remark: Two very important assumptions come from the first part of this lemma:

1) *$\mathcal{F}(\theta,\mathcal{I})/\theta^2 \leq a_M$, independently from θ_1,*
2) *$\exists\ \mu > 0$, such that $\theta(t)\delta_t$ must always be less than μ.*

The first one is taken care of in lemma 7. The second one implies that, since we don't know which value of θ renders convergence effective, we cannot upper bound $\theta(t)$ yet. This bound can be afterwards compensated by δ_t in order to respect the inequality $\theta(t)\delta_t < \mu$. It implies that until we set a maximum value for θ, the proof must not depend on δ_t.

E. Technical lemmas

The two following lemmas have been proven in the appendix of [4]. Lemma 7 defines the adaptation function.

Lemma 5: Let $\{x(t) > 0, t \geq 0\} \subset \mathbb{R}^n$ be absolutely continuous and satisfying:

$$\frac{dx(t)}{dt} \leq -k_1x + k_2x\sqrt{x},$$

for almost all $t > 0$, for $k_1, k_2 > 0$. Then, when $x(0) < \frac{k_1^2}{4k_2^2}$, $x(t) \leq 4\,x(0)\,e^{-k_1t}$. $\diamondsuit$

Lemma 6: Consider $\widetilde{b}(\tilde{z}) - \widetilde{b}(\tilde{x}) - \widetilde{b}^*(\tilde{z})\,\tilde{\varepsilon}$ which appears in the inequality (12) (for clarity, omitting u in $\widetilde{b}$) and suppose $\theta \geq 1$.

Then $\left\|\widetilde{b}(\tilde{z}) - \widetilde{b}(\tilde{x}) - \widetilde{b}^*(\tilde{z})\,\tilde{\varepsilon}\right\| \leq K\theta^{n-1}\left\|\tilde{\varepsilon}\right\|^2$, for some $K > 0$. $\diamondsuit$

Lemma 7: For any $\Delta T > 0$, there exists a positive constant M such that for any $\theta_1 > 1$ and any $\gamma_1 > \gamma_0 > 0$, there is a function $\mathcal{F}(\theta, \mathcal{I})$ such that, considering the following equation for any initial value $1 \leq \theta(0) < 2\theta_1$ and any measurable positive function $\mathcal{I}(t)$

$$\dot{\theta} = \mathcal{F}(\theta, \mathcal{I}(t)) \tag{19}$$

we have:

1) (19) has a unique solution $\theta(t)$ defined for all $t \geq 0$, and this solution satisfies $1 \leq \theta(t) < 2\theta_1$,
2) $|\frac{\mathcal{F}(\theta,\mathcal{I})}{\theta^2}| \leq M$,
3) if $\mathcal{I}(t) \geq \gamma_1$ for $t \in [\tau, \tau + \Delta T]$ then $\theta(\tau + \Delta T) \geq \theta_1$,
4) while $\mathcal{I}(t) \leq \gamma_0$, $\theta(t)$ decreases to 1. $\diamondsuit$

Remark:

The main non-obvious property is that if $\mathcal{I}(t) \geq \gamma_1$, $\theta(t)$ can reach an arbitrary large θ_1 in an arbitrary small time ΔT, and that this property can be achieved by a function satisfying $\mathcal{F}(\theta, \mathcal{I}) \leq M\theta^2$ with M independent from θ_1 (but dependent on ΔT).

Proof: Let $\mathcal{F}_0(\theta)$ be defined as follows:

$$\mathcal{F}_0(\theta) = \begin{cases} \frac{1}{\Delta T}\theta^2 & \text{if } \theta \leq \theta_1 \\ \frac{1}{\Delta T}(\theta - 2\theta_1)^2 & \text{if } \theta > \theta_1 \end{cases}$$

(the choice $2\theta_1$ is more or less arbitrary) and let us consider the system

$$\begin{cases} \dot{\theta} &= \mathcal{F}_0(\theta) \\ \theta(0) &= 1 \end{cases}$$

Simple computations give the solution (with $\theta(0) = 1$):

$$\theta(t) = \begin{cases} \dfrac{\Delta T}{\Delta T - t} & \text{while } \theta \leq \theta_1 \\ 2\theta_1 - \dfrac{\theta_1 \Delta T}{\theta_1 t + (2 - \theta_1)\Delta T} & \text{when } \theta > \theta_1 \end{cases}$$

