Asymptotic Behaviour of Reduced-Order Filters

A. Le Breton, M. C. Roubaud

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Asymptotic Behaviour of Reduced–Order Filters

A. Le Breton and M. C. Roubaud

N° 3311

Décembre 1997

THÈME 4

Rapport de recherche

ISSN 0249-6399

Asymptotic Behaviour of Reduced–Order Filters

A. Le Breton* and M. C. Roubaud†

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet SYSDYS

Rapport de recherche n° 3311 — Décembre 1997 — 40 pages

Abstract: Reduced–order filters are proposed for linear and nonlinear systems and their long time behaviour is studied. Using the results of Ocone and Pardoux [10] on the asymptotic stability of the optimal filter with respect to its initial condition, the asymptotic efficiency of these filters is established in various cases.

Key-words: Nonlinear filtering, Kalman filter, reduced-order filters, asymptotic efficiency

M.C. Roubaud (Université Joseph Fourier) worked in the project IDOPT (CNRS, UJF, INPG, INRIA) until September 1997 and now she is associated with the project SYSDYS (INRIA–LATP–UNIVERSITÉ DE PROVENCE).

* Idopt (CNRS, UJF, INPG, INRIA) LMC–IMAG, BP 53, 38041 Grenoble Cedex 9 (France), `Alain.Le-Breton@imag.fr`

† Sysdys–LATP, 39 rue Joliot–Curie, 13453 Marseille Cedex 13 (France), `mcroubau@sophia.inria.fr`

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, B.P. 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : 04 93 65 77 77 - International : +33 4 93 65 77 77 — Fax : 04 93 65 77 65 - International : +33 4 93 65 77 65
à partir du 01/01/1998
Téléphone : 04 92 38 77 77 - International : +33 4 92 38 77 77 — Fax : 04 92 38 77 65 - International : +33 4 92 38 77 65

Comportement asymptotique de filtres d'ordre réduit

Résumé : Nous proposons des filtres d'ordre réduit pour des systèmes linéaires et non linéaires et nous étudions leur comportement en temps long. En utilisant les résultats d'Ocone et Pardoux [10] sur la stabilité asymptotique du filtre optimal par rapport à la condition initiale, nous montrons l'efficacité asymptotique de ces filtres dans divers cas.

Mots-clés : Filtrage non linéaire, filtre de Kalman, filtres d'ordre réduit, efficacité asymptotique

Contents

1 Introduction

Consider the filtering problem of estimating at time t an unobserved state process $Z = \{Z_s\,,\ s \geq 0\}$ based on the observation of Y up to time t. The observation process $Y = \{Y_s\,,\ s \geq 0\}$ is given by

$$Y_s = \int_0^s h(Z_u)\,du + W_s\,,$$

where h is a continuous function and $W = \{W_s\,,\ s \geq 0\}$ is a Brownian motion which is independent of the state process Z.

Let $\mathcal{Y}_t = \sigma(Y_s\,,\ 0 \leq s \leq t)$. The conditional distribution Π_t of Z_t given $\mathcal{Y}_t$ is the optimal solution of the filtering problem. It will be referred to as the optimal filter.

In the case where Z is a Gaussian diffusion and h is a linear function, it is well known that the conditional law is Gaussian and moreover, its mean and covariance matrix are respectively solutions to a stochastic differential equation and a Riccati differential equation (the Kalman filter equations). However, for large scale models, for instance arising in data assimilation problems in the geosiences (e.g. [12]), the high computational demand and the required storage of the Kalman filter become prohibitive.

On the other hand, in the nonlinear case the filtering problem generally is infinite dimensional. Indeed, the conditional probability density (when it exists) is in most cases given only by the solution of a parabolic partial differential equation called Zakai's equation (see [13] or e.g. [11]). The difficulties encountered in numerical computations of this equation grow with the dimension of the space variable. So, both in the linear case and in the nonlinear case, it is interesting to propose reduced–order filters which are more easily computable than the optimal one and to study their efficiency when the time goes to infinity. This is the aim of the present work.

