1

Adaptive-gain observers and applications

Nicolas Boizot[1] and Eric Busvelle[2]

[1] LASSY, Université du Luxembourg, 6 rue Coudenhove Kalergi, L-1359 Luxembourg nicolas.boizot@uni.lu
[2] LE2I, IUT d'Auxerre, Route des plaines de l'Yonne, 89000 Auxerre, France busvelle@u-bourgogne.fr

1.1 Introduction

We distinguish two kinds of observers for nonlinear systems which are used by scientists and engineers: empirical observers and converging observers.

The first class of observers are based on some approximation of the nonlinear system or approximation of a theoretical best estimation. The most common example is of course the extended Kalman filter. Although, for linear systems, the Kalman filter is a converging observer and an optimal observer for some quadratic cost function, the nonlinear version is based on a linearization of the nonlinear system in a neighborhood of its estimation. Hence, the extended Kalman filter is a good – almost optimal – local observer but it is not a globally converging observer. Intuitively, if the *a priori* estimation is far from the actual state value, the linearization around the estimate has no sense (Section 1.2.2).

There are a lot of empirical observers, based on neural networks, genetic algorithms, fuzzy logic, and so on. These observers are also based on an approximation of the process.

An other type of observers are based on the approximation of the exact solution. Indeed, setting the problem as a stochastic problem, the optimal solution is given by the Duncan-Mortensen-Zakaï (DMZ) equation. The solution of this nonlinear stochastic partial differential equation is the law of the state knowing observations. Hence, the conditional expectation of the state knowing observations can be expressed using the solution of the DMZ equation. However, this PDE equation is very complicated. There exist some algorithms in order to calculate an approximation of the solution, and therefore to obtain an approximate observer. For instance, some Monte-Carlo methods can be used in order to calculate the conditional density of probability of the conditional law. In this case, these methods are called particle filtering methods. It consists in the simulation (by Monte-Carlo methods) of several processes, which allows the calculation of the law of the state. The observation appears in the DMZ equation as a killing process. Although this approach has some

theoretical justifications (it converges when a finite parameter – the number of particles – goes to infinity), observers based on this approach are always approximative observers.

Although these empirical observers are not proved to converge, they are used by many engineers for many processes, including some critical processes. During normal operation, these observers are often very reliable and gives very good practical results.

The second class of observers are theoretically converging observers. In this book, we mainly discuss about high-gain observers. Nevertheless, there exist also some other classes of converging observers. Most of them only deal with a small class of nonlinear systems. Most of them have also some bad performances in presence of noise.

In this chapter, we will not speak about sliding observers, algebraic observers, or finite dimensional filters, but we will focus on high-gain observers, and their performances comparing to extended Kalman filter.

Our purpose is to present a uniform framework where nonlinear filtering, empirical observers and exponentially converging observers are compared. We mainly discuss about their similarities, and we propose an observer based on empirical observers (as those used by engineers), which is an exponentially converging observer.

Despite the lack of theoretical justification, the extended Kalman filter (EKF) is one of the most famous algorithm used to estimate unknown state variables from measurements in dynamical nonlinear systems. It is also used to estimate unknown constant or slowly varying parameters in linear systems and sometimes to perform failure detection. In this last case, it is necessary to quantify the efficiency of the EKF with time. This task is usually based on the innovation process, which is the integrated difference between actual measurements and predicted measurements. The innovation process can be monitored, and a large value of the innovation can be used to send an alarm or to switch from an old model to a new one. It can also be used to estimate the noise entering into the process or to estimate the measurement noise.

The empirical EKF is even used for critical processes. Therefore, in order to increase the performance and the reliability of the EKF several engineers and researchers already tried to develop an adaptive version. Using innovation and state estimation, it seems possible to estimate parameters that characterize the state of the process. These parameters can then be used to adapt the gain matrix by online automatic tuning of some of the covariance matrices used in the computation of the gain matrix. These kind of adaptive EKF are empirical but seem to have nice behavior compared to the EKF.

Because of the difficulty to ensure robustness when adaptive quantity is continuously updated, some authors used an adaptive algorithm based on switching between several models. For instance, in [33], authors have developed an application on a highly critical process (from robustness point of view). They proposed to switch between two covariances matrix Q_1 and Q_2 depending on the state of the process.

There exist many papers dealing with adaptive observers and adaptive extended Kalman filtering especially in the GPS and DGPS community, see [22, 12, 26]. In [12] for instance, authors present an adaptive extended Kalman filter using innovation in order to adapt Q and R matrices, exactly in the same spirit than in the present chapter, except that they do not give any theoretical proof. Nevertheless, the need for this kind of observer is clearly established.

In those papers, adaptation of the filter is done using empirical rules (genetic algorithms [35], neural networks [44], statistics [33]...), and no proofs are given. But in all cases, efficiency of the adaptive observer is highlighted. Let us remark that for neural networks based extended Kalman filters (N-EKF), the system is splitted into a linear part and a nonlinear part, and the extended Kalman filter is applied to the nonlinear part, which is approximated by neurons. The weights of neurons can be calculated using EKF, making the algorithm adaptive. In this case, some proofs can be established, but only if the neural network can approximate the system.

An intuitive theoretical justification of adaptive gain is based on the high gain observer theory. It has been shown from a long time ([17]) that high gain observers have very nice theoretical properties. The first one is that they required to study the observability property of the model. This study prevents from developing an observer for a non-observable system. But high gain observers are also exponential observers: one can prove the convergence of the high gain observer. In our opinion, the convergence property is a minimum requirement for an observer which is used on some critical processes, and sometimes as a diagnostic tool. Therefore, it is a good idea to adapt the gain of observers in the following way:

- use an EKF when the estimation is close to the true state, because EKF is a good (optimal) local observer (as already stated) and
- use a high-gain observer when large perturbations occur, because these observers are nonlinear converging observers.

In [14, 15, 20], the high-gain extended Kalman filter (HG-EKF) has been introduced. Compared with the Luenberger observer, HG-EKF is also an exponentially converging observer, but with the property that it is more efficient in the presence of noise. Indeed, the high sensitivity of high-gain observers is a well known drawback: the high gain ensures convergence but also increases noise effects. In [8], a new algorithm, based on classical and high-gain EKF, has been developed. This algorithm is based on a theoretical result, which states that a time-dependant HG-EKF, which is asymptotically equivalent to a classical EKF, may be an exponentially converging observer, if the transition from HG-EKF to EKF is slow enough. But this result is based on a time-dependant observer and, in order to make its convergence property persistent, it is necessary to use several observers and to switch from one to another, depending on the innovation process. Although it is an efficient observer, as shown in the reference above, but also in [9, 10], it is rather complicated and CPU intensive. Moreover, even if the final algorithm can be considered as an adaptive

high-gain extended Kalman filter (AG–EKF), its implementation is far from classical observers as used by engineers.

In this chapter, we will present a time-independant adaptive-gain extended Kalman filter. The adaptation of θ will depend on the innovation process.

As usual for the HG-EKF, the parameter θ appears in the Riccati equation of the Kalman filter, and more precisely in the matrix Q, denoted Q_θ. But in this new case, the high-gain parameter appears also in the matrix R (denoted R_θ), as in [12] (for a practical application). It is the first difference with result in [8]. The second difference is that θ may increase if the innovation is high and decrease if the innovation is low. This idea is the basis of practical applications: it is also the cornerstone of the proof of the theorem.

Before considering extended Kalman filtering, we will present in the next section some results concerning nonlinear filtering. A nonlinear filter is similar to a nonlinear observer, in the sense that it is supposed to estimate the state of a system given some measurements. But nonlinear filtering deals with stochastic equations. In the deterministic case, one have in mind that the model approximates the system, that some unmodelized and unmeasured perturbations can enter continuously into the system, and that measurements are corrupted by noise. Therefore, an observer should be robust to these perturbations. In the filtering problem, these perturbations are taken into account in the synthesis of the algorithm. Hence, the stochastic approach seems to be more adapted to the problem, which is better defined (and the stochastic problem is completely solved by the DMZ equation).

As we will see however, both approaches yields to similar tools. In fact, the main difference between the two theories is the observability property:

- In the stochastic case, the system has not to be observable. A nonlinear filter can be developed even for unobservable systems since it gives only the conditional law of the state knowing observations. Typically, an observable system gives rise to an unimodal law.
- In the deterministic case, an observer has no sense for a non observable system (except perhaps if the system is globally asymptotically stable in which case the model itself is a – slow – observer).

The "nonlinear filtering" section may be read even by a reader which is not specialist in probability. It can also be omitted by a reader which is not interested by the filtering/observation comparison.

1.2 Nonlinear filtering

1.2.1 Duncan-Mortensen-Zakaï equation

We study the observer problem in a stochastic setting. Let us consider the following stochastic system

$$\begin{cases} dX(t) = f(X(t), u)\, dt + Q^{\frac{1}{2}} dW(t) \\ dY(t) = h(X(t), u)\, dt + R^{\frac{1}{2}} dV(t) \end{cases} \tag{1.1}$$

where

- $X(t) \in \mathbb{R}^n$, $X(0)$ being a random variable, $Y(t) \in \mathbb{R}^p$, and u is a $\mathbb{R}^d$–valued measurable function,
- $W(t)$ and $V(t)$ are two independent Wiener processes (also independent from $X(0)$).

In this chapter, we will omit to specify the time variable whenever no confusions are possible, writing X instead of $X(t)$.

Therefore,

$$E\left[\left(Q^{\frac{1}{2}} W(t)\right)\left(Q^{\frac{1}{2}} W(t)\right)'\right] = Q.t$$

(where M' denotes the transpose of a matrix M) so Q is the covariance matrix of the state noise, and R is the covariance matrix of the measurement noise (the notation $Q^{\frac{1}{2}}$ represents the Cholesky decomposition of Q, also called square root of Q).

In this section, we denote $X(t)$ a process or random variable and $x(t)$ its realization, that is $x(t) = X(t)(\omega)$.

$X(0)$ is supposed to be an $L^2(\mathbb{R}^n)$ random variable independent from W and V. For simplicity, we will assume that this random variable admits a density function, denoted $p(0,x) = \frac{dP(\{X(0)\leq x\})}{dx}$.

Considering equations in the Ito sense, if f is a Lipschitz function w.r.t. x with a Lipschitz constant independent of u, then the system (1.1) admits a unique solution.

In this stochastic context, the observer problem is an estimation problem: we want to calculate the best estimation of $X(t)$ knowing measurements Y from 0 to t, denoted by the σ–algebra $\mathcal{F}_t^Y$. Hence, we want to calculate the conditional expectation $E\left[X(t) \mid \mathcal{F}_t^Y\right]$, or more generally $E\left[\phi(X(t)) \mid \mathcal{F}_t^Y\right]$ for any test function ϕ. Finally, this is equivalent to calculate the conditional law of $X(t)$ knowing $\mathcal{F}_t^Y$.

We assume that this law admits a density denoted by $p(t,x)$, *i.e.* the conditional law is absolutely continuous with respect to Lebesgue measure (this restrictive assumption is not necessary but it simplifies some formulas, especially the DMZ equation). Then, $p(t,x)$ is the solution of the well known Duncan–Mortensen–Zakaï (DMZ) equation. We will not explain this equation here: it is a stochastic partial differential equation, which has to be regularized before to be used, and which is difficult to use for practical problems, especially if n is large (see [37] for a clear statement of the DMZ equation).

The DMZ equation has been used in several ways:

- First, this equation may be simplified in some very special cases. One of them is the linear case, where the solution of the DMZ equation is the Kalman filtering equation. There exists also some nonlinear cases where

the DMZ equation gives a computable solution, for instance for systems which are linearizable up to a change of coordinates, or an immersion. In these cases, it is of course a very good approach to build an optimal observer.

- Second, despite its complexity, the (regularized version) of the DMZ equation can be approximately solved, for instance using Monte-Carlo methods. In this context, Monte–Carlo methods are called particle methods. The main idea is to approximate the initial law of $X(0)$, given by its density $p(0,x)$, by a set of "particles", *i.e.* a set of independent random variables $X_i(0)$ such that

$$p(0,x) \simeq \sum_{i=1}^{N} \delta_{X_i(0)}$$

where δ_x denotes the Dirac measure at x. The notation $\simeq$ will be precisely defined in Theorem 2.

The principle of a particle method is then to approach the probability law of $X(t)$ knowing $\mathcal{F}_t^Y$ by a (weighted) sum of Dirac measures at points $X_i(t)$. When applied to filtering, this just consists in approaching the law of the current state knowing observations by means of a particular weighted sum of Dirac distributions. This kind of method is well adapted to the case in which the dimension of the state is large, because in this case one usually uses the Monte-Carlo method to compute the conditional expectation

$$E\left[\phi(X(t)) \mid \mathcal{F}_t^Y\right] = \int \phi(x) p(t,x) dx$$

and this method requires a sample of the law $p(t,x)$ which is given by $X_i(t)$, $i = 1, \ldots, N$.

To characterize a particle method, it is sufficient to give some rules such as

- how to calculate weights of particles (e.g. Dirac measures)
- how to move particles $X_i(t)$ in the state space

Let us give an example of a particle filtering. As we will see in next section, this algorithm have some similarities with the observer construction (Section 1.3.3), although it has been obtained by a totally different way.

We will study the nonlinear filtering problem with linear discrete-time observation, that is to say, the second equation in (1.1) is replaced by

$$Y_k = CX(t_k) + R^{\frac{1}{2}} V(k) \tag{1.2}$$

where $(t_k)_{k\in\mathbb{N}}$ is the sample time and $(V(k))_{k\in\mathbb{N}}$ is an independent (w.r.t. W and $X(0)$) Gaussian white noise. The limitation to a linear observation function is not necessary but is a simplification when one wants to implement this

algorithm. The choice of discrete-time observation simplifies the mathematical background necessary to define the DMZ equation. Indeed, in this case, the conditional density $p(t,x)$ is given by the discrete version of the DMZ equation:

$$p(t_k,x) = \frac{1}{f_{Y_k}^{Y^{k-1}=y^{k-1}}(y_k)} f_{Y_k}^{X(t)=x}(y_k) \int_{\underline{X}} f_{X(t_k)}^{X(t_{k-1})=\xi}(x) p(t_{k-1},\xi) d\xi \quad (1.3)$$

where notations

- $f_{Y_k}^{X(t)=x}(y_k)$ represents the conditional density of Y_k knowing $X(t)=x$;
- $f_{X(t)}^{X(s)=\xi}(x)$ represents the conditional density of $X(t)$ knowing $X(s)=\xi$;
- $f_{Y_k}^{Y^{k-1}=y^{k-1}}(y_k)$ represents the conditional density of Y_k knowing Y between time 0 and time t_{k-1} is equal to $(y_0,\ldots,y_{k-1})$ so that for instance,

$$p(t,x) = f_{X(t)}^{Y^k=y^k}(x)$$

Equation (1.3) is nothing else than the Bayes formula applied to the problem.