Therefore, since the system is autonomous, $\theta(t)$ reaches θ_1 at time $t < \Delta T$ (for any value of $\theta(0) \in [1, 2\theta_1[$). Let us remark also that $\mathcal{F}_0$ is Lipschitz. Now, let us define

$$\mathcal{F}(\theta, \mathcal{I}) = \mu(\mathcal{I})\mathcal{F}_0(\theta) + (1 - \mu(\mathcal{I}))\lambda(1 - \theta)$$

for a $\lambda > 0$ and with

$$\mu(\mathcal{I}) = \begin{cases} 1 & \text{if } \mathcal{I} \geq \gamma_1 \\ 0 & \text{if } \mathcal{I} \leq \gamma_0 \end{cases}$$

and $0 \leq \mu(\mathcal{I}) \leq 1$ for $\gamma_0 \leq \mathcal{I} \leq \gamma_1$. We claim that all properties are satisfied.

If $\mathcal{I} \geq \gamma_1$, $\mathcal{F}(\theta, \mathcal{I}) = \mathcal{F}_0(\theta)$ ensuring *Property 3*, due to the first part of the proof. Conversely, if $\mathcal{I} \leq \gamma_0$, $\mathcal{F}(\theta, \mathcal{I}) = \lambda(1 - \theta)$ and this implies *Property 4*. Moreover, $\mathcal{F}(\theta, \mathcal{I})$ is Lipschitz and so *Property 1* is verified. Finally:

$$\left|\frac{\mathcal{F}(\theta, \mathcal{I})}{\theta^2}\right| \leq \left|\frac{\mathcal{F}_0(\theta)}{\theta^2}\right| + \left|\frac{\lambda(1 - \theta)}{\theta^2}\right| \tag{20}$$

however the first term is such that if $\theta \leq \theta_1$, $\left|\frac{\mathcal{F}_0(\theta)}{\theta^2}\right| = \frac{1}{\Delta T}$ and if $\theta \geq \theta_1$ (and $\theta < 2\theta_1$) :

$$\left|\frac{\mathcal{F}_0(\theta)}{\theta^2}\right| = \frac{1}{\Delta T}\left(\frac{\theta - 2\theta_1}{\theta}\right)^2 \leq \frac{1}{\Delta T}$$

and the second term satisfies

$$\left|\frac{\lambda(1 - \theta)}{\theta^2}\right| = \lambda\frac{\theta - 1}{\theta^2} = \lambda\left(\frac{1}{4} - \frac{\frac{\theta^2}{4} - \theta + 1}{\theta^2}\right) \leq \frac{\lambda}{4}$$

Property 2 is ensured because of (20), with $M = \frac{1}{\Delta T} + \frac{\lambda}{4}$. ■

F. Proof of the theorem

First of all, let us set a time T such that $0 < T < T^*$. Let λ be a strictly positive number and $M = \frac{1}{\Delta T} + \frac{\lambda}{4}$ as in Lemma 7. Let α and β be the bounds of Lemma 4. For $t \in [k\delta_t; (k+1)\delta_t[$, inequality (12) can be written, (i.e. using $\alpha Id \leq \tilde{S}$)

$$\frac{d\tilde{\varepsilon}'\tilde{S}\tilde{\varepsilon}(t)}{dt} \leq -\alpha q_m \theta \tilde{\varepsilon}'\tilde{S}\tilde{\varepsilon}(t) + 2\tilde{\varepsilon}'\tilde{S}\left(\tilde{b}(\tilde{z}) - \tilde{b}(\tilde{x}) - \tilde{b}^*(\tilde{z})\tilde{\varepsilon}\right) \tag{21}$$

with $q_m > 0$ such that $q_m Id < Q$ (and omitting to write the control variable u).

From (21) we can deduce two bounds: the first one, local, will be useful when $\tilde{\varepsilon}'\tilde{S}\tilde{\varepsilon}(t)$ is small whatever the value taken by θ. The second one, global, will be useful mainly when $\tilde{\varepsilon}'\tilde{S}\tilde{\varepsilon}(t)$ is not in a neighborhood of 0.