In this report, in various cases we propose a reduced–order filter which corresponds to the conditional law with erroneous initial condition. Therefore to study the asymptotic optimality of this filter, we shall use the results obtained by Ocone and Pardoux in [10] on the asymptotic stability of the optimal filter with respect to its initial condition and we shall extend their results to a linear case with coloured noises.

The report is organized as follows. At first we consider the linear case in Section 2. We study the asymptotic efficiency of one reduced–order Kalman filter as the time tends to infinity. The idea is to apply correction in certain directions only. These directions depend on the system dynamics and they are those for which the system does not reduce the prediction error. Our first step is to change state coordinates so that these directions appear clearly. Relative to these coordinates, the form of the state equation will be referred to as the *stabilizability canonical form*. In the case where the stable component is noise–free, the results of [10] imply the convergence of the approximate filter towards the optimal one as time tends to infinity. If the stable component is noisy, under some assumptions a upper bound is given for the approximation error.

In Section 3, we consider a state process with two components $Z = (M, X)$. The system is linear in X and nonlinear in M where M is a Markov process which tends to zero as time tends to infinity in some sense. A finite dimensional approximate filter which is asymptotically optimal is proposed. In fact this case, called the *semi-linear case*, is interpreted as a linear filtering problem with coloured noises, X_t being the state to be estimated by means of the observation until time t.

In Section 4, a *nonlinear case* when the state process is ergodic is studied. An approximate filter which satisfies a Zakai equation whose dimension in the spatial variable is smaller than the exact equation is proposed. Using the results of [10], the asymptotic efficiency of this filter is shown.

Notations : Throughout, we shall let $\hat{Z}_t$ denote the conditional expectation $\mathbf{E}[Z_t|\mathcal{Y}_t]$, where $\mathcal{Y}_t = \sigma\{Y_s\,,\, 0 \le s \le t\}$. If Π_t denotes the conditional distribution of Z_t given $\mathcal{Y}_t$, $\Pi_t(\varphi)$ denotes the conditional expectation $\mathbf{E}[\varphi(Z_t)\,|\,\mathcal{Y}_t]$ for every $\varphi \in \mathcal{C}_b(\mathbb{R}^n)$ i.e. for every bounded continuous function $\varphi : \mathbb{R}^n \to \mathbb{R}$. The notation $\mathcal{N}(z, Q)$ will be used for the Gaussian law on $\mathbb{R}^n$ with mean z and covariance matrix Q and for every $\varphi \in \mathcal{C}_b(\mathbb{R}^n)$:

$$\mathcal{N}(z,Q)(\varphi) = \int_{\mathbb{R}^n} \varphi(y)\, d\mathcal{N}(z,Q)(y)\,.$$

The Euclidean norm of vector x is denoted by $|x|$. The $n \times n$ identity matrix is written I_n. For a matrix M, M' and $\mathrm{Tr}(M)$ stand respectively for the transpose and the trace and the norm is defined as $||M|| = \sqrt{\text{Tr } (MM')}$. If M is symmetric positive (nonnegative) definite, we write $M > 0$ ($M \geq 0$); $M > N$ means $M - N > 0$ and similarly with $\geq$ in place of $>$. Let $\lambda_1(M), \ldots, \lambda_n(M)$ denote the eigenvalues of a $n \times n$–matrix M and $\text{Re}\lambda_i(M)$ denote the real part of $\lambda_i(M)$, $i = 1 \ldots n$. All the eigenvalues are ordered such that their real parts are nonincreasing, i.e.,

$$\text{Re}\lambda_1(M) \geq \text{Re}\lambda_2(M) \geq \cdots \geq \text{Re}\lambda_n(M)\,.$$

Finally, let c and C denote any constant. These constants may then stand for different real values, from line to line.

2 The linear case

2.1 Some results on Riccati equations and linear Kalman filtering

Let us first present some useful and classical results of the stability theory and on the solutions of Riccati differential equations.