Remark 1. We point out that the DMZ equation (1.3) gives an exhaustive information on $X(t)$ knowing all informations available at time t. Hence it gives the best possible estimate and, if the system is observable (Definition 1), it is a very good observer.

As usual with equations describing evolution of a density of probability, the un–normalized version of the DMZ is more tractable: (1.3) is equivalent to

$$q(t_k,x) = f_{Y_k}^{X(t)=x}(y_k) \int_{\underline{X}} f_{X(t_k)}^{X(t_{k-1})=\xi}(x) q(t_{k-1},\xi) d\xi \quad (1.4)$$

with

$$p(t_k,x) = \frac{q(t_k,x)}{\int_{\underline{X}} q(t_k,\xi) d\xi}$$

There are several ways to solve the un–normalized DMZ equation using particle methods. The first way is to recognize the composition/rejection theorem in this formula ([27]), and therefore to consider this equation as a simulation formula, which is the basis of a Monte-Carlo method. The algorithm consists in simulating the process (by "particles" Z_i) and killing some of them thanks to measurements (the "bad" particles). At a time $t_k < t \leq t_{k+1}$, the number of particles which are still alive is a random variable $N(k)$. If this random number is large enough, the conditional density is approximated by

$$p(t,x) \simeq \sum_{i=1}^{N(k)} \delta_{Z_i(t)}(x)$$

This approach can not be applied exactly as explained here, since $N(k)$ is a decreasing integer which goes almost surely to 0 (each measurement kill particles). In order to obtain a more efficient algorithm, one usually consider a weighted sum of Dirac measures.

Let us introduce coefficients $a_i(t) \in [0,1]$. These numbers represent the degree of confidence in each particle, and replace binary coefficients 1 (the particle is alive) or 0 (the particle is dead). As for the DMZ equation itself, we consider an un–normalized set of coefficients $b_i(t) \in \mathbb{R}^+$ such that

$$a_i(t) = \frac{b_i(t)}{\sum_{j=1}^{N} b_i(t)}$$

We consider an algorithm $\mathcal{P}$ which describes the trajectory of particles $z_i(t)$ and weight coefficients $b_i(t)$. The law truncated at n particles given by $\mathcal{P}$ is denoted as $P_n(t)$ ($dP_n(t) = p_n(t,x)\,dx$) and defined by

$$P_n(t) = \frac{\sum_{i=1}^{n} b_i(t)\,\delta_{z_i(t)}}{\sum_{i=1}^{n} b_i(t)} = \sum_{i=1}^{n} a_i(t)\,\delta_{z_i(t)}$$

Algorithm 1 *Initialization*

$z_i(0)$ *is the realization of a random variable with respect to the initial law* $p(0)$*;*

$b_i(0) = 1$*;*

Loop

$z_i(t_k)$ *is a Gaussian variable with respect to* $f_{X(t_k)}^{X(t_{k-1})=z_i(t_{k-1}),Y_k=y_k}$*;*

$b_i(t_k)$ *is defined by*

$$b_i(t_k) = b_i(t_{k-1})\ \ f_{Y_k}^{X(t_{k-1})=z_i(t_{k-1})}(y_k)$$

Let us remark that this algorithm is easy to implement on a computer, in particular on a parallel computer.

Theorem 2. *Let us consider the system*

$$\begin{cases} dX = f(X,u)\,dt + Q^{\frac{1}{2}}\,dw(t) \\ Y_k = CX(t_k) + R^{\frac{1}{2}}V(k) \end{cases}$$

and $P(t)$ *being the conditional law of* $X(t)$ *knowing* $\mathcal{F}_t^Y$*. If* $P_n(t)$ *represents the law given by the algorithm* $\mathcal{P}$ *with* n *particles, then we have*

$$P_n(t) \to P(t) \ \text{ as } n \to \infty \text{ weakly almost surely}$$

Remark 2. This theorem is true at t fixed. It is never true for any t. In order to obtain an asymptotic result (as in observer theory), it is necessary to add some correlations between particles. This is particularily simple here (see [39]).

In order to illustrate this theorem, we consider a continuous stirred tank reactor (CSTR). The dimensionless form of the model is:

$$dX_1 = \left(-X_1 + D_A(1-X_1)\exp\{\frac{X_2}{1+X_2/\gamma}\}\right)dt + dW_1$$
$$dX_2 = \left(-X_2(1+\beta) + H_a D_a(1-X_1)\exp\{\frac{X_2}{1+X_2/\gamma}\} + \beta u\right)dt + dW_2$$

where W_1 and W_2 are two independent Wiener processes. X_1 is the reactant concentration and X_2 is the temperature into the tank. We suppose that X_2 is measured in discrete time and that we want to control X_1 using the control variable u. The system can also be written in the following generic form

$$\begin{aligned} X(t_{k+1}) &= X(t_k) + \int_{t_k}^{t_{k+1}} f(X(s))ds + \int_{t_k}^{t_{k+1}} BdW(s) \\ Y_k &= CX(t_k) + V_k \end{aligned}$$

with $C = \begin{pmatrix} 0 & 1 \end{pmatrix}$. We suppose that W is a two-dimensional Wiener process and that V_k is a Gaussian process independent of W and with covariance R. We propose the following discretization scheme for the continuous-time equation

$$X(t_{k+1}) = \Phi(t_k, t_{k+1}, X(t_k)) + \frac{\partial \Phi(t_k, t_{k+1}, X(t_k))}{\partial x} B\sqrt{t_{k+1}-t_k}W_k$$

where $\Phi(s,t,x)$ is the solution of

$$\begin{cases} \frac{dx(t)}{dt} = f(x(t)) \\ x(s) = x \end{cases}$$

at time t.

The right-hand part of this scheme is the first order development of

$$\Phi(t_k, t_{k+1}, X(t_k) + \int_{t_k}^{t_{k+1}} BdW(s))$$

which comes naturally from the diffusion equation. A classical theorem of probability, see for instance [21], shows that this scheme converges in law to the solution of the diffusion equation when the step of the discretization goes to zero.

Our main goal is to estimate the reactant concentration X_1 and its confidence intervals, in order to control as well as possible the CSTR.

If we solve the equations, we can see that for each particle $z(t)$ at time t and for each weight $b(t)$, we have, thanks to the algorithm of the theorem

- **Correction** at time t_k

$$\begin{cases} z\left(t_k\right) = z\left(t_k^-\right) \ + \ P\left(t_k^-\right) C^T (CP\left(t_k^-\right) C^T + R)^{-1} (y_k - Cz\left(t_k^-\right)) \\ \qquad\qquad + \ \left(P\left(t_k^-\right) - P\left(t_k^-\right) C^T (CP\left(t_k^-\right) C^T + R)^{-1} CP\left(t_k^-\right)\right) \bar{w}_k \\ P\left(t_k\right) = BB^T (t_{k+1} - t_k) \\ b\left(t_k\right) = b\left(t_{k-1}\right) \frac{\exp\left(-\frac{1}{2}(y_k - Cz\left(t_k^-\right))^T (CP_{t_k^-} C^T + R)^{-1} (y_k - Cz\left(t_k^-\right))\right)}{\sqrt{2\pi . det(CP\left(t_k^-\right) C^T + R)}} \end{cases} \tag{1.5}$$

where $\bar{w}_k$ is a Gaussian white noise.

- **prediction** between t_k and t_{k+1}

$$\begin{cases} \frac{dz}{dt} = f(z\left(t\right)) \\ \frac{dP}{dt} = f^*(\xi\left(t\right)) P\left(t\right) \ + \ P\left(t\right) f^*(\xi\left(t\right))' \\ \frac{db}{dt} = 0 \end{cases} \tag{1.6}$$

1.2.2 Extended Kalman filter

The previous algorithm is CPU-time consuming and rather complicated to implement, especially in the linear case. Indeed, for a linear system, there exist a very simple and famous solution. Let us consider the following linear system:

$$\begin{cases} dX = \left(A\left(t\right) X + B\left(t\right) u\right) \, dt + Q^{\frac{1}{2}} dW\left(t\right) \\ dY = C\left(t\right) X dt + R^{\frac{1}{2}} dV\left(t\right) \end{cases} \tag{1.7}$$

with $X\left(0\right)$ a random variable with Gaussian law $\mathcal{N}\left(m_0, P_0\right)$, the DMZ equation reduces itself to the well-known Kalman filter. More precisely, solving the DMZ equation yields to the following result: the conditional law of $X\left(t\right)$ knowing $y\left(s\right)$ from 0 to t $(\mathcal{F}_t^Y)$ is the Gaussian law $\mathcal{N}\left(z\left(t\right), P\left(t\right)\right)$ where, for an output trajectory $y\left(t\right)$, $z\left(t\right)$ and $P\left(t\right)$ are the solutions of the finite-dimensional system of ordinary differential equations:

$$\begin{cases} dz \ = \left(A\left(t\right) z + B\left(t\right) u\right) \, dt + PC\left(t\right)' R^{-1} (dy - C\left(t\right) z \, dt) \\ \frac{dP}{dt} = A\left(t\right) P + PA\left(t\right)' + Q - PC\left(t\right)' R^{-1} C\left(t\right) P \end{cases} \tag{1.8}$$

with $z\left(0\right) = m_0$ and $P\left(0\right) = P_0$. Therefore, $z\left(t\right) = E\left[X\left(t\right) \mid \mathcal{F}_t^Y\right]\left(\omega\right)$ is the best estimation of $X\left(t\right)$ knowing measurements up to time t. When applied to a deterministic observable linear system, Q and R being considered as tuning parameters, the Kalman filter is called the Kalman observer. The observable property is not crucial in the stochastic case since the conditional law is defined even for non observable systems. But the observability property implies that the covariance matrix of the conditional expectation of $X\left(t\right)$ knowing $Y\left(s\right)$, $0 \leq s \leq t$ is bounded.

In the deterministic case, this property is crucial. Recall also that, for linear systems, observability does not depends from inputs.

The Kalman filter/observer algorithm has been used for long by engineers for linear systems. For nonlinear systems, engineers introduced and successfully used the extended Kalman filter (EKF), either in its stochastic or its

deterministic form. The EKF is just the standard Kalman filter for linear time-dependant systems, applied to the linearized system along the estimate trajectory. The EKF is the heart of our approach.

Let us consider a nonlinear system

$$\begin{cases} dX = f(X,u)\,dt + Q^{\frac{1}{2}} dW(t) \\ dY = h(X)\,dt + R^{\frac{1}{2}} dV(t) \end{cases} \tag{1.9}$$

where f and h are smooth Lipschitz functions, the linear Kalman filter does not apply anymore, and the exact solution should be obtained by solving the DMZ equation. But if one want an approximated solution, it is very common to consider the first order approximation of the previous system. The right way to do this is to consider an *a priori* solution $\hat{x}(t)$ of the deterministic system associated to (1.9) and to use the Kalman filter to estimate the first order difference $\delta x(t) = x(t) - \hat{x}(t)$ between the *a priori* solution and the estimated solution. This approach yields to the following first order Kalman filter, for a given output trajectory:

$$\begin{cases} \frac{d(\delta x)}{dt} = f^*(\hat{x},u)\delta x + Ph^*(\hat{x},u)' R^{-1}(y(t) - h(\hat{x},u)) \\ \frac{dP}{dt} = f^*(\hat{x},u)\,P + Pf^*(\hat{x},u)' + Q \\ \qquad -Ph^*(\hat{x},u)' R^{-1} h^*(\hat{x},u)\,P \end{cases} \tag{1.10}$$

where f^* and h^* are the Jacobian of f and h w.r.t. x respectively. But this approach has a major weakness: the choice of the *a priori* solution $\hat{x}(t)$ is not obvious if there is no precise *a priori* information on the initial state. This is usually the case, especially in the deterministic case, since the only missing information on the system is precisely the initial state. Moreover, if one make a bad choice of $\hat{x}(t)$, the first order equation has no significant meaning since the actual state is far from the initial guess. At the opposite, if $\delta x(0)$ is small (that is the *a priori* solution is closed to the actual solution, at least at time 0), then $\hat{x}(t) + \delta x$ will be a good approximation of the optimal filter, when state and measurement noises are small ([38]).

To overcome this difficulty, engineers have an attractive idea: to replace the *a priori* solution by the estimated solution at current time. The main advantage of this approach is that the estimated solution is supposed to be closed to the actual solution, hence the first order approximation should be small and hence the linear approximation should be a good approximation. This remark yields to the extended Kalman filter:

$$\begin{cases} \frac{dz}{dt} = f(z,u) + Ph^*(\hat{x},u)' R^{-1}(y(t) - h(z,u)) \\ \frac{dP}{dt} = f^*(z,u)\,P + Pf^*(z,u)' + Q \\ \qquad -Ph^*(z,u)' R^{-1} h^*(z,u)\,P \end{cases} \tag{1.11}$$

where z is the estimated state. Here again, if P_0, Q and R are small, this filter is closed to the optimal filter (see all works of Picard, [38] for instance).

In a deterministic context, the extended Kalman filter is a converging local observer (see [4, 8]), that is if $z(0) \simeq x(0)$ then $z(t) - x(t) \longrightarrow 0$

as $t \longrightarrow +\infty$ (exponentially). Nevertheless, the extended Kalman filter has no global converging properties. Indeed, it is well known that, if the initial guess $z(0)$ is far from $x(0)$, the extended Kalman filter may not converge. Moreover, the mathematical study of (1.11) is difficult because it has no clear mathematical meaning: it is not a first order approximation of a nonlinear object around a given trajectory. In other words, the behavior of (1.11) is not intrinsic and depends on a choice of coordinates. Hopefully, this mathematical difficulty will give us a way to chose a good system of coordinates and to prove some convergence results, thanks to this crucial choice of coordinates.

To conclude, the EKF is very efficient in a lot of practical problems. It is used as a filter or as an observer in many various systems. From a theoretical point of view, it is not an optimal filter (it differs from the DMZ equation). Nevertheless, when the system has some observability properties, it has very nice local properties: in the stochastic case, it is a good filter when noises are small (see [38]) and in the deterministic case, it is a local observer ([4, 8]).

1.2.3 Continuous-discrete stochastic systems

Before considering deterministic systems and observers, let us recall a result concerning discrete measurements. Continuous-discrete time are very common in practise: the nonlinear differential equation describes a mechanical, physical or chemical process. Therefore, it is a continuous time system. But measurements are usually sampled at times t_k. Therefore, the system can be written

$$\begin{cases} dX(t) = f(X(t), u(t))dt + dW(t) \\ y_k = h(X(t_k)) + V(k) \end{cases} \tag{1.12}$$

where h is a differentiable function from the state space to $\mathbb{R}^p$.

For this system, the EKF has two set of equations: the correction step which is applied at each measurement time and the prediction step which is used to predict the system according to the model.