Global bound: Starting from:

$$\left\|\tilde{b}(\tilde{z}) - \tilde{b}(\tilde{x}) - \tilde{b}^*(\tilde{z})\tilde{\varepsilon}\right\| \leq 2L_b \|\tilde{\varepsilon}\|$$

together with $\alpha Id \leq \tilde{S} \leq \beta Id$ (Lemma 4), (21) becomes

$$\frac{d\tilde{\varepsilon}'\tilde{S}\tilde{\varepsilon}(t)}{dt} \leq \left(-\alpha q_m \theta + 4\frac{\beta}{\alpha}L_b\right)\tilde{\varepsilon}'\tilde{S}\tilde{\varepsilon}(t) \tag{22}$$

Local bound: Thanks to Lemma 6

$$\left\|\tilde{b}(\tilde{z}) - \tilde{b}(\tilde{x}) - \tilde{b}^*(\tilde{z})\tilde{\varepsilon}\right\| \leq K\theta^{n-1}\|\tilde{\varepsilon}\|^2$$

which implies, since $1 \leq \theta \leq 2\theta_1$

$$\frac{d\tilde{\varepsilon}'\tilde{S}\tilde{\varepsilon}(t)}{dt} \leq -\alpha q_m \tilde{\varepsilon}'\tilde{S}\tilde{\varepsilon}(t) + 2K(2\theta_1)^{n-1}\left\|\tilde{S}\right\|\|\tilde{\varepsilon}\|^3$$

but $\|\tilde{\varepsilon}\|^3 = \left(\|\tilde{\varepsilon}\|^2\right)^{\frac{3}{2}} \leq \left(\frac{1}{\alpha}\tilde{\varepsilon}'\tilde{S}\tilde{\varepsilon}(t)\right)^{\frac{3}{2}}$ and therefore

$$\tilde{\varepsilon}'\tilde{S}\tilde{\varepsilon}(t) \leq -\alpha q_m \tilde{\varepsilon}'\tilde{S}\tilde{\varepsilon}(t) + \frac{2K(2\theta_1)^{n-1}\beta}{\alpha^{\frac{3}{2}}}\left(\tilde{\varepsilon}'\tilde{S}\tilde{\varepsilon}(t)\right)^{\frac{3}{2}} \tag{23}$$

Let us apply Lemma 5: if there exists ξ such that

$$\tilde{\varepsilon}'\tilde{S}\tilde{\varepsilon}(\xi) \leq \frac{\alpha^5 q_m^2}{16K^2(2\theta_1)^{2n-2}\beta^2}$$

then for any $k\delta_t \leq \xi \leq t \leq (k+1)\delta_t$

$$\tilde{\varepsilon}'\tilde{S}\tilde{\varepsilon}(t) \leq 4\tilde{\epsilon}'\tilde{S}\tilde{\varepsilon}(\xi)\,\mathrm{e}^{-\alpha q_m(t-\xi)}.$$

If $\gamma \in \mathbb{R}$ such that

$$\gamma \leq \frac{1}{(2\theta_1)^{2n-2}}\min\left(\frac{\alpha\varepsilon^*}{4}, \frac{\alpha^5 q_m^2}{16K^2\beta^2}\right) \tag{24}$$

then $\tilde{\varepsilon}'\tilde{S}\tilde{\varepsilon}(\xi) \leq \gamma$ implies

$$\tilde{\varepsilon}'\tilde{S}\tilde{\varepsilon}(t) \leq \frac{\alpha\varepsilon^*}{(2\theta_1)^{2n-2}}\mathrm{e}^{-\alpha q_m(t-\xi)}. \tag{25}$$

Given any value of δ_t, there exists $k \in \mathbb{N}$ such that $T \in [k\delta_t;(k+1)\delta_t[$. From the global bound (22), with $\theta(t) \geq 1$:

$$\tilde{\varepsilon}^{'} \tilde{S}\tilde{\varepsilon}\left(T\right) \leq \tilde{\varepsilon}^{'} \tilde{S}\tilde{\varepsilon}\left(k\delta_t\right) \mathrm{e}^{\left(-\alpha q_m + 4\frac{\beta}{\alpha}L_b\right)(T-k\delta_t)}$$

But when we consider $t \in [k\delta_t;(k+1)\delta_t[$ this means that $\left(\tilde{\epsilon}^{'} \tilde{S}\tilde{\epsilon}\right)(k\delta_t) = \left(\tilde{\epsilon}^{'} \tilde{S}\tilde{\epsilon}\right)_k(+)$. We know from (17) that

$$\left(\tilde{\epsilon}^{'} \tilde{S}\tilde{\epsilon}\right)_k(+) \leq \left(\tilde{\epsilon}^{'} \tilde{S}\tilde{\epsilon}\right)_k(-) \qquad (26)$$