Definition 2.1 *The $n \times n$–matrix M is said to be asymptotically stable if and only if all its eigenvalues have a strictly negative real parts, i.e. $Re\lambda_1(M) < 0$.*

Lemma 2.2 *For any M, $N \in \mathbb{R}^{n \times n}$, if $N = N' > 0$ then*

$$\frac{1}{2}\lambda_n(M + M')tr\ (N) \leq tr\ (MN) \leq \frac{1}{2}\lambda_1(M + M')tr\ (N)\,.$$

Proof See [9].

Lemma 2.3 *Let M be any real $n \times n$–matrix. For all $t > 0$, We have*

$$\lambda_1(e^{Mt}e^{M't}) \leq \exp \lambda_1(M + M')t\,.$$

Proof See [2].

Consider the Riccati differential equation (RDE)

$$\begin{cases} \dot{P}_t^R &= BP_t^R + P_t^R B' + FF' - P_t^R H' DHP_t^R\,, \\ P_0^R &= R\,. \end{cases} \tag{1}$$

where $B \in \mathbb{R}^{n\times n}$, $F \in \mathbb{R}^{n\times q}$, $H \in \mathbb{R}^{p\times n}$, $D \in \mathbb{R}^{p\times p}$, $D > 0$ and $R \in \mathbb{R}^{n\times n}$, $R \geq 0$.

Let $\Phi_t^R \in \mathbb{R}^{n\times n}$ be the transition matrix associated with $B - P_t^R H' H$ i.e. the matrix function $t \to \Phi_t^R$, $t \in \mathbb{R}$, which satisfies

$$\dot{\Phi}_t^R = (B - P_t^R H' H)\Phi_t^R\,, \quad \Phi(0) = I_n\,.$$

Lemma 2.4 *If (B,H) is detectable and (B,F) is stabilizable, then there exists a unique symmetric nonnegative definite solution P_∞ to the algebraic Riccati equation (ARE) associated to the RDE (1).*

$$0 = BP_\infty + P_\infty B' + FF' - P_\infty H' DHP_\infty\,, \tag{2}$$

such that $B - P_\infty H' H$ is asymptotically stable. Moreover for any initial condition $P_0^R = R$ and any $0 < \sigma < -Re\lambda_1(B - P_\infty H' H)$ there exist a constant C_σ and a $t_\sigma < \infty$ such that

$$||P_t^R - P_\infty|| \leq C_\sigma e^{-\sigma t}\,, \forall t \geq 0\,, \tag{3}$$

and

$$||\Phi_t^R (\Phi_s^R)^{-1}|| \leq C_\sigma e^{-\sigma t}\,, \forall t_\sigma \leq s < t\,, \tag{4}$$

where P_t^R is the solution of the RDE (1) and Φ_t^R is the transition matrix associated with $B - P_t^R H' H$.

Proof See [7] for (3) and see e.g. [10] for (4).

Now we state the result of Ocone and Pardoux [10] about the stability of the Kalman filter with respect to its initial condition in the linear case.

Consider the following filtering model:

$$\begin{cases} dZ_t &= BZ_t\, dt + F\, dV_t\,, \\ dY_t &= HZ_t\, dt + dW_t\,, \quad Y_0 = 0\,, \end{cases} \tag{5}$$

where $B \in \mathbb{R}^{n\times n}$, $F \in \mathbb{R}^{n\times q}$, $H \in \mathbb{R}^{p\times n}$, and V and W are independent standard Brownian motions taking values in $\mathbb{R}^q$ and $\mathbb{R}^p$, respectively. Z_0 is assumed to be a random vector independent of (V, W).