Correction step

$$\begin{cases} Z\left(t_k^+\right) = Z(t_k) + G(k)\left(y_k - h(Z(t_k))\right) \\ G(k) = P(t_k) h^*(Z(t_k))' \left(h^*(Z(t_k)) P(t_k) h^*(Z(t_k))' + R\right)^{-1} \\ P\left(t_k^+\right) = \left(I - G(k) h^*(Z(t_k))\right) P(t_k) \end{cases} \tag{1.13}$$

Prediction step

$$\begin{cases} \frac{dZ}{dt} = f(Z, u) \\ \frac{dP}{dt} = f^*(Z, u) P + P f^*(Z, u)' + Q \end{cases} \tag{1.14}$$

These equations presents some similarities with equations (1.5,1.6). As we will see in the end of Section 1.3.4, if the system is observable, then equations

(1.13,1.14) may gives an observer. In the non observable case, one should use (1.5,1.6).

Although this kind of model is closer to the practical case, it is less used than continuous-time systems. The main reason is a practical one: the sampled time is usually chosen small enough w.r.t. time constants of the process. Therefore, the continuous EKF can be applied. Sometimes (for very fast processes or for slow measurement devices), the sampled time is a constraint and can not be neglected. In this case, continuous-discrete EKF should be applied.

1.3 Nonlinear observers

1.3.1 Canonical form of observability

From now, we study deterministic nonlinear systems of the general form

$$\Sigma \left\{ \begin{array}{l} \frac{dx}{dt} = f(x,u) \\ y = h(x,u) \end{array} \right. \tag{1.15}$$

on a smooth n–dimensional manifold X, $y \in \mathbb{R}^p$, $u \in U$, subset of $\mathbb{R}^d$. We want to develop an observer. Our approach is closely related to observation theory, as explained in the book from Gauthier and Kupka [20], which is itself a summary of the papers [16, 17, 18, 19, 32].

This theory leads to the consideration of systems under the normal form (1.21), or similar multi-output normal forms. Here, by "observability", we mean "observability for every fixed input function $u(t)$". For details, see [20].

In this introduction part, we summarize the main observability results of the observation theory developed in [20].

First of all, the state–output mapping $PX_{\Sigma,u}$ is the function $x(0) \longrightarrow (y(t))_{t\geq 0}$. In this definition (and the following ones), we do not speak about explosion times, in order to simplify the notations.

Definition 1. *The system (1.15) is said uniformly* ***observable****, or just* ***observable****, w.r.t. a certain class $\mathcal{C}$ of inputs ($L^\infty(U)$ in most cases) if, for each $u(.) \in \mathcal{C}$, the state output mapping $PX_{\Sigma,u}$ is injective.*

This first definition is the natural definition of observability. Nevertheless, injectivity is not a very tractable property, since it is not stable (even for standard mappings between finite dimensional spaces -example: $x \to x^3, \mathbb{R} \to \mathbb{R}$). Therefore, in order to state results, we need a few other definitions. The uniform infinitesimal observability make the observable property stable.

Let us define the lift of Σ on TX, also called the first variation of Σ. Let us consider $T_X f : TX \times U \longrightarrow TTX$(the tangent bundle of TX) the tangent mapping of $f : X \times U \longrightarrow TX$ and $d_X h : TX \times U \longrightarrow \mathbb{R}^p$ the Jacobian of $h : X \times U \longrightarrow \mathbb{R}^p$. Then

$$T\Sigma \begin{cases} \frac{d\xi}{dt} = T_X f(\xi, u) = T_X f_u(\xi) \\ \eta = d_X h(\xi, u) = d_X h_u(\xi) \end{cases} \tag{1.16}$$

The state–output mapping of $T\Sigma$ is denoted by $PTX_{\Sigma,u}$. It is also the first order approximation of $PX_{\Sigma,u}$ denoted $TPX_{\Sigma,u}$.

Definition 2. *System Σ is said uniformly infinitesimally observable if, for each $u(.) \in L^\infty(U)$, each $x_0 \in X$, all the tangent mappings $TPX_{\Sigma,u}|x_0$ are injective.*

Remark 3. This definition of observability is stable in the sense of discretization: if a system is uniformly infinitesimally observable, its continuous–discrete version (1.12) remains uniformly infinitesimally observable for a sampling time small enough. It is not the case for a system which is only observable (see [2]).

The two following definitions are an other way to define observability in a stable way. Note that these definitions are important for practical purpose, since they give a way to prove observability for nonlinear systems.

Definition 3. *System Σ is said differentially observable (of order k) if for all $j^k\hat{u}$, the extension to k-jets mapping*[3] *$\Phi_k : x_0 \to j^k\hat{y}$; $X \to \mathbb{R}^{km}$ is injective.*

Definition 4. *System Σ is said strongly differentially observable (of order k) if for all $j^k u$, the extension to k-jets mapping $\Phi_{k,j^k u} : x_0 \to j^k y$; $X \to \mathbb{R}^{km}$ is an injective immersion*[4]*.*

Clearly, strong differential observability implies differential observability, which implies observability for the C^∞ class, (and L^∞-observability).

It is also a consequence of the theory that for analytic systems, uniform infinitesimal observability implies observability of the restrictions of (1.15) to small open subsets of X, the union of which is dense in X.

The main result concerning observability of systems 1.15 is that, depending from the number of outputs w.r.t. the number of inputs, the property may be generic or not generic. More precisely, we distinguish two cases:

1. More measurements than control inputs ($p > d$): in that case, observability is a generic property, and generically, a system can be put globally under a normal form similar to (1.21), but the dimension of the state in the normal

[3] k-jets $j^k u$, of smooth functions u at $t = 0$ are defined as

$$j^k u = (u(0), u'(0), ..., u^{(k-1)}(0)).$$

Then, for a smooth function u and for each $x_0 \in X$, the k-jet $j^k y = (y(0), y'(0), ..., y^{(k-1)}(0))$ is well define: this is the k-jets state-output mapping Φ_k.

[4] immersion means that all the tangent mappings $T_{x_0}\Phi_{k,j^k\hat{u}}$ to this map, have full rank n at each point

form is bigger than the dimension of the state of the original system: it is at most double plus one. Also, the control in the normal form contains a certain number of derivatives of the control of the initial system. But this is more or less unimportant for observation problems, where the control, and hence its derivatives, are known.
Hence, if $p > d$, and for sufficiently smooth inputs, generic systems are very good from the point of view of observability.

2. Less or same number of measurements than control inputs ($p \leq d$): in that case observability is a non generic property. It is even a property of infinite codimension. This high degeneracy leads to the fact that, in the control affine case, all observable systems can be put locally under normal forms similar to (1.21) (with $a_i = 1$, $i = 1, ..., n$).
In the analytic case $p = 1$, $d \geq 1$, we can be more precise. If (1.15) is uniformly infinitesimally observable, then locally almost everywhere on X, the system (1.15) can be put in the form

$$\begin{cases} y = h(x_1, u) \\ \frac{dx_1}{dt} = f_1(x_1, x_2, u) \\ \frac{dx_2}{dt} = f_2(x_1, x_2, x_3, u) \\ \quad \vdots \\ \frac{dx_{n-1}}{dt} = f_{n-1}(x_1, x_2, .., x_n, u) \\ \frac{dx_n}{dt} = f_n(x_1, x_2, ..., x_n, u) \end{cases} \tag{1.17}$$

with

$$\frac{\partial h}{\partial x_1} \text{ and } \frac{\partial f_i}{\partial x_{i+1}}, i = 1, .., n-1 \tag{1.18}$$

does not vanishes on $V_x \times U$.
In the control affine case, where (1.15) can be written:

$$\begin{aligned} \dot{x} &= f(x) + \sum_{i=1}^{d} g_i(x) u_i \\ y &= h(x) \end{aligned} \tag{1.19}$$

then the canonical form of observability is

$$\begin{cases} y = x_1 \\ \frac{dx_1}{dt} = x_2 + \sum_{i=1}^{p} g_{1,i}(x_1) u_i \\ \frac{dx_2}{dt} = x_3 + \sum_{i=1}^{p} g_{2,i}(x_1, x_2) u_i \\ \quad \vdots \\ \frac{dx_{n-1}}{dt} = x_n + \sum_{i=1}^{p} g_{n-1,i}(x_1, x_2, .., x_{n-1}) u_i \\ \frac{dx_n}{dt} = \psi(x) + \sum_{i=1}^{p} g_{n,i}(x_1, x_2, .., x_{n-1}, x_n) u_i \end{cases} \tag{1.20}$$

These two results are very important since they allow us to restrict our study to systems of the form (1.17) and (1.20) (and also because of course, these results are based on a constructive diffeomorphism).

1.3.2 High-gain extended Kalman filter

We describe observers for nonlinear systems in canonical form of observability (1.20 and 1.21 below), on $\mathbb{R}^n$. The control space $\mathcal{U}_{\text{adm}}$, is supposed to be a closed subset of $\mathbb{R}^d$. In this section, **t**he observation is assumed to be single-valued: it is a u–dependant linear form on $\mathbb{R}^n$. This hypothesis is not necessary and our observers constructions also applies for multi–output systems. From an observability point of view, the multi–output case is a little bit more complicated since canonical form of observability are less natural. But from the observer point of view, except in section 1.3.4, the problem is exactly the same, since we simply apply some kind of EKF.

We consider systems of the form

$$\begin{cases} \frac{dx}{dt} = A(u)x + b(x,u) \\ y = C(u)x \end{cases} \tag{1.21}$$

where $A(u)$, $C(u)$ are matrices:

$$A(u) = \begin{pmatrix} 0 & a_2(u) & 0 & \cdots & 0 \\ & & a_3(u) & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & & a_n(u) \\ 0 & & \cdots & & 0 \end{pmatrix} \tag{1.22}$$

$$C(u) = (a_1(u), 0, ..., 0) \tag{1.23}$$

and where $a_i(.)$, $i = 1, ..., n$, are positive smooth functions, bounded from above and below:

$$0 < a_m \leq a_i(u) \leq a_M$$

Also, $b(x,u)$ is a smooth, $u-$dependant vector field, depending triangularly on x and compactly supported:

$$b(x,u) = \begin{pmatrix} b(x_1,u) \\ b(x_1,x_2,u) \\ \vdots \\ b(x_1,\ldots,x_n,u) \end{pmatrix} \tag{1.24}$$

These assumptions look very strong, but as we already seen, under either genericity hypotheses or observability hypotheses, for the purpose of synthesis of observers, it is sufficient to restrict to these systems, under the normal form (1.21) (or similar multi-output normal forms), and meeting these assumptions. In fact, this form generalizes the canonical form of observability (1.20) for control affine systems. We call (1.21) (together with (1.22–1.23)) the generalized canonical form of observability. There are several reasons to study (1.21) rather than (1.20):

- It is sometimes easiest to put the system into this form, using intuitive transformations, rather than a more restrictive normal form, the last transformation being based on Lie derivatives. This point will be illustrated in the application sections;
- Since we want to apply an EKF which use the model to filter noises, and a high–gain approach to kill the nonlinear part of the system, it is better to leave the largest part of the nonlinear system in A rather to put it in b. This technical point will be developed later;
- Last but not least, our observer construction still work for these systems.

However, this form does not include the canonical form of observability for systems (1.15) when the control is not affine. For those systems, there exist a change of coordinates that put the equivalent system (1.17) into a system of the generalized canonical form of observability (1.21) [10, 23]. For this, we just need to suppose that u admits a time derivative almost everywhere.

Consider a system (1.17) on $\mathbb{R}^n$, and set:

$$z = \Phi_u(x) = (h(x,u), L_f h(x,u), ..., L_f^{n-1} h(x,u)). \tag{1.25}$$

Let $K \subset \mathbb{R}^n$ be any fixed open relatively compact subset. We deal with semi-trajectories of Σ that remain in K, only. It follows from (1.18) that, for all $u \in U$, Φ_u is an injective immersion (this is easily checked by induction on the components of Φ_u). Therefore, Φ_u is a u–dependent diffeomorphism from K onto its image. Consider the image of the system (1.17) restricted to K by the time dependant diffeomorphism Φ_u. It is of the form:

$$\begin{cases} \frac{d\xi}{dt} = A\xi + g(\xi, u, \frac{du}{dt}) \\ y = \xi_1 \end{cases} \tag{1.26}$$

where A is the antishift matrix, and where g is smooth and depends in a triangular way of ξ.

Even if some technical difficulties remains in the general theoretical case (see [10] for a precise result), it is clear that the new system is of the form (1.21) except that we use explicitly $\frac{du}{dt}$, considered as a new input.

Thanks to this result, our observers (Sections 1.3.2, 1.3.3 and 1.3.4) applies to general uniformly infinitesimally observable systems.

Let us come back to the system (1.21) and its properties. The assumption $0 < a_m \leq a_i(u) \leq a_M$ is not more restrictive than $a_i(u) \neq 0$[5]. It just implies observability of systems in the normal form (1.21), by the following reasoning:

[5] Modulo a trivial change of variables, and the fact that the a_i being smooth, restricting to a compact subset of the set of values of control implies that we can find the a_m and a_M.

1. If the output $y(t)$ is known, the input being also known, the fact that $a_1(u)$ is nonzero implies that we can compute $x_1(t)$ from $y(t)$,
2. The fact that $a_2(u) \neq 0$ implies that we can compute $x_2(t)$ from the knowledge of $x_1(t)$,
3. By induction, we can reconstruct the whole state $x(t)$ from the knowledge of $y(t)$.

The compact support of b can be trivially achieved, by multiplying by a cut-off function, compactly supported, leaving the original vector field b unchanged on an arbitrarily large compact subset of R^n. Let us mention that this restriction to compact sets (unavoidable in a general observation theory), has not so important consequences: for instance, the high gain observers can be used in general for **global** dynamic output stabilization (again, see [20]).

The following results have been proved in [13, 14, 20].

We consider the equations of the extended Kalman filter (1.11), in which the covariance matrix Q depends on a real parameter θ, $\theta \geq 1$, in the following way:

$$Q_\theta = \theta \Delta^{-1} Q \Delta^{-1}$$

where

$$\Delta = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{\theta} & 0 & & \vdots \\ 0 & 0 & \frac{1}{\theta^2} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & \frac{1}{\theta^{n-1}} \end{pmatrix}$$

The EKF become the high-gain extended Kalman filter (HG-EKF):

$$\begin{cases} \frac{dz}{dt} = A(u)z + b(z,u) + PC'R^{-1}(y(t) - Cz) \\ \frac{dP}{dt} = (A(u) + b^*(z,u))P + P(A(u) + b^*(z,u))' \\ \qquad +Q_\theta - PC'R^{-1}CP \end{cases} \tag{1.27}$$

If $\theta = 1$, the HG-EKF is equivalent to the EKF. If θ is large, Q_θ is a large symmetric definite positive (s.d.p.) matrix and since it appears in the Riccati equation in a positive way, P will becomes large (in the s.d.p. sense). Therefore, the gain of the observer, namely $PC'R^{-1}$, will be large. This is why the observer (1.27) is called high-gain extended Kalman filter.