which means,

$$\tilde{\varepsilon}^{'} \tilde{S}\tilde{\varepsilon}\left(T\right) \leq \left(\tilde{\varepsilon}^{'} \tilde{S}\tilde{\varepsilon}\right)_k(-)\mathrm{e}^{\left(-\alpha q_m + 4\frac{\beta}{\alpha}L_b\right)(T-k\delta_t)}$$

and since $\left(\tilde{\varepsilon}^{'} \tilde{S}\tilde{\varepsilon}\right)_k(-)$ is the end value of the equation (12) for $t \in [(k-1)\delta_t; k\delta_t[$, then:

$$\left(\tilde{\varepsilon}^{'} \tilde{S}\tilde{\varepsilon}\right)_k(-) \leq \left(\tilde{\varepsilon}^{'} \tilde{S}\tilde{\varepsilon}\right)_{k-1}(+)\mathrm{e}^{\left(-\alpha q_m + 4\frac{\beta}{\alpha}L_b\right)\delta_t}.$$

Consequently, independently from δ_t, we can walk down to:

$$\tilde{\varepsilon}^{'} \tilde{S}\tilde{\varepsilon}\left(T\right) \leq \tilde{\varepsilon}^{'} \tilde{S}\tilde{\varepsilon}\left(0\right) \mathrm{e}^{\left(-\alpha q_m + 4\frac{\beta}{\alpha}L_b\right)T}. \qquad (27)$$

We suppose now that $\theta \geq \theta_1$ for $t \in [T, T^*]$, $T^* \in [\tilde{k}\delta_t;(\tilde{k}+1)\delta_t[$ and use (22):

$$\tilde{\varepsilon}^{'} \tilde{S}\tilde{\varepsilon}\left(T^*\right) \leq \tilde{\epsilon}^{'} \tilde{S}\tilde{\varepsilon}\left(\tilde{k}\delta_t\right) \mathrm{e}^{\left(-\alpha q_m \theta_1 + 4\frac{\beta}{\alpha}L_b\right)(T^*-\tilde{k}\delta_t)} \qquad (28)$$

which can be rewritten (with the same argument as before), independently from δ_t

$$\tilde{\varepsilon}^{'} \tilde{S}\tilde{\varepsilon}\left(T^*\right) \leq \tilde{\epsilon}^{'} \tilde{S}\tilde{\varepsilon}\left(T\right) \mathrm{e}^{\left(-\alpha q_m + 4\frac{\beta}{\alpha}L_b\right)T} \mathrm{e}^{\left(-\alpha q_m \theta_1 + 4\frac{\beta}{\alpha}L_b\right)(T^*-T)} \qquad (29)$$

$$\tilde{\varepsilon}^{'} \tilde{S}\tilde{\varepsilon}\left(T^*\right) \leq M_0 \mathrm{e}^{-\alpha q_m T} \mathrm{e}^{4\frac{\beta}{\alpha}L_b T^*} \mathrm{e}^{-\alpha q_m \theta_1 (T^*-T)} \qquad (30)$$

where $M_0 = \sup_{x,z\in\chi} \varepsilon^{'} S\varepsilon\left(0\right)$ and using inequality (27).

Now, we choose θ_1 and γ such that

$$M_0 \mathrm{e}^{-\alpha q_m T} \mathrm{e}^{4\frac{\beta}{\alpha}L_b T^*} \mathrm{e}^{-\alpha q_m \theta_1 (T^*-T)} \leq \gamma \qquad (31)$$

and (24) are satisfied simultaneously, which is possible since $\mathrm{e}^{-\mathrm{cte}\times\theta_1} < \frac{\mathrm{cte}}{\theta_1^{2n-2}}$ for θ_1 large enough.

The condition $2\theta_1\delta_t < \mu$ (from Lemma 7: $\theta_{max} = 2\theta_1$) is checked and δ_t shortened if needed (as all the parameters we use until now do not depend on δ_t).

We set $d \in \mathbb{N}^*$ such that $0 < d\delta_t < T < T^*$ and such that the condition $d \geq n-1$ is satisfied (we still can shorten δ_t) . Now that innovation is defined, so is the parameter λ_d^0 of Lemma 2. We design a function $\mathcal{F}$ as in Lemma 7 with $\Delta T = T - d\delta_t$ and $\gamma_1 = \frac{\lambda_d^0\gamma}{\beta}$.

We claim that there exists $\xi \leq T^*$ such that $\tilde{\varepsilon}^{'} \tilde{S}\tilde{\varepsilon}\left(\xi\right) \leq \gamma$.