Let $(Z_t^{z,R}, P_t^R)$ be the solution to the Kalman filtering equations corresponding to (5) initialized with (z, R), where $z \in \mathbb{R}^n$ and R is a symmetric nonnegative definite $n \times n$–matrix:

$$\left\{\begin{array}{rcl} dZ_t^{z,R} & = & BZ_t^{z,R}\,dt + P_t^R H'[dY_t - HZ_t^{z,R}dt]\,, \\ Z_0^{z,R} & = & z\,, \end{array}\right. \tag{6}$$

$$\left\{\begin{array}{rcl} \dot{P}_t^R & = & BP_t^R + P_t^R B' + FF' - P_t^R H' H P_t^R\,, \\ P_0^R & = & R\,. \end{array}\right. \tag{7}$$

For any (z, R), the distribution $\mathcal{N}(Z_t^{z,R}, P_t^R)$ will be referred to as the Kalman filter driven by the filtering equations (6)-(7) initialized with the prior distribution $\mathcal{N}(z, R)$ or, more simply, a Kalman filter initialized with (z, R).

It is well known that in model (5) if Z_0 is normal with mean z_0 and covariance matrix R_0, then the conditional law of Z_t given $\mathcal{Y}_t$ is Gaussian and the conditional mean $\hat{Z}_t$ and the conditional covariance are given by:

$$\hat{Z}_t = Z_t^{z_0,R_0}\ ,\ \mathbf{E}[(Z_t - \hat{Z}_t)(Z_t - \hat{Z}_t)'] = P_t^{R_0}\,.$$

In other words, in this case the Kalman filter $\mathcal{N}(Z_t^{z_0,R_0}, P_t^{R_0})$ is the optimal filter in model (5).

In the case where the initial condition Z_0 is non–Gaussian, then in general the optimal filter is not a Kalman filter but it can be expressed in terms of an appropriate modified Kalman filter (see Makowski [8] and Benes and Karatzas [1]).

Actually, under stabilizability and detectability assumptions, whatever is the distribution of Z_0 provided that it has a second–order moment, asymptotically in time the optimal filter is approached by any Kalman filter.

These properties are stated more precisely by Ocone and Pardoux in [10] as follows:

Theorem 2.5 *Assume that in model (5), (B, H) is detectable, (B, F) is stabilizable and $\mathbf{E}\left[|Z_0|^2\right] < \infty$. Then*

$$\lim_{t \to \infty} \hat{Z}_t - Z_t^{z,R} = 0 \quad \textit{almost surely,} \tag{8}$$

and in the L^2 sense for any $z \in \mathbb{R}^n$, $R \geq 0$. Moreover, if Π_t denotes the conditional distribution of Z_t given $\mathcal{Y}_t$ then

$$\lim_{t \to \infty} \Pi_t(\varphi) - \mathcal{N}(Z_t^{z,R}, P_t^R)(\varphi) = 0 \quad \textit{almost surely,} \tag{9}$$

for every bounded, uniformly continuous function $\varphi : \mathbb{R}^n \to \mathbb{R}$.

Proof See Theorem 2.6 in [10].

Remark 2.6 The above statement says in particular that the Kalman filter is "asymptotically insensitive to perturbations of its initial condition" in the sense that if the true initial condition is Gaussian then a Kalman filter initialized with an incorrect Gaussian prior distribution approaches the optimal Kalman filter asymptotically in time. Due to Lemma 2.4, so does for instance the Kalman filter initialized with (z, P_∞) where P_∞ is the solution of the ARE (2) associated with the RDE (7). Moreover, of course in this specific filter one has $P_t^{P_\infty} \equiv P_\infty$.

2.2 The stabilizability canonical form and the reduced–-order Kalman filter

In order to construct a reduced–order Kalman filter in model (5) "which corrects in unstable directions only", we introduce a change of state coordinates so that these directions appear clearly. Relative to these coordinates, the form of the state equation in (5) will be referred to as the *stabilizability canonical form.*

Let $Spect(B)$, the spectrum of B, be divided into two disjoint subsets, S_- and S_+, as follows:

$$Spect(B) = S_- \cup S_+ = \{\lambda_1, \cdots, \lambda_k\} \cup \{\mu_1, \cdots, \mu_d\},$$

The subset S_- is formed by the eigenvalues with strictly negative real parts. It is convenient to refer to these k eigenvalues as the "stable poles" of the system and to the associated real linear subspace E_- as the stable subspace. The subset S_+ is formed by the eigenvalues with nonnegative real parts; these d eigenvalues will be referred to as the "unstable poles". Similarly, we define the unstable subspace E_+ as the real linear subspace associated with the unstable poles. It is well known that

$$\mathbb{R}^n = E_- \oplus E_+ \,,$$

and E_- and E_+ are two B–invariant subspaces (See e.g. [7]).