This observer has some very nice properties. From a practical point of view, since it is based on extended Kalman filtering approach, it is well designed for filtering noise using the model. Moreover, the HG-EKF is applied to a system written in the canonical form of observability. As a matter of fact, it clearly improves the convergence of the observer, both in simulation and in practical situations. Moreover, the parameter θ has a clear meaning and can be used to

tune efficiently the observer: if the observer is too slow, θ should be increased, and if the noise is not enough filtered, θ should be decreased.

This last point has also been validated from a theoretical point of view: the estimation error has arbitrarily large exponential decay, depending on θ. This holds whatever the initial error is, (that is, this is a global result). The theorem is the following:

Theorem 3. *For θ large enough and for all $T > 0$, the HG-EKF (1.27) satisfies for $t > \frac{T}{\theta}$*

$$\|z(t) - x(t)\| \leq \theta^{n-1} k(T) \left\| z\left(\frac{T}{\theta}\right) - x\left(\frac{T}{\theta}\right) \right\| \mathrm{e}^{-(\theta\omega(T)-\mu(T))\left(t-\frac{T}{\theta}\right)}$$

for some positive continuous functions $k(T)$, $\omega(T)$ and $\mu(T)$.

Remark 4. In a stochastic setting, the HG-EKG is a nonlinear filter with bounded variance ([13]).

1.3.3 High-gain and non high-gain extended Kalman filter

The EKF is a local converging observer, and has very good properties w.r.t. noise. It is close to the Kalman filter, which is an optimal solution to estimate the unknown state.

The HG-EKF is a globally converging observer. Moreover, it converges exponentially as fast as wanted, depending on the choice of the parameter θ.

The EKF cannot be used to estimate the state from a poor *a priori* estimation, or when large unmodelized perturbations occurs. The HG-EKF is designed to do this. This is the basis of the observer construction proposed in this section. More precisely, let us recall that:

1. if one sets θ to 1 in system (1.27) then one obtains the classical extended Kalman filter, which is a local optimal observer (in the sense explained above)
2. if θ is large enough then one obtains a high-gain observer, which is a global exponential observer.

The first application of this remark was presented in [8]: we just added the equation

$$\frac{d\theta}{dt} = \lambda(1 - \theta) \tag{1.28}$$

to the system (1.27). If $\theta(0) = \theta_0$ is large enough (and the parameter λ small enough) then we obtain an observer which is a high-gain observer for small time and which converges asymptotically to a classical extended Kalman filter. Hence we can expect its convergence since the observer should converge exponentially to the state (high-gain observer property) and then stays in a neighborhood of the state (since extended Kalman filter is a local observer).

Indeed this result has been proved in [8]. More precisely, the observer can be written (where Q_θ has be defined in the previous section):

$$\begin{cases} \frac{dz}{dt} = A(u)z + b(z,u) + PC'R^{-1}(y(t) - Cz) \\ \frac{dP}{dt} = (A(u) + b(z,u))P + P(A(u) + b^*(z,u))' \\ \qquad +Q_\theta - PC'R^{-1}CP \\ \frac{d\theta}{d\tau} = \lambda(1-\theta) \end{cases} \tag{1.29}$$

and the theorem says that the asymptotic behavior of the observer is the one of the extended Kalman filter, the "short term behavior" is the one of the HG-EKF. More precisely, let us denote $\varepsilon(t) = z(t) - x(t)$:

Theorem 4. *For all $0 \leq \lambda \leq \lambda_0$, ($\lambda_0$ small enough), for all $\theta(0) = \theta_0$ large enough, depending on λ, for all $S(0) = S_0 \geq c\ Id$, for all $K \subset \mathbb{R}^n$, K a compact subset, for all z_0 such that $\varepsilon(0) = z_0 - x(0) \in K$, the following estimation holds, for all $\tau \geq 0$:*

$$\begin{aligned} ||\varepsilon(\tau)||^2 &\leq R(\lambda, c)\, e^{-a\ \tau} ||\varepsilon_0||^2 \Lambda(\theta_0, \tau, \lambda), \\ \Lambda(\theta_0, \tau, \lambda), &= {\theta_0}^{2(n-1)+\frac{a}{\lambda}} e^{-\frac{a}{\lambda}\theta_0(1-e^{-\lambda\tau})}, \end{aligned} \tag{1.30}$$

Moreover the short term estimate

$$||\varepsilon(\tau)||^2 \leq \theta(\tau)^{2(n-1)} R(\lambda_0, c)\, e^{-(a_1\theta(T) - a_2)\tau} ||\varepsilon(0)||^2. \tag{1.31}$$

holds for all $0 \leq \tau \leq T$ and for all θ_0 large enough. $R(\lambda, c)$ is a decreasing function of c, and a, a_1 and a_2 are three positive constants.

Remark 5. (1.31) means that, provided that λ is smaller than a certain constant λ_0, and θ_0 is large in front of λ, the estimation error goes exponentially to zero, and can be made arbitrarily small in arbitrary short time. Moreover, in (1.30), the function $\Lambda(\theta_0, \tau, \lambda)$ being a decreasing function of τ, for all $\tau > 0$, $\lambda > 0$, $\Lambda(\theta_0, \tau, \lambda)$ can be made arbitrarily small, increasing θ_0, hence the observer is an exponential observer. Therefore, the observer is an exponential observer but the asymptotic rate of convergence does not depend on $\theta(t)$ (because $\theta(t) \simeq 1$), hence this observer does not converge as fast as we want after a given time τ.

The main drawback of this observer, as presented here, is that it converge exponentially for any initial condition only in the beginning, in order to estimate the initial state of the system: if a large perturbation occurs after time τ, this observer will have the same behavior as an EKF (since $\theta(t) \simeq 1$ for t larger that τ).

In order to construct a persistent observer, we should take into account this property and construct a time-dependant observer. The simplest way is to use several observers of the form (1.29), each one initialized at different times, and using some delays between each initialization. Thus we obtain several

estimations of the state, given by each one of the observers: the final estimation is the one corresponding to the observer that minimizes the innovation process. The whole construction is clearly explained in [8, 9] and we will recall the algorithm:

We consider a one parameter family $\{O_\tau, \tau \geq 0\}$ of observers of type (1.29), indexed by the time, each of them starting from S_0, θ_0, at the current time τ. In fact, in practice, it will be sufficient to consider, at time τ, a slipping window of time, $[\tau - T, \tau[$, and a finite set of observers $\{O_{t_i}, \tau - T \leq t_i \leq \tau\}$, with $t_i = \tau - i\frac{T}{N}$, $i = 1, ..., N$.

As usual, we call the term $I(\tau) = \hat{y}(\tau) - y(\tau)$, (the difference at time τ between the estimate output and the real output), the "innovation". Here, for each observer O_{t_i},we have an innovation $I_{t_i}(\tau)$.

Our suggestion is to take as the estimate of the state, the estimation given by the observer O_{t_i} that minimizes the absolute value of the innovation.

This is a very natural choice, according to probability theory (Section 1.2). The innovation process will also have an important role in Section 1.3.4, but we will consider its integral over small past time, which is another possible choice here.

Let us analyze what will be the effect of this procedure in a deterministic setting: after the transient part and if no unmodelized perturbation occurs, the best estimation is given by the oldest observer. Indeed, the oldest observer has converged and moreover, it is close to a classical EKF and therefore, it is more robust to measurement noise. But if a large perturbation occurs, making a jump on the state, the oldest (EKF) observer will no more converge. The youngest observer, which is a HG-EKF, will converge since it is in transient time (it's life time is less than τ). After an (arbitrary) short transient, the youngest observer will then give the best estimate and hence the smallest innovation.

This analysis is validated by our experience and we can even use these remarks to detect jumps, which correspond to abnormal operations or sensor failures.

Another remark is that this approach may be compared to a particle filtering method where the *a posteriori* estimation of the state is the maximum likelihood one. There exist several differences between these two algorithms and in fact, their use depends as usual on the observability study. If the system is not observable, a filtering approach should be used. If the system is observable, an observer can be used.

1.3.4 Adaptive gain extended Kalman filter

Here, we present a much simple observer. In place of equation (1.28), we introduce the equation

$$\frac{d\theta}{dt} = F(\theta, \mathcal{I}) \tag{1.32}$$

where

$$\mathcal{I} = \int_{t-T}^{t} \|y(s) - \bar{y}_{t-T}(s)\|^2 \, ds = \|y - \bar{y}_{t-T}\|^2_{L^2(t-T,t)} \tag{1.33}$$

is the innovation from time $t-T$ to current time t. More precisely, in (1.33), y represents the output, but $\bar{y}_{t-T}$ represents the prediction of the output from the state estimation at time $t-T$ (given by the observer, $z(t-T)$). Hence $\bar{y}_{t-T}(s)$ is the solution at time s of

$$\begin{cases} \frac{d\xi}{d\tau} = A(u)\xi(\tau) + b(\xi(\tau), u) \\ \xi(t-T) = Z(t-T) \\ \bar{y}_{t-T}(\tau) = C(u)\,\xi(\tau) \end{cases}$$

T is a tuning parameter, representing the length of the window used to calculate the innovation. In the following theorem, the function F will be chosen in the form

$$F(\theta, \mathcal{I}) = \lambda(1-\theta) + K(\theta_{\max} - \theta)\,\mathcal{I} \tag{1.34}$$

In fact, F can be chosen in a more general form. We will give a version of F that is better adapted in the presence of noise in the application part of this chapter (Section 1.5). Intuitively, the role of the function F is:

- to let θ decrease if the innovation is small, because in this case the observer has already converged and a Kalman-like observer will be sufficient to correctly estimate the state
- to let θ increase if the innovation is too large, because in this case, the observer gives a bad estimation of the state and θ has to be large enough in order to ensure convergence, thanks to the exponential property of high-gain observers.

Finally, the adaptive extended Kalman filter can be written

$$\begin{cases} \frac{dZ}{dt} = A(u)Z + b(Z,u) + S^{-1}C'R_\theta^{-1}(CZ - y(t)) \\ \frac{dS}{dt} = -(A(u) + b(Z,u))'S - S(A(u) + b^*(Z,u)) \\ \qquad\quad + C'R_\theta^{-1}C - SQ_\theta S \\ \frac{d\theta}{dt} = \lambda(1-\theta) + K(\theta_{\max} - \theta)\,\mathcal{I} \end{cases} \tag{1.35}$$

We define Q_θ and R_θ from Q and R thanks to the matrix

$$\Delta = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{\theta} & 0 & & \vdots \\ 0 & 0 & \frac{1}{\theta^2} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & \frac{1}{\theta^{n-1}} \end{pmatrix}$$

by $Q_\theta = \theta\Delta^{-1}Q\Delta^{-1}$ and $R_\theta = \theta^{-1}R$. Let us remark that this change of coordinates is different from the previous one (high-gain extended Kalman filters of Section 1.3.2 and Section 1.3.3).

Our main result is the following:

Theorem 5. *Let us consider a system in the canonical form of observability. We consider the adaptive-gain extended Kalman filter (1.35). Let us suppose that λ, K and $\theta_{\max}$ (in (1.34)) are three constant parameters such that λ is small enough, K is large enough, and $\theta_{\max}$ is large enough. Then, (1.35) is an exponentially converging observer.*

The proof is based on the following crucial lemma:

Lemma 1. *Let x_1^0, $x_2^0 \in \mathbb{R}^n$. Let us consider the outputs $y_1(t)$ and $y_2(t)$ with initial conditions respectively x_1^0 and x_2^0. The following condition (called persistant observability) holds:*

$$\forall T > 0 \quad \forall u \in L_b^1(\mathcal{U}_{\text{adm}}) \quad \exists \lambda_T > 0$$

$$\left\| x_1^0 - x_2^0 \right\| \leq \frac{1}{\lambda_T} \int_0^T \left\| y_1(\tau) - y_2(\tau) \right\| d\tau$$

The main difference between the previous observer is the fact that now, the matrix R depends on θ, which was not necessary when θ was only a decreasing parameter. The behavior of this adaptive–gain extended Kalman filter is illustrated on a DC–motor, in Section 1.5.

We point out that this AG-EKF is a very promising tool: it is a small modification of already existing adaptive–gain EKF proposed by engineers to improve the performance of EKF during abnormal operations. We propose the same approach in a theoretical framework, ensuring the exponential convergence of the algorithm.

1.3.5 Observer for continuous–discrete systems

As already explain in Section 1.2.2, practical problems may often be written in continuous-discrete form (1.12). There exist also some observability results concerning these systems. Let us suppose, for simplicity, that the sampling time is constant, *i.e.* $t_k = k\,\Delta t$.

A generalized canonical form of observability for these systems is the natural extension of the generalized canonical form of observability (1.21)

$$\begin{cases} \frac{dx}{dt} = A(u)x + b(x,u) \\ y_k = C(u)\,x(t_k) \end{cases} \tag{1.36}$$

were A, b and C are defined as in (1.22), (1.24) and (1.23) and satisfies the same hypothesis. In the affine control case (1.20), with a discrete observation,

the change of coordinates is the same as in the continuous case. In fact, (1.21) and (1.36) are exactly equivalent with $y_k = y(t_k)$.

The HG–EKF for continuous–discrete systems has the (not surprising) form:

Correction step

$$\begin{cases} z\left(t_k^+\right) = z(t_k) + G(k)\left(y_k - C(u)\, z(t_k)\right) \\ G(k) = P(t_k)\, C(u)' \left(C(u)\, P(t_k)\, C(u)' + \frac{1}{\Delta t} R\right)^{-1} \\ P\left(t_k^+\right) = \left(I - G(k)\, C(u)\right) P(t_k) \end{cases} \tag{1.37}$$

Prediction step

$$\begin{cases} \frac{dz}{dt} = A(u)x + b(x,u) \\ \frac{dP}{dt} = (A(u) + b^*(z,u))P + P(A(u) + b^*(z,u))' + Q_\theta \end{cases} \tag{1.38}$$

Then we have:

Theorem 6. *([14]) Under same assumptions as in continuous case and for Δt small enough, there is an interval $[\theta_0, \theta_1]$ such that for any $\theta \in [\theta_0, \theta_1]$, the continuous–discrete high–gain extended Kalman filter (1.37–1.38) is an exponential observer.*

Genericity and observability have also been studied for continuous–discrete systems. One can expect that same results hold when sampling time is small enough. Roughly speaking, it is more or less true. There exist continuous–discrete versions of theorems from Section 1.3.1 in the continuous–discrete case ([1, 2]).

1.3.6 A "weak" separation principle

In this section, we just want to give an important application concerning high-gain observers and particularly the high-gain extended Kalman filter.

Usually, observers are used in order to control nonlinear systems with a state-feedback control law. This control law $u(x)$ is calculated in order to achieve a good performance and, at least, to ensure the stability of the closed loop system. An observer is developed in order to estimate the state (which is not completely measured, in most applications) and the control law applied to the process is $u(z)$ (where z is the estimation of x given by the observer)[6].