Indeed, if $\tilde{\varepsilon}^{'} \tilde{S}\tilde{\varepsilon}\left(\xi\right) > \gamma$ for all $\xi \leq T^*$ then $\tilde{\varepsilon}^{'} \tilde{S}\tilde{\varepsilon}\left(k\delta_t\right) > \gamma$ for all $k \in \{0, ..., k^*\}$ with $k^* = max\{k \in \mathbb{N}, k\delta_t \leq T^*\}$. Then thanks to Lemma 2:

$$\begin{aligned} \gamma < \tilde{\varepsilon}^{'} \tilde{S}\tilde{\varepsilon}\left(k\delta_t\right) \leq \quad & \beta \left\|\tilde{\varepsilon}\left(k\delta_t\right)\right\|^2 \\ \leq \quad & \beta \left\|\varepsilon\left(k\delta_t\right)\right\|^2 \quad \leq \frac{\beta}{\lambda_d^0}\mathcal{I}_d\left(k\delta_t + d\delta_t\right) \end{aligned}$$

Therefore, $\mathcal{I}_{k+d,d} \geq \gamma_1$ for all $k \in \{0, ..., k^*\}$ hence $\mathcal{I}_{k,d} \geq \gamma_1$ for all $k \in \{d, ..., k*\}$. Hence $\theta\left(t\right) \geq \theta_1$ for $t \in [T, T^*]$ which gives a contradiction ($\tilde{\varepsilon}^{'} \tilde{S}\tilde{\varepsilon}\left(T^*\right) \leq \gamma$) thanks to (30) and (31).

Finally, for $t \geq \xi$,

$$\|\epsilon(t)\|^2 \leq (2\theta_1)^{2n-2}\|\tilde{\epsilon}(t)\|^2 \leq \frac{(2\theta_1)^{2n-2}}{\alpha}\tilde{\epsilon}^{'} \tilde{S}\tilde{\epsilon}(t)$$

which gives from (25),

$$\|\epsilon(t)\|^2 \leq \epsilon^* e^{-\alpha q_m (t-\xi)} \leq \epsilon^* e^{-\alpha q_m (t-T^*)}.$$

REFERENCES

[1] S. Ammar, J-C. Vivalda, "On the Preservation of Observability Under Sampling", *Systems and Controls Letters, Vol. 52;2004.*

[2] *J. S. Baras, A. Bensoussan, M. R. James, "Dynamic Observers as Asympotic Limits of Recursive Filters: Special Cases", SIAM J. Appl. Math,. 48, pp 1147-1158; 1988.*

[3] *G Besançon (Ed.), "Nonlinear Observers and Applications,LNCIS 363, Springer; 2007.*

[4] *E. Busvelle, J-P. Gauthier, "High-gain and Non High-gain Observers for Nonlinear Systems", in Contemporary Trends in Nonlinear Geometric Control Theory and Its Applications, World Scientific Publishing; 2002.*

[5] *N. Boizot, "Adaptive High-gain Extended Kalman Filter and Applications ", Ph.D. Thesis, to be defended.*

[6] *N. Boizot, E. Busvelle, J. Sachau, "High-gain Observers and Kalman Filtering in Hard Real-time", 9th Real-Time Linux Workshop, Linz, Austria; 2007.*

[7] *M. Boutayeb, D. Aubry, "A Strong Tracking Extended Kalman Observer for Nonlinear Discrete-Time Systems ", IEEE Transactions on Automatic Control, Vol. 44, No.8; 1999.*

[8] *F. Deza, E. Busvelle, J-P Gauthier, D. Rakotopara, "High-gain Estimation for Nonlinear Systems", System and Control Letters 18, pp 295-299; 1992.*

[9] *J-P. Gauthier, H. Hammouri, S. Othman, "A Simple Observer for Nonlinear Systems", IEEE Trans. Aut. Contr., 37, pp 875-880; 1992.*

[10] *J-P Gauthier, I Kupka, "Deterministic Observation Theory and Applications", Cambridge University Press, Cambridge, UK; 2001.*

[11] *R. E. Kalman, A new Approach to Linear Filtering and Prediction Problems, J. of Basic Eng., 82 (1), 3545; 1960.*

[12] *D. G. Luenberger, Observers for Multivariable Systems, IEEE trans. Aut. Contr., 11(2), pp 190-197; 1966.*

[13] *J. Picard, Efficiency of the Extended Kalman Filter for Nonlinear Systems with Small Noise, SIAM J. Appl. Math., 51 (3), pp 843-885; 1991.*