Throughout, we shall suppose:

$(A2_1)$ The d unstable poles are controllable and observable; that is,

Controllable: $rank_{\mathbf{C}}([\mu_i I_n - B|F]) = n$, for $i = 1, \cdots, d$,

Observable: $rank_{\mathbf{C}} \begin{pmatrix} \mu_i I_n - B \\ H \end{pmatrix} = n$, for $i = 1, \cdots, d$.

Now, in order to decouple the dynamics of the state components associated with the stable poles from that of state components associated with the unstable poles, we introduce the following change of basis :

$$Z = [U_-|U_+] \begin{bmatrix} X^- \\ X^+ \end{bmatrix} ,$$

where $U_- \in \mathbb{R}^{n\times k}$, $U_+ \in \mathbb{R}^{n\times d}$ and the matrix $U = [U_-|U_+]$ is orthogonal (i.e. $U'U = I_n$) and such that the columns of U_+ span the unstable subspace E_+. Of course, X^- is the projection of Z on the stable subspace E_- along E_+ and X^+ is the projection of Z on the unstable subspace E_+ along E_-.

The orthogonal matrix U can be obtained from a real Schur decomposition of the matrix B. In the orthonormal basis formed of columns of matrix U, the system (5) is rewritten as

$$\left\{ \begin{array}{lcl} dX_t &=& U'BUX_t\,dt + U'F\,dV_t\,, \\ dY_t &=& HUX_t\,dt + dW_t\,, \quad Y_0 = 0\,, \end{array} \right. \tag{10}$$

where

$$U'BU = \begin{pmatrix} U'_-BU_- & U'_-BU_+ \\ U'_+BU_- & U'_+BU_+ \end{pmatrix} .$$

Define $B_+ = U'_+BU_+$. The equality $U'_+U_+ = I_d$ implies $BU_+ = U_+B_+$ and we obtain $U'_-BU_+ = O_{k\times d}$.

Then the state equation of system (5) is transformed into the *stabilizability canonical form*

$$\left\{ \begin{array}{rcl} dX_t &=& B_U X_t\,dt + F_U\,dV_t\,, \\ dY_t &=& H_U X_t\,dt + dW_t\,, \quad Y_0 = 0\,, \end{array} \right. \tag{11}$$

where

$$B_U \triangleq \begin{pmatrix} B_- & O \\ B_\pm & B_+ \end{pmatrix} \triangleq \begin{pmatrix} U'_-BU_- & O \\ U'_+BU_- & U'_+BU_+ \end{pmatrix}, \; F_U \triangleq \begin{pmatrix} F_- \\ F_+ \end{pmatrix} \triangleq \begin{pmatrix} U'_-F \\ U'_+F \end{pmatrix},$$

and

$$H_U \triangleq (H_-|H_+) \triangleq (HU_-|HU_+) \;.$$

Note that the $k\times k$–matrix B_- is asymptotically stable and that the eigenvalues of B_+ are the unstable poles of the system.

To summarize, the study of the linear filtering problem (5) under $(A2_1)$ is equivalent to the study of the problem (11) under the following assumption:

$(A2_2)$: The matrix B_- is asymptotically stable, (B_+, F_+) is a controllable pair and (B_+, H_+) is an observable pair.

Note that under $(A2_2)$ the system (11) is stabilizable and detectable.

From here, we shall consider the linear filtering problem in the *stabilizability canonical form* and the subscript U will be dropped in the notations B_U, F_U and H_U.