Therefore, the closed loop system consist in a control law and an observer, and both are developed independently.

[6] If a filter has been developped, then one should apply the more accurate control law $u(t) = E\left[u(X(t)) \mid \mathcal{F}_t^Y\right]$ which is usually different from $u(z)$ where $z = E\left[X(t) \mid \mathcal{F}_t^Y\right]$.

In the linear-quadratic case, the "separation principle" stated that, if an optimal state-feedback control law is applied with an optimal observer, the result is optimal. It is a very strong "superposition" result which is false for nonlinear systems.

Nevertheless, we can expect to prove a weaker version of the linear separation principle.

Let us consider again our system 1.21. Let us suppose that there exist a positively invariant compact subset of $\mathbb{R}^n$ for any control law $u(t)$.

Theorem 7. *If $u(x)$ is a state feedback which make the system 1.21 globally asymptotically stable, then the system*

$$\begin{cases} \frac{dx}{dt} = A(u(z))x + b(x, u(z)) \\ \frac{dz}{d\tau} = A(u)z + b(z,u) - S(t)^{-1}C'R^{-1}(Cz - y(t)) \\ \frac{dS}{d\tau} = -(A(u) + b^*(z,u))'S - S(A(u) + b^*(z,u)) \\ \qquad +C'R^{-1}C - SQ_\theta S \end{cases}$$

is globally asymptotically stable for θ large enough.

Hence, this theorem show that the state-feedback control law can be replaced by an observer based control law and that the stability is preserved.

Remark 6. It has to be point out that this result is not true for the adaptive-gain extended Kalman filter (with these hypothesis) because it is necessary to have an exponentially converging observer with an arbitrary fast convergence.

1.4 Identifiability and identification

1.4.1 Definitions

The problem of identification is a generalization of the observation problem: very often, practical control systems depend on some functions, (with physical meaning), that are not well known, and that have to be determined on the basis of experiments. Systems under consideration have the following form

$$\begin{cases} \frac{dx}{dt} = f(x, u, \varphi(x,u)) \\ y = h(x, u, \varphi(x,u)) \end{cases} \tag{1.39}$$

If x denotes the state of the system, if $\varphi(x,u)$ is the unknown function, and $y(t)$ is the observed data, the identification problem is the problem of reconstructing the piece of the graph of $\varphi(.)$,visited during the experiment. That is, for an experiment of duration T, we want to determine the trajectories $(x(t), u(t), \varphi(x(t), u(t))$, for all $t \in [0,T]$, using only the observed data $\{y(t), t \in [0,T]\}$. We say that a system is identifiable if this is possible, whatever the experiment.

An identifier is a device performing this task. We will be interested with "on-line identifiers" only, *i.e.* identifiers that estimate the graph of φ simultaneously to the experiment.

The two problems, of observation and identification, are of course strongly connected for two reasons:

1. we do not suppose that $x(0)$ is known. Hence, the identification problem include an observation problem: we want to estimate both $x(t)$ and $\varphi(.)$. It is the main difference with the right-inversion problem, also known as the input identification problem.
2. identification requires an identifiability study, and this study is closely related to observability study. Moreover, our main tools to perform identification are based on (high–gain) observers.

Let us explain the second point, in the uncontrolled case. We consider smooth (C^ω or C^∞, depending on the context) systems of the form Σ

$$\Sigma \begin{cases} \frac{dx}{dt} = f(x, \varphi(x)) \\ y = h(x, \varphi(x)) \end{cases} \tag{1.40}$$

where the state $x = x(t)$ lies in a n–dimensional analytic manifold[7] X, $x(0) = x_0$, the observation y is $\mathbb{R}^p$–valued, and f, h are respectively a smooth (parametrized) vector field and a smooth function. The function φ is an unknown function of the state. Each trajectory is supposed to be defined on some interval $[0, T_{x_0,\varphi}[$.

- If the number of outputs is three ore more, then, identifiability is a generic property,
- If there is only one or two outputs, then, identifiability is a nongeneric property, so strong that it can be characterized by four very rigid normal forms.

Our goal is to estimate both state variable x and unknown function $\varphi : X \longrightarrow I$, I being a compact interval of $\mathbb{R}$ (the theory, developed in [10], clearly has extensions to higher dimension). More precisely, we want to reconstruct the piece of the graph of φ visited during experiment.

Let us recall some definitions and results from this last paper. For this introduction, we will only consider uncontrolled systems such as (1.40). Some results can be extended to controlled systems.

Let $\Omega = X \times L^\infty[I]$, where

$$L^\infty[I] = \{\hat{\varphi} : [0, T_{\hat{\varphi}}] \mapsto I,\ \hat{\varphi} \text{ measurable}\}$$

Then we can define the input/output mapping

[7] analytic manifold stands for analytic connected paracompact Hausdorf manifold

$$\begin{array}{rc} P_\Sigma : & \Omega \longrightarrow L^\infty\left[\mathbb{R}^{d_y}\right] \\ & (x_0, \hat{\varphi}(\cdot)) \longrightarrow y(\cdot) \end{array}$$

Definition 5. *Σ is said to be identifiable if P_Σ is injective.*

As for observability, we define an infinitesimal version of identifiability. Let us consider the first variation of the system (1.40) (where $\hat{\varphi}(t) = \varphi \circ x(t)$):

$$T\Sigma_{x_0,\hat{\varphi},\xi_0,\eta} \begin{cases} \frac{dx}{dt} = f(x, \hat{\varphi}) \\ \frac{d\xi}{dt} = T_x f(x, \hat{\varphi})\,\xi + d_\varphi f(x, \hat{\varphi})\,\eta \\ \hat{y} = d_x h(x, \hat{\varphi})\,\xi + d_\varphi h(x, \hat{\varphi})\,\eta \end{cases}$$

and the input/output mapping of $T\Sigma$

$$\begin{array}{rc} P_{T\Sigma,x_0,\hat{\varphi}} : T_{x_0}X \times L^\infty[\mathbb{R}] & \longrightarrow L^\infty\left[\mathbb{R}^{d_y}\right] \\ (\xi_0, \eta(\cdot)) & \longrightarrow \hat{y}(\cdot) \end{array}$$

Definition 6. *Σ is said to be infinitesimally identifiable if $P_{T\Sigma,,x_0,\hat{\varphi}}$ is injective for any $(x_0, \hat{\varphi}(\cdot)) \in \Omega$ i.e. $\ker(P_{T\Sigma,x_0,\hat{\varphi}}) = \{0\}$ for any $(x_0, \hat{\varphi}(\cdot))$.*

Both identifiability and infinitesimal identifiability mean injectivity of some mapping. Clearly injectivity depends on the domain. Therefore, it seems that these notions are not well defined. In fact these notions do not depend on the domain. Indeed, if an analytic system Σ is not (infinitesimally) identifiable because there exists a L^∞ function which make the system not (infinitesimally) identifiable, then there exist an analytic function which make the system not (infinitesimally) identifiable.

We consider again a system Σ of the form (1.40). In [10], we have shown that identifiability is a generic property if and only if the number of observation p is greater or equal to 3. On the contrary, if p is equal to 1 or 2, identifiability is a very restrictive hypothesis (infinite codimension) and we have completely classified infinitesimally identifiable systems by giving certain geometric properties that are equivalent to the normal forms presented in Theorems 8 and 9 [10] below.

These theorems are the basis of our identifier construction: since every identifiable systems may be put, up to a change of coordinates, in one of these canonical form of identifiability, then it is sufficient to develop an identifier for these forms (exactly as observers for observable systems).

Theorem 8. *($p = 1$) If Σ is uniformly infinitesimally identifiable, then, there is a subanalytic closed subset Z of X, of codimension 1 at least, such that for any $x_0 \in X \backslash Z$, there is a coordinate neighborhood $(x_1, \ldots, x_n, V_{x_0})$, $V_{x_0} \subset X \backslash Z$ in which Σ (restricted to V_{x_0}) can be written:*

$$\Sigma_1 \begin{cases} \dot{x}_1 &= x_2 \\ &\vdots \\ \dot{x}_{n-1} &= x_n \\ \dot{x}_n &= \psi(x,\varphi) \\ y &= x_1 \end{cases} \quad \text{and } \frac{\partial}{\partial \varphi}\psi(x,\varphi) \neq 0 \tag{1.41}$$

Theorem 9. *($p = 2$) If Σ is uniformly infinitesimally identifiable, then, there is an open-dense semi-analytic subset $\tilde{U}$ of $X \times I$, such that each point (x_0, φ_0) of $\tilde{U}$, has a neighborhood $V_{x_0} \times I_{\varphi_0}$, and coordinates x on V_{x_0} such that the system Σ restricted to $V_{x_0} \times I_{\varphi_0}$, denoted by $\Sigma_{|V_{x_0} \times I_{\varphi_0}}$, has one of the three following normal forms:*

- ***type 1 normal form:***

$$\Sigma_{2,1} \begin{cases} y_1 &= x_1 & y_2 &= x_2 \\ \dot{x}_1 &= x_3 & \dot{x}_2 &= x_4 \\ &\vdots & &\vdots \\ \dot{x}_{2k-3} &= x_{2k-1} & \dot{x}_{2k-2} &= x_{2k} \\ \dot{x}_{2k-1} &= f_{2k-1}(x_1, ..., x_{2k+1}) \\ \dot{x}_{2k} &= x_{2k+1} \\ &\vdots \\ \dot{x}_{n-1} &= x_n \\ \dot{x}_n &= f_n(x,\varphi) \end{cases} \tag{1.42}$$

  with $\frac{\partial f_n}{\partial \varphi} \neq 0$.
- ***type 2 normal form:***

$$\Sigma_{2,2} \begin{cases} y_1 &= x_1 & y_2 &= x_2 \\ \dot{x}_1 &= x_3 & \dot{x}_2 &= x_4 \\ &\vdots & &\vdots \\ \dot{x}_{2r-3} &= x_{2r-1} & \dot{x}_{2r-2} &= x_{2r} \\ \dot{x}_{2r-1} &= \psi(x,\varphi) & \dot{x}_{2r} &= F_{2r}(x_1, \ldots, \\ & & & \quad x_{2r+1}, \psi(x,\varphi)) \\ & & \dot{x}_{2r+1} &= F_{2r+1}(x_1, \ldots, \\ & & & \quad x_{2r+2}, \psi(x,\varphi)) \\ & & &\vdots \\ & & \dot{x}_{n-1} &= F_{n-1}(x, \psi(x,\varphi)) \\ & & \dot{x}_n &= F_n(x,\varphi) \end{cases} \tag{1.43}$$

  with $\frac{\partial \psi}{\partial \varphi} \neq 0, \frac{\partial F_{2r}}{\partial x_{2r+1}} \neq 0, ..., \frac{\partial F_{n-1}}{\partial x_n} \neq 0$
- ***type 3 normal form:***

$$\Sigma_{2,3}\begin{cases} y_1 = x_1 & y_2 = x_2 \\ \dot{x}_1 = x_3 & \dot{x}_2 = x_4 \\ \quad\vdots & \quad\vdots \\ \dot{x}_{n-3} = x_{n-1} & \dot{x}_{n-2} = x_n \\ \dot{x}_{n-1} = f_{n-1}(x,\varphi) & \dot{x}_n = f_n(x,\varphi) \end{cases} \tag{1.44}$$

with $\frac{\partial}{\partial\varphi}(f_{n-1}, f_n) \neq 0$.

Theorem 10. *($p \geq 3$) If Σ is an infinitesimally identifiable generic system, then there is a connected open dense subset Z of X such that for any $x_0 \in X\backslash Z$, there exist a smooth C^∞–function F and a $\left(\check{y}, \check{y}', \ldots, \check{y}^{(2n)}\right)$–dependant embedding $\Phi_{\check{y},\ldots,\check{y}^{(2n)}}(x)$ such that outside Z, trajectories of $\Sigma_{x_0,\varphi}$ are mapped via $\Phi_{\check{y},\ldots,\check{y}^{(2n)}}$ into trajectories of the following system*

$$\Sigma_{3+}\begin{cases} \frac{dz_1}{dt} = z_2 \\ \frac{dz_2}{dt} = z_3 \\ \quad\vdots \\ \frac{dz_{2n}}{dt} = z_{2n+1} \\ \frac{dz_{2n+1}}{dt} = F\left(z_1, \ldots, z_{2n+1}, \check{y}, \ldots, \check{y}^{(2n+1)}\right) \\ \bar{y} = z_1 \end{cases}$$

where z_i, $i = 1, \ldots, 2n+1$ has dimension $p-1$, and with

$$\begin{cases} x = \Phi^{-1}_{\check{y},\ldots,\check{y}^{(2n)}}(z) \\ \varphi = \Psi(x, \check{y}) \end{cases} \tag{1.45}$$

($\check{y}$ is a selected output).

1.4.2 Identifiers

As explained before, we have to build an identifier for each canonical form of identifiability. The basic idea is the same for all these forms, and leads to the use of the nonlinear observers developed previously: we assume, along the trajectories visited, a local model for φ. For instance, a simple local model is: $\varphi^{(k)} = 0$.

This does not mean, at the end, that we will identify φ as a polynomial in t: the question is not that this polynomial models the function φ globally as a function of t, but only locally, on reasonable time intervals (reasonable w.r.t. the performances of the observer that we will use).

This idea is just an extension of the classical way to identify constant or slowly varying parameters m. In this case, one use to add the parameter in the state variables and to add the equation $\frac{dm}{dt} = 0$. Therefore, the local model is a constant polynomial. In our case, such local model is too constrained (since φ is not supposed to vary slowly), so we add a polynomial local model.

Let us consider a system Σ in the identifiability normal form 1.41. Adding the local model for φ, we get the system:

$$\begin{aligned} y &= x_1, \qquad (1.46)\\ \dot{x}_1 &= x_2, ..., \dot{x}_{n-1} = x_n,\\ \dot{x}_n &= \Psi(x, \varphi_1), \dot{\varphi}_1 = \varphi_2, ..., \dot{\varphi}_{k-1} = \varphi_k, \dot{\varphi}_k = 0,\\ \frac{\partial \Psi}{\partial \varphi_1} &\neq 0 \quad \text{(never vanishes)}. \qquad (1.47) \end{aligned}$$

This is a system on $\mathbb{R}^{n+k}$, which is not controlled (however, for the considerations that follow, Ψ could depend on a control u), and this system is under the normal form (1.17, 1.18).

Therefore, we may apply high gain Luenberger observer, or we may apply the trick in Section 1.3.2. Then, for instance, the observer of Sections 1.3.2, 1.3.3 and 1.3.4 may be applied to this system. It will provide estimations of $x(t), \varphi(t)$, that is, just an estimation of the piece of the graph of φ visited during the experiment.

The cases of normal forms (1.42), (1.43), (1.44), corresponding to Type 1 to 3 systems can be treated in a similar way to the single-output case, with some more or less easy adaptations of the methods of the previous sections. This exercise is left to the reader.