The reduced–order Kalman filter

According to the above considerations, we are interested in the following linear filtering problem

$$\left\{ \begin{array}{rcl} dX_t^- &=& B_-X_t^-\,dt + F_-\,dV_t\,, \\ dX_t^+ &=& B_+X_t^+\,dt + B_\pm X_t^-\,dt + F_+\,dV_t\,, \\ dY_t &=& H_+X_t^+\,dt + H_-X_t^-\,dt + dW_t\,, \quad Y_0 = 0\,, \end{array} \right. \tag{12}$$

where $X = (X^{-'}, X^{+'})'$ is the n–dimensional state process with $n = k+d$ and Y is the p–dimensional observation process, V and W are independent standard Brownian motions, X_0 is a random vector assumed independent of (V, W), $B_- \in \mathbb{R}^{k\times k}$, $B_+ \in \mathbb{R}^{d\times d}$, $B_\pm \in \mathbb{R}^{d\times k}$,$F_- \in \mathbb{R}^{k\times q}$, $F_+ \in \mathbb{R}^{d\times q}$, $H_- \in \mathbb{R}^{p\times k}$, $H_+ \in \mathbb{R}^{p\times d}$.

Let us define $((\bar{X}_t^-, \bar{X}_t^+), \bar{P}_t^+)$ as the solution to the following equations:

$$\left\{ \begin{array}{rcl} d\bar{X}_t^- &=& B_-\bar{X}_t^- dt\,, \\ d\bar{X}_t^+ &=& B_+\bar{X}_t^+\, dt + B_\pm\bar{X}_t^-\, dt \\ && +\bar{P}_t^+ H_+'[dY_t - H_+\bar{X}_t^+ dt - H_-\bar{X}_t^- dt]\,, \\ \bar{X}_0 &=& \bar{x}_0\,, \end{array} \right. \tag{13}$$

$$\left\{ \begin{array}{rcl} \dot{\bar{P}}_t^+ &=& B_+\bar{P}_t^+ + \bar{P}_t^+ B_+' + F_+F_+' - \bar{P}_t^+ H_+' H_+ \bar{P}_t^+\,, \\ \bar{P}_0^+ &=& \bar{R}^+\,. \end{array} \right. \tag{14}$$

where $\bar{x}_0 \in \mathbb{R}^n$ and $\bar{R}^+$ is a symmetric nonnegative definite $d\times d$–matrix.

The above equations can be written

$$\left\{ \begin{array}{rcl} d\bar{X}_t &=& B\bar{X}_t\, dt + \bar{P}_t H'[dY_t - H\bar{X}_t dt]\,, \\ \bar{X}_0 &=& \bar{x}_0\,, \end{array} \right. \tag{15}$$

$$\left\{ \begin{array}{rcl} \dot{\bar{P}}_t &=& B\bar{P}_t + \bar{P}_t B' + \bar{F}\bar{F}' - \bar{P}_t H' H \bar{P}_t\,, \\ \bar{P}_0 &=& \bar{R}\,, \end{array} \right. \tag{16}$$

with

$$B = \begin{pmatrix} B_- & O \\ B_\pm & B_+ \end{pmatrix} ; H = (H_-|H_+)\ ,$$

and

$$\bar{X}_t = (\bar{X}_t^{-'}, \bar{X}_t^{+'})'\ ,\ \bar{R} = \begin{pmatrix} 0 & 0 \\ 0 & \bar{R}^+ \end{pmatrix}\ ,\ \bar{P}_t = \begin{pmatrix} 0 & 0 \\ 0 & \bar{P}_t^+ \end{pmatrix}\ \text{and}\ \bar{F} = \begin{pmatrix} 0 \\ F_+ \end{pmatrix}.$$

Notice that the algorithm (13)–(14), or equivalently (15)–(16), corrects components in $\bar{X}$ which correspond to the unstable signal component X^+ only. The distribution $\mathcal{N}(\bar{X}_t, \bar{P}_t)$ will be referred to as a reduced-order Kalman filter for the filtering problem defined by (12), initialized with $(\bar{x}_0, \bar{P}_0)$.

2.3 Asymptotic behaviour of the reduced–order Kalman filter

Under the assumption $(A2_2)$ (then the system is detectable and stabilizable), it follows readily from Lemma 2.4 that there exists a unique solution $\bar{P}_\infty \geq 0$ to the ARE associated with the RDE (16) which is nothing but

$$\bar{P}_\infty = \begin{pmatrix} 0 & 0 \\ 0 & \bar{P}^+_\infty \end{pmatrix} ,$$

where $\bar{P}^+_\infty \geq 0$ is the unique solution to the ARE associated to (14).

Moreover the matrix $B - \bar{P}_\infty H^{'} H$ is asymptotically stable,

$$\bar{P}_t \rightarrow \bar{P}_\infty \text{ as } t \rightarrow \infty ,$$

and for any initial condition $\bar{P}_0 \geq 0$ and for any $0 < \sigma < -\mathrm{Re}\lambda_1(B - \bar{P}_\infty H^{'} H)$, there is a constant C_σ such that

$$||\bar{P}_t - \bar{P}_\infty|| \leq C_\sigma e^{-\sigma t} , \ \forall t \geq 0 .$$

Remark 2.7 Like the Kalman filter, the reduced-order Kalman filter is "asymptotically insensitive to perturbations of its initial condition" in the sense that for two different initializations the long time behaviour of the filter is unchanged. In particular for the reduced-order Kalman filter initialized with $(\bar{x}_0, \bar{P}_\infty)$ where $\bar{P}_\infty$ is the solution of the ARE associated with (16), of course one has $\bar{P}_t \equiv \bar{P}_\infty$.

In the sequel, we will consider the case "$F_- = O$" and the case "$F_- \neq O$".

A) The case "$F_- = O$"

Here we consider the particular case where the stable component is not noisy in the sense that $F_- = 0$.

In this case, if we initialize the Kalman equations (6)–(7) with $z = (0, x^+)$ where $x^+ \in \mathbb{R}^d$ and

$$R = \begin{pmatrix} 0 & 0 \\ 0 & \bar{R}^+ \end{pmatrix} , \ \bar{R}^+ \in \mathbb{R}^{d \times d} ,$$

it is readily seen that the corresponding Kalman filter coincides with the just defined reduced–order Kalman filter initialized in the same way (i.e. $\bar{x}_0 = (0, x^+)$ and R as above). In other words

$$\bar{X}_t = Z_t^{(0,x^+),R} \;,\quad \bar{P}_t = P_t^R = \begin{pmatrix} 0 & 0 \\ 0 & \bar{P}_t^+ \end{pmatrix} ,$$

where $\bar{P}_t^+ \in \mathbb{R}^{d\times d}$ is the solution of the Riccati equation (14). From Theorem 2.5, the following result is obtained immediately:

Proposition 2.8 *Suppose that in the model (12) $F_- = 0$, assumption $(A2_2)$ holds and $\mathbf{E}\left[|X_0|^2\right] < \infty$. Then if $\hat{X}_t$ denotes the conditional expectation $E(X_t|\mathcal{Y}_t)$ of $X_t = (X_t^{-'}, X_t^{+'})'$ and if $\bar{X}_t = (\bar{X}_t^{-'}, \bar{X}_t^{+'})'$ is defined by (13)–(14),*

$$\lim_{t\,\to\,\infty} \hat{X}_t - \bar{X}_t = 0 \;\; \textit{almost surely,}$$

and in the L^2 sense for any $\bar{x}_0 \in \mathbb{R}^n$, $R \geq 0$. Moreover, if Π_t denotes the conditional distribution of X_t given $\mathcal{Y}_t$ then

$$\lim_{t\,\to\,\infty} \Pi_t(\varphi) - \mathcal{N}(\bar{X}_t, \bar{P}_t)(\varphi) = 0 \;\; \textit{almost surely}\,,$$

for every bounded, uniformly continuous function $\varphi : \mathbb{R}^n \to \mathbb{R}$.