An application of this technique in a difficult case (the local polynomial model does not apply) is presented in Section 1.6. Some important remarks and practical considerations are discussed in this section.

1.5 Series-connected DC motor

In this first application we present (in simulation) the design of the adaptive-gain extended Kalman filter (AG–EKF, see Section 1.3.4) for a single input single output (SISO) system, namely a series-connected DC motor.

Basically, an electric motor converts electrical energy into mechanical energy. In a DC motor, the stator (also denoted field) is composed of an electromagnet, or a permanent magnet, that immerses the rotor in a magnetic field. The rotor (also denoted armature) is made of an electromagnet that once supplied with current creates a second magnetic field. Movement is then caused by the attraction/repelling behavior of magnets. As far as the magnetic field created by the stator remain fixed the rotor windings are connected to a commutator. The direction of the current flowing through the armature coils is then switched during the rotation and the polarity of the armature magnetic field is reversed. Successive commutations then maintain the rotating motion of the machine. A DC motor whose field circuit and armature circuit are connected in series, and therefore fed by the same power supply, is referred to as a series-connected DC motor [34].

1.5.1 Mathematical model

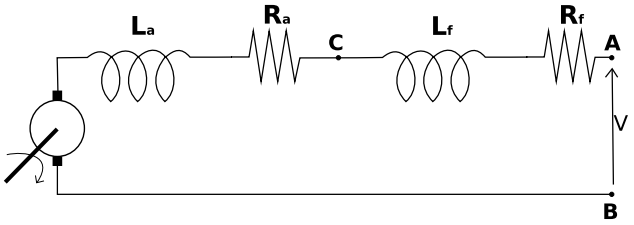


Fig. 1.1. Series-connected DC motor equivalent circuit representation

The model of the series-connected DC motor is obtained from the equivalent circuit representation shown in Figure 1.1. We denote I_f the current flowing through the field part of the circuit (between points A and C) and I_a the current through the armature circuit (between points C and B). When the shaft of the motor is turned by an external force, the motor acts as a generator and produces an electromotive force. In the case of the DC motor, this force will act against the current applied to the circuit and is then denoted *back or counter electromotive force* (BEMF or CEMF). The electrical balance leads to

$$L_f \dot{I}_f + R_f I_f = V_{AC}$$

for the field circuit, and to

$$L_a \dot{I}_a + R_a I_a = V_{CB} - E$$

where L_f and R_f are the inductance and the resistance of the field circuit, L_a and R_a are the inductance and the resistance of the armature circuit, and E denotes the Back EMF. Kirchoff's laws give us the relations

$$\begin{cases} I = I_a = I_f \\ V = V_{AC} + V_{CB} \end{cases}$$

which gives for the total electrical balance

$$L\dot{I} + RI = V - E$$

where $L = L_f + L_a$ and $R = R_f + R_a$. We now denote Φ the field flux, we have $\Phi = f(I_f) = f(I)$, and $E = K_m \Phi \omega_r$ where K_m is a constant and ω_r is the rotational speed of the shaft.

The second equation of the model is given by the mechanical balance of the shaft of the motor using the well known Newton's law. We consider that the only forces applied to the shaft are the electromechanical torque T_e, the viscous friction torque and the load torque T_a leading to

$$J\dot{\omega_r} = T_e - B\omega_r - T_a$$

where J denotes the rotor inertia, and B the viscous friction coefficient. The electromechanical torque is given by $T_e = K_e \Phi I$ with K_e denoting a constant parameter. We consider that the motor is operated **below saturation**: the field flux can be expressed by the linear expression $\Phi = L_{af} I$ where L_{af} denotes the mutual inductance between the field and the rotating armature coils. To conclude with the modeling of the DC motor we suppose the ideal hypothesis of 100% efficiency in the energy conversion expressed by $K = K_m = K_e$, and for notation simplicity we write L_{af} instead of KL_{af}. The voltage is the input of the system $u(t)$ and the current I is the measured output. We finally obtain the following SISO model for the series-connected DC motor

$$\begin{pmatrix} L\dot{I} \\ J\dot{\omega}_r \end{pmatrix} = \begin{pmatrix} u - RI - L_{af}\omega_r I \\ L_{af}I^2 - B\omega_r - T_a \end{pmatrix} \qquad (1.48)$$
$$y = I$$

This model will be used to simulate the DC motor by mean of a Matlab/Simulink S-function.

1.5.2 Observability canonical form

Before implementing the observer in order to reconstruct the state vector of this system we test (quite easily) its observability property. We use the *differentiation* approach that is we verify the differential observability (Definition 3) which implies observability.

- $I(t)$ is known with time, then $\dot{I} = (1/L)(u - R.I - L_{af}\omega_r I)$ is known and as far as $u(t)$, R, and L_{af} are known then ω_r can be computed
- now that $\omega_r(t)$ is known, $\dot{\omega_r} = (1/J)(L_{af}I^2 - B\omega_r - T_a)$ can be computed and because of the knowledge we have of $I(t)$, L_{af}, B,and J, T_a can be estimated

We deduce from this that a third variable may be added to the state vector in order to reconstruct both the state of the system and the load torque applied to the shaft of the motor. We assume that the load torque is constant over time. Sudden changes of the load torque will then be considered as unmodeled perturbations. The observer we use is the adaptive-gain Kalman filter as described in Section 1.3.4 because it has the classical EKF structure when no perturbations occur and the structure of a HG–EKF when the system faces a perturbation. Estimation of the load torque is made possible by the addition of the equation $\dot{T_a} = 0$ to (1.48) (see remarks in Section 1.4.2). We now need to find the coordinate transformation that puts this systems into the observability canonical form.

From the equation $y = I$, we choose $z_1 = I$ and then

$$\dot{z_1} = \frac{1}{L}(u(t) - RI - L_{af} I\omega_r)$$

which by setting $z_2 = I\omega_r$ becomes

$$\dot{z_1} = -\frac{L_{af}}{L} z_2 + \frac{1}{L}(u(t) - Rz_1) = \alpha_2(u) z_2 + b_1(z_1, u) \tag{1.49}$$

we now compute the time derivative of z_2

$$\dot{z_2} = \dot{I}\omega_r + I\dot{\omega_r} = -\frac{1}{J} T_a I - \frac{B}{J} I\omega_r + \frac{L_{af}}{J} I^3 - \frac{L_{af}}{L}\omega_r^2 I + \frac{u(t)}{L}\omega_r - \frac{R}{L}\omega_r I$$

when $I > 0$ and consequently $z_1 > 0$ which sounds as a reasonable assumption as far as I is the current of the circuit which is equal to zero only when there is no power supplied to the engine (and therefore nothing to observe), we set $\omega_r = \frac{z_2}{z_1}$, and by setting $z_3 = T_a I$ this equation becomes

$$\dot{z_2} = -\frac{1}{J} z_3 - \frac{B}{J} z_2 + \frac{L_{af}}{J} z_1^3 - \frac{L_{af}}{L}\frac{z_2^2}{z_1} + \frac{u(t)}{L}\frac{z_2}{z_1} - \frac{R}{L} z_2 = \alpha_3(u) z_3 + b_2(z_1, z_2, u) \tag{1.50}$$

and identical remark as above lead us to the expression $T_a = \frac{z_3}{z_1}$ and recalling that $\dot{T}_a = 0$ we obtain

$$\dot{z_3} = -\frac{L_{af}}{L}\frac{z_2 z_3}{z_1} + \frac{u(t)}{L}\frac{z_3}{z_1} - \frac{R}{L} z_3 = b_3(z_1, z_2, z_3, u) \tag{1.51}$$

Thus the application from $\mathbb{R}^{*+} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^{*+} \times \mathbb{R} \times \mathbb{R}$ defined by $(I, \omega_r, T_a) \to (I, I\omega_r, IT_a)$ with $(z_1, z_2, z_3) \to \left(z_1, \frac{z_2}{z_1}, \frac{z_3}{z_1}\right)$ as its inverse, is a change of coordinates that puts the system (1.48) into the observer canonical form defined by (1.49), (1.50) and (1.51). It is necessary to compute the coefficients of the matrix b^*.

1.5.3 Observer implementation

We now recall the equations of the AG–EKF

$$\begin{cases} \dfrac{dZ}{dt} = A(u)Z + b(Z, u) + PC' R_\theta^{-1}(CZ - y(t)) \\ \dfrac{dS}{dt} = P(A(u) + b^*(Z, u))' + (A(u) + b^*(Z, u)) \\ \qquad -PC' R_\theta^{-1} CP + Q_\theta \\ \dfrac{d\theta}{dt} = \lambda(1 - s(\mathcal{I})).(1 - \theta) + K.s(\mathcal{I}).(\theta_{max} - \theta) \end{cases} \tag{1.52}$$

where $R_\theta = \theta^{-1} R$ and $Q_\theta = \theta \Delta Q \Delta$ with $\Delta_\theta = diag\left(\theta, \theta^2, \ldots, \theta^n\right)$, $s(\mathcal{I}) = \left[1 + e^{-\beta(\mathcal{I} - m)}\right]^{-1}$ and

$$\mathcal{I} = \int_{t-T}^{t} \|y(s) - \bar{y}_{t-T}(s)\|^2 \, ds = \|y - \bar{y}_{t-T}\|^2_{L^2(t-T,t)} \tag{1.53}$$

In fact, these equations are a slight modification of (1.34): the function F has been modified in order to take into account noise effects, as we will explain below.

The simulation of the DC motor is straightforward, we then only comment the implementation of the observer. A Matlab/Simulink bloc diagram representing the DC machine and the observer is shown in Figure 1.2 (this figure is incomplete as far as one would surely want to plot errors between real and estimated states). As it may be seen from the simulink bloc diagram shown in Figure 1.3 the observer is decomposed into three parts: two level 1 S-functions and a transport delay bloc. As written on the diagram, the rightmost S-function is dedicated to the computation of the three main equations of the observer which are equations (1.52). This bloc has three type of inputs: the measured output of the observed system, the input delivered to the observed system and the innovation. The innovation is computed using a distinct S-function because unlike the main equations that may be processed continuously (or quasi-continuously), a **discrete** S-function is needed to compute the innovation. This choice was made because:

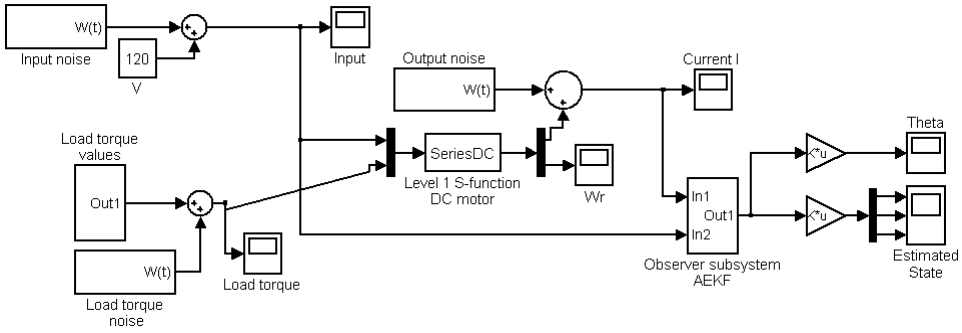


Fig. 1.2. Simulation and observation of the DC motor

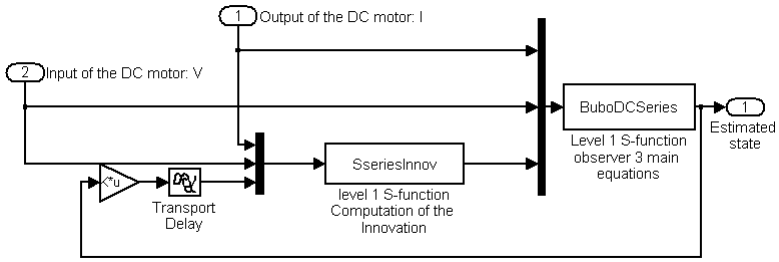


Fig. 1.3. Observer subsystem

- the computation of the integral is made by mean of a fixed step trapezoidal method
- we have to keep memory of the input and the output trajectories over a time interval $[0;T]$ where T is the delay of (1.53) which is easily done with a fixed step process.

The codes to implement those different functions may be downloaded from http://www.u-bourgogne.fr/monge/e.busvelle/springer/ or obtained from the authors if the link happens to be disabled.

1.5.4 Simulation parameters and observer tuning

The parameters used to simulate the DC engine, motivated by measures made on a real system, are $L = 1.22\,H$, $Res = 5.4183\,\Omega$, $L_{af} = 0.0683\,N.m.Wb^{-1}.A^{-1}$, $J = 0.0044\,kg.m^2$, and $B = 0.0026\,N.m.s^{-1}.rad^{-1}$.

We now need to set the observer parameters d, Dt, R, Q, θ_{max}, λ, K, β, and m. Before explaining how those parameters may be tuned, we want to stress that the last four ones do not need to be reseted for each new observer. Those parameters appear in the last equation in (1.52) and drive the evolution of the parameter θ. The values $\lambda = K = 500$, $\beta = 2000$, and $m = m_1 + m_2$ where $m_1 = 0.005$ (m_2 will be specific to each new process) may be kept each time a new observer is implemented. The procedure used to tune the parameters R, Q, θ_{max} is inspired by the one proposed in [9, part. 5.2.2].

1. As a first step, we determine the (symmetric positive definite) matrices R and Q by using an EKF. This observer can be obtained from the AG–EKF when the parameters of the adaptation function are set to 0 and $\theta(0) = 1$. Large perturbations are not considered and the observer is initialized to the proper (or previously estimated) values of the state vector.

2. As a second step, we set the R and Q matrices to the values previously found and use a HG-EKF in order to tune θ. As above the observer needed is obtained from the AG–EKF when the parameters of the adaptation function are set to 0. Then $\theta(0)$ is the value that is tuned. Here we will try to find a value for the high-gain parameter that allows fast and reasonable convergence (with respect to noise amplification) when large unmodeled perturbations are applied to the system. θ_{max} is then taken equal to the value estimated at this step.

3. As a last step we now set the parameters of the adaptation function. We remark that when $m = 0$ then $s(0) = 0.5$. Thus we need to shift the sigmoid function to the right if we want $s(0)$ to be close to zero. Choosing y_1 as small as we want and solving the equation $s(0) = y_1$ allows to obtain the parameter m. This solution is easily computed provided that the parameter β is known. As the sigmoid function is centered on $(0, 0.5)$ when $m = 0$, the computation of β is made by setting a length l for the transition part and solving the nonlinear equation (with $m = 0$): $s(l/2) - s(-l/2) = (1 - y_1) - y_1$. Of course, another approach is to graphically define β and m from trial and error. Figure 1.4

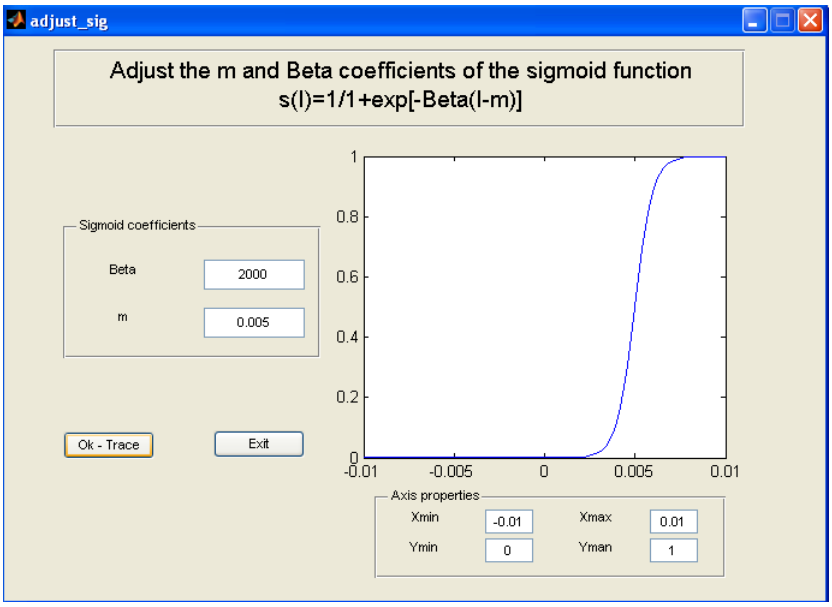


Fig. 1.4. estimation of β and m_1 by trial and error

shows a simple Matlab GUI implemented to ease this latter method (the result displayed is for the values of β and m_1 given above). The code of this GUI is also available from *http://www.u-bourgogne.fr/monge/e.busvelle/springer/*.

Now that the transition part is small, we want the gain to increase and decrease quickly. If we suppose that $\theta(t) = 1$ and that we want it to reach θ_{max} within a time τ then the equation $\dot{\theta} = \frac{\theta_{max}-1}{\tau} = K.(\theta_{max} - 1)$ allows the computation of K. As far as the equation used to compute K is only an approximation, a bigger value (e.g. twice the computed value) may be used. Finally, a reasonable choice for the last parameter remaining is $\lambda = K$.

The parameter T, the length of the window on which innovation is computed, is related to the rise time of the system when it is facing perturbations: it has to be sufficiently big so as to give an account of perturbations that occur on the system. The sample time Dt of the discrete S-function should ideally be chosen as small as possible, leading to a significant increase of the amount of time and of the memory needed to compute the innovation (we need to keep track of $\frac{T}{Dt}+1$ system outputs and $\frac{T}{Dt}$ system inputs). $Dt = T/3$ or $Dt = T/4$ seems to be reasonable, fewer values will of course give more flexibility to the system.

Because of measurement noise the innovation will never be equal to zero and therefore the observer will stay in a high-gain mode. To avoid this problem, the parameter m is rewritten $m = m_1 + m_2$ where m_1 is the previously computed quantity and m_2 will represent the influence of the noise on the

system. As a result, when $\mathcal{I} \leq m_2$ we will have $s(\mathcal{I}) \leq y_1$ and θ won't increase. We denote by σ the standard deviation of the output noise, which can be estimated from output measurements, and then $m_2 = T.\sigma^2$ where T is the delay used in the definition of the innovation. Figure 1.5 shows the output of the simulated DC motor (with addition of noise) and that $\sigma = 0.7$ is a reasonable value for the standard deviation.

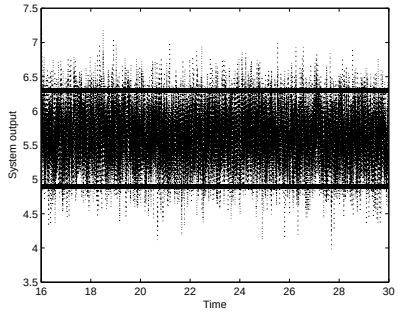


Fig. 1.5. Estimation of the standard deviation

Finally all those steps allow us to set the parameters to $R = 1$, $Q = [1, 0, 0; 0, 5, 0; 0, 0, 5]$, $\theta_{max} = 3$, $\lambda = K = 500$, $\beta = 2000$, $T = 0.1$, $Dt = 0.01$, and $m = 0.005 + 0.049$.

1.5.5 Simulation results

Figures 1.6 and 1.7 shows the performance of the designed observer, all the observers identify the values taken by the load torque but with different behaviors. The EKF rejects noise but converges slowly when the system faces unmodeled perturbations. We may add that in order to speed up a little bit the EKF the Q matrix was set to $[25, 0, 0; 0, 25, 0; 0, 0, 50]$ in this special case, it was kept to the value given in the previous chapter for all the other simulations.

The HG–EKF is on the contrary very sensitive to measurement noise but is very fast regarding convergence when a perturbation arises.

The AG–EKF presents both the advantages of the two previous filters, namely noise rejection and speed of convergence under perturbations. We observe that the adaptive-gain observer is a little bit slower than the fixed high-gain one. This is due to the delay induced by the computation of innovation, in fact the value chosen for Dt will have an impact on this delay as far as the behavior of θ (increasing toward θ_{max} or decreasing toward 1) will only change with the innovation. In all the parameters tuned for this last observer one will have a major impact, this is m_2. Indeed if it is set to a too big value

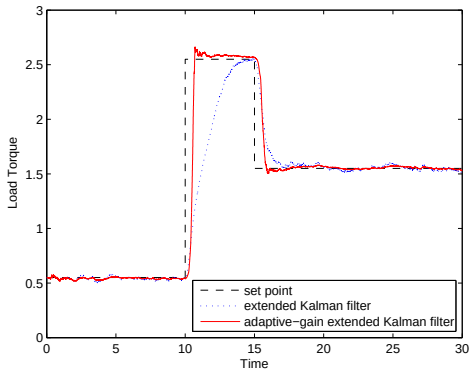


Fig. 1.6. EKF VS AEKF

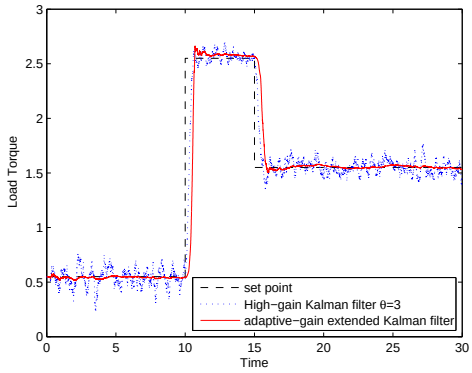


Fig. 1.7. HGEKF VS AEKF

then θ won't increase every time it is needed what does not constitutes a major drawback because the EKF rejects noise (this is true provided that m_2 is not such as big that it totally prevent θ from increasing). On the contrary, if m_2 is too small then θ will increase when it is not needed (only because of the noise) having the only effect to amplify noise. However as it can be seen from Figure 1.5, σ and therefore m_2 is not difficult to estimate from output measurements. To illustrate this comment Figure 1.8 shows the evolution of θ for two different values of m_2 (the value 0.049 corresponds to the simulations which results are shown above).

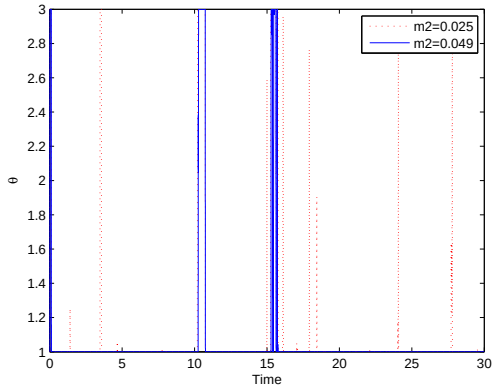


Fig. 1.8. Different values of m_2

1.6 Electronical neuron circuit

With this second application we illustrate how observers can assist system modeling and, in the case considered here, prototype assessment (as in Section 1.4). Identifiability study of this model has been presented in [5].

The modelization of neurons is extensively studied in neuroscience research. A large quantity of models of isolated neuron cells or of neuron cells networks are available in the literature each one of them presenting variable degrees in their accuracy. The model we use here, a modification of the model proposed by Fitzhugh, Nagumo & al. in the early 1960's, is a simplification of the one of a single isolated biological neuron proposed by Hodgkin and Huxley [24]. Historical informations on the development of this model can be found in [28].

1.6.1 The modified Fitzhugh-Nagumo model (MFHN)

From the biological point of view this model is composed of two variables, V representing the membrane voltage and W that represents the recovery variable

$$\begin{cases} \dot{V} = V - \frac{V^{\cdot 3}}{3} - W \\ \dot{W} = \epsilon \left(g(V) - W - \eta \right) \end{cases} \tag{1.54}$$

where ϵ and η are constant parameters and g is the piecewise linear function

$$g(V) = \begin{cases} \beta V \; if \; V > 0 \\ \alpha V \; if \; V \leq 0 \end{cases}$$

where α and β are constant parameters.

This model was implemented as an analogue circuit at LE2I laboratory (university of Burgundy), the exact description of this circuit is given in [6]. The analyze of this physical system is made by mean of an observer based approach, real data being available.

1.6.2 Identifiability and observability

From the analogue circuit point of view, V corresponds to a voltage and W to a current therefore both of them can be measured. Although in the case of a real biological system it will only be possible to measure V, the membrane voltage. Thus we will consider that only V is actually measured. The objective of this study is the identification of the function g (i.e. the part of its graph visited during the experiment) and the study of the identifiability property of the system constitutes a first step. In Section 1.4, we described an identifiability normal form for single output uncontrolled systems (normal forms for systems with more than one output are also given)

$$\begin{cases} \dot{x_1} &= x_2 \\ & \vdots \\ \dot{x_{n-1}} &= x_n \\ \dot{x_n} &= \psi(x,g) \\ y &= x_1 \end{cases} \tag{1.55}$$

We now want to find a change of coordinates that allow the MFHN equations to match this normal form. This coordinate transformation is easily found: set $x_1 = V$ and $x_2 = \dot{V}$.

$$\begin{cases} \dot{x_1} = \dot{V} \\ \quad = x_2 \\ \dot{x_2} = \ddot{V} - \dot{V}V^2 - \dot{W} \\ \quad = (1-x_1^2)x_2 - \epsilon\left(g(x_1) - x_1 + \frac{x_1^3}{3} + x_2 - \eta\right) \\ \quad = \psi(x,g) \end{cases} \tag{1.56}$$

Since $\epsilon \neq 0$, the system is clearly identifiable. We see that if the parameter η is unknown we have the possibility to redefine the unknown function g as $g(x_1) = g(x_1) - \eta$ with no change in the normal form.

In order to identify the function g, we extend the state vector by making g a state variable. As it is clear that g is not constant over time we model it as a local polynomial of time

$$g(V(t)) = g(t) = a_0 + a_1 t + ... + a_n t^n$$

which imply that $\frac{d^{n+1}g(t)}{dt^{n+1}} = 0$. The model is completed by the addition of n new state variables corresponding to the n first derivatives of g with respect to time (for a total of $n+1$ new variables). It appears that when the system

defined by (1.56) is extended in that manner it is in the observability canonical form. However there exist a much more simpler way to obtain the canonical form that is not to do any change of variables. This latter form is the one we will consider so as to avoid change of variables while implementing the observer

$$\frac{d}{dt}\begin{pmatrix} V \\ W \\ h_0 \\ h_1 \\ ... \\ h_{n-1} \\ h_n \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 & 0 & ... & 0 & 0 \\ 0 & 0 & \epsilon & 0 & ... & 0 & 0 \\ 0 & 0 & 0 & 1 & ... & 0 & 0 \\ 0 & 0 & 0 & 0 & ... & 0 & 0 \\ ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & 0 & ... & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} V \\ W \\ h_0 \\ h_1 \\ ... \\ h_{n-1} \\ h_n \end{pmatrix} + \begin{pmatrix} V - \frac{V^3}{3} \\ -\epsilon(W - \eta) \\ 0 \\ 0 \\ ... \\ 0 \\ 0 \end{pmatrix}$$

where $h_i = \frac{d^i g(t)}{dt^i}$ for $i = 0, ..., n$ and with $\frac{d^0 g}{dt} = g$.

One could think that the choice of a local representation for the function g (here a polynomial of time) and the transformation of the model into the canonical observability form suffices to prove identifiability. It is in fact not the case. This subtle difference has been well illustrated in [9, part 6] where the authors exhibit the example

$$\begin{cases} \dot{x} = \quad \varphi(x) \\ y = x + \varphi(x) \end{cases} \quad x \in \mathbb{R}$$

indeed, keeping the notations used above for the function g and setting $n = 1$, then the change of coordinates $(x, h_0, h_1) \to (z_1, z_2, z_3) = (x + h_0, h_0 + h_1, h_1)$ leads to an observability canonical form. However the authors showed that this system is not identifiable !

1.6.3 Implementation

The high-gain extended Kalman filter is adapted to the problem of identification of the unknown function g. The implementation of this observer is much more easy to carry on than the previous one: only one S-function is needed. Even if our objective is to use real data get from the analogue circuit mentioned above we use a continuous S-function. This is motivated by the fact that our data's sample time is smaller than the average time step used by the software to compute the continuous solutions (but a continuous–discrete observer (1.37)–(1.38) will be another possible choice). The corresponding Matlab/Simulink diagram is shown Figure 1.9.

Codes may be downloaded (together with a set of data) from http://www.u-bourgogne.fr/monge/e.busvelle/springer/.

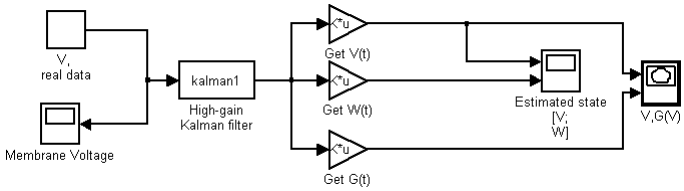


Fig. 1.9. Identification from real data

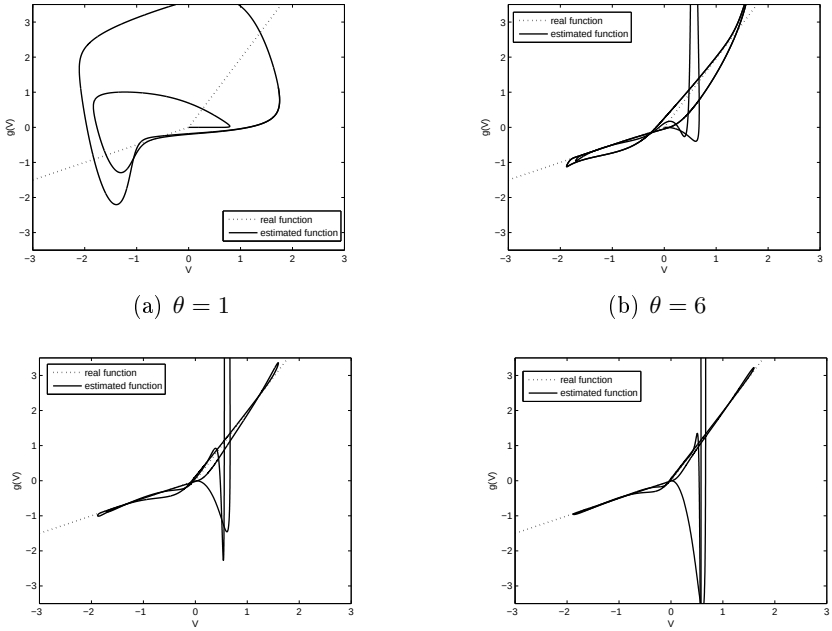
1.6.4 Results

A first series of simulations of the MFHN model are done in order to tune the three parameters n, Q, and θ. The parameters for the MFHN model are set to $\alpha = 0.5$, $\beta = 1.96$, $\epsilon = 0.2966$, $\eta = 0.20531$ $V(0) = 1.0656$, and $W(0) = 2.6903$. Since we are using an observer that only has a high-gain behavior, Q is set to the identity matrix $Id_{(3+n)\times(3+n)}$. The high-gain parameter θ is then chosen to ensure an accurate identification of the function. Several simulations shows that $\theta = 1$ (corresponding to an extended Kalman filter) does not lead to the identification of the function. The identification is made possible when $\theta \in [5; 10]$, and is very accurate when $\theta > 10$. Figure 1.10 shows identification results for four different values of the high-gain parameter when the data fed to the observer are simulated. No noise have been added during those simulations and then even if $\theta = 15$ gives the best result, the trade-off between speed of convergence and sensibility to noise lead us to choose a smaller value.

The values for V got from the analogue circuit are shown Figure 1.11 and the result of the identification (with $\theta = 10$ and $n = 1$) is shown Figure 1.12(a). We see that the unknown function is identified as a loop and from the shape of the data used, we expect four of them.

We isolated the first values given by the observer in order to obtain the clearer graphic Figure 1.12(b) in which we highlighted the overshoot due to the inaccurate initialization of the observer. After this overshoot the observer converges to the values taken by the unknown function and while $V < 0$ the estimation is quite good. When V becomes positive the estimation is not that accurate anymore. Two reasons can be pointed out to explain this phenomenon: the real data do not correspond exactly to the output the theoretic model would give for the same set of parameters (which is analogous to modeling errors) and the fact that the function we want to identify is not differentiable in 0, a very specific property that is not reflected by our polynomial approximation.

We rewrite the model used to perform the identification so as to take this into consideration



(a) $\theta = 1$ (b) $\theta = 6$

(c) $\theta = 10$ (d) $\theta = 15$

Fig. 1.10. Identification of g from simulations (without noise addition)

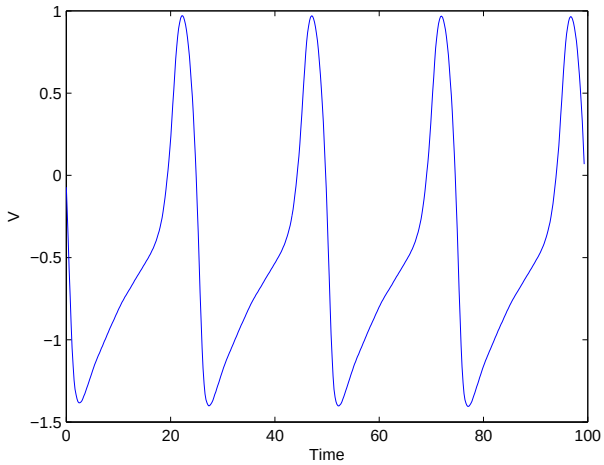


Fig. 1.11. Circuit voltage V

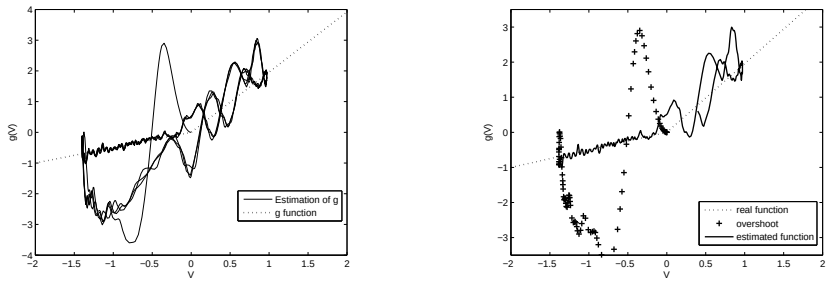
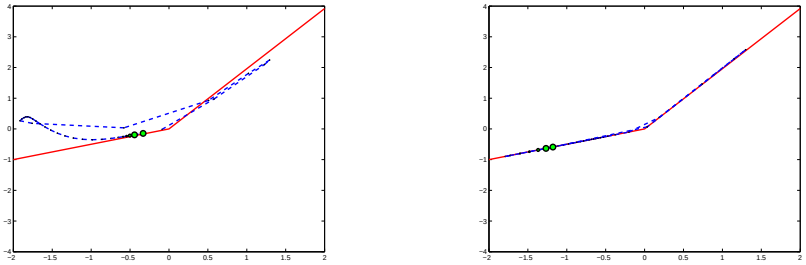


Fig. 1.12. Decomposed identification of the function g



(a) after 4 cycles (b) after 15 cycles

Fig. 1.13. Estimation of g

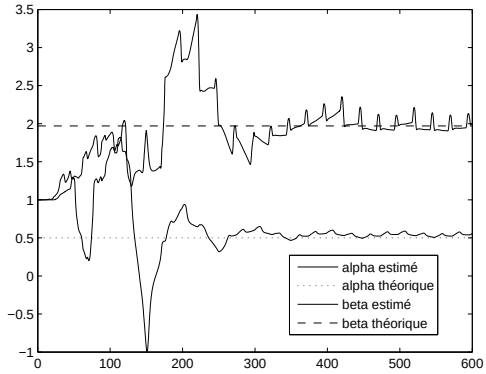


Fig. 1.14. $\hat{\alpha}$ and $\hat{\beta}$ against time

$$\begin{cases} \dot{x_1} = x_2 & \dot{x_2} = \bar{\psi}(x, \hat{\alpha}, \hat{\beta}) \\ \dot{\hat{\alpha}} = \alpha_1 & \dot{\hat{\beta}} = \beta_1 \\ \dot{\alpha_1} = \alpha_2 & \dot{\beta_1} = \beta_2 \\ \dot{\alpha_2} = \alpha_3 & \dot{\beta_2} = \beta_3 \\ \dot{\alpha_3} = 0 & \dot{\beta_3} = 0 \end{cases} \tag{1.57}$$

The results of this new identification are shown Figure 1.13(a-b). This new estimation is very accurate after a few cycles. Small errors both for the positive and negative values of V are still visible, they can also be spotted when we trace the values taken by $\hat{\alpha}$ and $\hat{\beta}$ against time as in Figure 1.14. Those errors are due to the fact that real data differ from the ideal mathematical model.

References

1. Ammar S., Vivalda J.-C. (2005) On the genericity of the observability of controlled discrete-time systems, in ESAIM: Control, Optimisation and Calculus of Variations, vol. 11, 161–179
2. Ammar S., Vivalda J.-C. (2004) On the preservation of observability under sampling, Systems & Control Letters, vol. 52, no 1, 7–15
3. Balde M., Jouan P. (1988) Observability of control affine systems, ESAIM/COCV, Vol. 3, pp. 345-359, 1998.
4. Baras J.S. , Bensoussan A. , James M.R. , (1988) Dynamic observers as asymptotic limits of recursive filters: special cases", SIAM J. Appl. Math., 48, 1147–1158
5. Binczak S, Busvelle E, Gauthier JP, Jacquir S (2006) Identification of unknown functions in dynamic systems. 17th International Symposium on Mathematical Theory of Networks and Systems (MTNS 2006), Kyoto:2157–2162.
6. Binczak S, Kazantsev V, Nekorkin V I, Bilbault JM (2003) Experimental study of bifurcations in modified Fitzhugh-Nagumo cell. Electronic letters 39,13:961–962.
7. Boizot N, Busvelle E, Gauthier JP, Sachau J (2007) Adaptive-Gain Extended Kalman Filter: Application to a Series-Connected DC Motor. In Conference on Systems and Control (CSC'2007), Marrakech.
8. Busvelle E., Gauthier J.-P. (2002) High-Gain and Non-High-Gain Observers for Nonlinear Systems, Contemporary Trends in Nonlinear Geometric Control Theory and its Applications, World Scientific, 233–256
9. Busvelle E., Gauthier J.-P. (2005) Observation and Identification Tools for Nonlinear systems. Application to a Fluid Catalytic Cracker. International Journal of Control, Vol. 78, 3, 2005
10. Busvelle E., Gauthier J.-P. (2003) On determining unknown functions in differential systems, with an application to biological reactors, ESAIM: COCV 9, 509–552
11. Bucy R. , Joseph P. (1987) Filtering for stochastic processes with applications to guidance, Chelsea publishing company, second edition
12. Busse F.D. , How J.P., Simpson J. (2003) Demonstration of Adaptive Extended Kalman Filter for Low-earth-orbit Formation Estimation using CDGPS Navigation. Journal of the Institute of Navigation, v 50, n 2, 79–93

13. Deza F. (1991) Contribution to the synthesis of exponential observers, Phd thesis, INSA de Rouen, France, June 1991.
14. Deza F., Busvelle E., Gauthier J.-P. (1992) High-gain estimation for nonlinear systems, Systems and Control Letters 18, 295–299
15. Deza F., Busvelle E., Gauthier J.-P., Rakotopara D., (1992)A stability result on the continuous-continuous and continuous-discret extended Kalman filters, Compte-Rendus de l'Académie des Sciences de Paris, 314, Février 1992
16. Gauthier J.-P., Hammouri H., Kupka I., (1991) Observers for nonlinear systems; IEEE CDC Conference, december 1991, Brighton, England, 1483–1489
17. Gauthier J.-P., Hammouri H., Othman S. (1992) A simple observer for nonlinear systems. IEEE Trans. Aut. Control, 37, 875–880
18. Gauthier J.P., Kupka I., (1994) Observability and observers for nonlinear systems. SIAM Journal on Control, vol. 32, N° 4, 975–994
19. Gauthier J.P., Kupka I., (1996) Observability for systems with more outputs than inputs. Mathematische Zeitschrift, 223, pp. 47–78
20. Gauthier, J.-P., and Kupka, I., (2001), Deterministic Observation Theory and Applications, Cambridge University Press
21. Gikhman, I. I., Skorokhod, A.V. (1969), Introduction to the Theory of Random Processes, W.B. Saunders, Philadelphia
22. Gevers M., Bastin G. (1986) A Stable Adaptive Observer for a class of Nonlinear Second-order Systems, Analysis and optimization of systems, ed. by Bensoussan, A., and Lions, J. L. Springer-Verlag, 143–155
23. Hammouri, H., Farza, M., (2003), Nonlinear observers for locally uniformly observable systems, COCV, 9, 343-352.
24. Hodgkin L A, Huxley A F (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. Journal of phisiology 117:500–544.
25. Holland C.D. (1963) Multicomponent Distillation, Englewood Cliffs, New-Jersey, USA: Prentice Hall, 1963.
26. Hu C., Chen W., Chen Y., Liu D. (2003) Adaptive Kalman filtering for vehicle navigation, Journal of Global Positioning Systems, Vol. 2, No 1, 42–47
27. Huillet T., Salut G. (1989) Interprétation des équations du filtrage non linéaire, GdR Automatique du CNRS, Paris, 8 Novembre 1989
28. Jacquir S (2006) Systèmes dynamiques non-linéaires, de la biologie à l'électronique. Thèse de doctorat, Université de Bougogne, France.
29. Jaswinski A., (1970) Stochastic processes and filtering theory, Academic Press, New York
30. Johansson A., Medvedev A. (2003), An observer for systems with nonlinear output map, Automatica 39, pp 909–918
31. Jouan P., (1995) Singularités des systèmes non linéaires, observabilité et observateurs, PHD thesis, Université de Rouen,
32. Jouan P., J.P. Gauthier J.P., (1996) Finite singularities of nonlinear systems. Output stabilization, observability and observers. Journal of Dynamical and Control Systems, vol. 2, N° 2, 255–288.
33. Kent K. C. Yu, N. R. Watson, J. Arrillaga, (2005) An adaptive Kalman filter for dynamic harmonic state estimation and harmonic injection tracking, IEEE Trans. on power delivery, Vol. 20, No 2
34. Krause P C, Wasunczuk O, Sudhoff S D (2002) Analysis of Electric Machinery and Drive Systems, 2nd edition. Wiley-interscience, IEEE series on power engineering.

35. La Moyne L., Porter L. L., Passino K. M. , (1995) Genetic adaptive observers, Engng Applic. Artif. Intell. Vol. 8, No 3, 261–269
36. Mehta S,Chiasson J (1998) Nonlinear Control of a Series DC Motor: Theory and Experiment. IEEE transactions on industrial electronics 45, 1.
37. E. Pardoux (1991) Filtrage non linéaire et équations aux dérivées partielles stochastiques associées, Ecole d'Eté de Probabilités de Saint-Flour XIX-1989, Lecture Notes in Math., 1464, Springer, 67–163
38. Picard J. (1991), "Efficiency of the extended Kalman filter for nonlinear systems with small noise", SIAM J. Appl. Math., 51, No3, 843–885.
39. Rigal G. (1993) Filtrage non-linéaire, résolution particulaire et applications au traitement du signal, Thèse de l'Université de Toulouse, LAAS.
40. Rosenbrock H. (1962), A Lyapunov function with applications to some nonlinear physical systems, Automatica, 1, 31–53
41. Ropuchon P. (1990), Simulation dynamique et commande non linéaire des colonnes à distiller, Thèse de l'école des mines de Paris,
42. Shen L., Wang H. (2003), Adaptive observer design for general nonlinear systems with linear output structure, 4th International Conference on Control and Automation, June 2003. ICCA '03. Proceedings, 48–52
43. Shim, H., Son, Y. I., and Seo, J. H., (2001), Semi-global observer for multi-output nonlinear systems, Systems Control Lett. 42 233–244.
44. Stubberud S., Lobbia R., Owen M., (1998) An adaptive extended Kalman filter using artificial neural networks, The international journal on smart system design, Vol. 1, pp 207-221
45. Viel, F (1994) Stabilité des systèmes non linéaires controlés par retour d'état estimé. Application aux réacteurs de polymérisation et aux colonnes à distiller, Thèse de l'université de Rouen
46. Viel F., Busvelle E., Gauthier J.P. (1997) A stable control structure for binary distillation columns, International Journal on Control, Vol 67, N°4, 475–505