Therefore in the case where the stable component is not noisy, the reduced–order Kalman filter is asymptotically optimal as time goes to infinity.

B) The case “$F_- \neq O$”

Here we assume that X_0 is Gaussian with mean x_0 and covariance R_0 and we analyze the long time behaviour of the mean square deviation of the state estimation $\bar{X}_t$ provided by a reduced-order Kalman filter from $\hat{X}_t$ i.e. the estimation given by the optimal filter which is here a Kalman filter.

Since the Kalman filter and the reduced-order Kalman filter are both “asymptotically insensitive to perturbations of their initial condition” (cf. Remarks 2.6 and 2.7), the initialization is not significant for our purpose. Hence we choose to reduce our investigation to the case where $\bar{x}_0 = x_0$, $R_0 = P_\infty$ and

$\bar{R}_0 = \bar{P}_\infty$ where P_∞ and $\bar{P}_\infty$ are the solutions of the ARE equations associated with the RDE (7) and (16) respectively. Recall that consequently in the corresponding filters one has $P_t^{P_\infty} \equiv P_\infty$ and $\bar{P}_t \equiv \bar{P}_\infty$.

We obtain

$$\begin{cases} d(\hat{X}_t - \bar{X}_t) & = (B - \bar{P}_\infty H'H)(\hat{X}_t - \bar{X}_t)\, dt + (P_\infty - \bar{P}_\infty)H'\, d\nu_t\,, \\ \hat{X}_0 - \bar{X}_0 & = 0\,, \end{cases} \tag{17}$$

where $\nu_t = Y_t - \int_0^t H\hat{X}_t\, dt$ defines the innovation process, which is a Brownian motion.

Therefore, setting $A = B - \bar{P}_\infty H'H$, we can write

$$\hat{X}_t - \bar{X}_t = \int_0^t e^{A(t-s)}(P_\infty - \bar{P}_\infty)H'\, d\nu_s\,. \tag{18}$$

Let $\Delta_t = \mathbf{E}(\hat{X}_t - \bar{X}_t)(\hat{X}_t - \bar{X}_t)'$. Clearly, Δ_t is the nonnegative definite solution of the following differential Lyapunov matrix equation:

$$\dot{\Delta}_t = A\Delta_t + \Delta_t A' + (P_\infty - \bar{P}_\infty)H'H(P_\infty - \bar{P}_\infty)\,,\ \Delta_0 = O\,. \tag{19}$$

Since A is an asymptotically stable matrix, Δ_t approaches the stationary solution Δ as t tends to infinity. The matrix Δ is the solution of the algebraic Lyapunov equation

$$O = A\Delta + \Delta A' + (P_\infty - \bar{P}_\infty)H'H(P_\infty - \bar{P}_\infty)\,. \tag{20}$$

Since $\mathbf{E}|\hat{X}_t - \bar{X}_t|^2 = \text{ tr } (\Delta_t)$, we have

$$\lim_{t \to \infty} \mathbf{E}|\hat{X}_t - \bar{X}_t|^2 = \text{ tr } (\Delta)\,.$$

Now, the upper bound on the trace of Δ is derived in the case where $F_+F_-' = O$. In this case, note that the noise on the stable component of the system state is independent of the noise on the unstable component.

Set $D = P_\infty - \bar{P}_\infty$ where P_∞ and $\bar{P}_\infty$ are solutions to the ARE associated with (7) and (16) respectively. Then D is solution to

$$O = AD + DA' - DH'HD + (FF' - \bar{F}\bar{F}')\,, \tag{21}$$

where

$$FF' - \bar{F}\bar{F}' = \begin{pmatrix} F_-F_-' & O \\ O & O \end{pmatrix}.$$

Since A is an asymptotically stable matrix, we obtain

$$D = \int_0^{+\infty} e^{At}(DH'HD + FF' - \bar{F}\bar{F}')e^{A't}\, dt\,.$$

Therefore the fact that $FF' - \bar{F}\bar{F}' \geq O$ implies $D = P_\infty - \bar{P}_\infty \geq O$.

Proposition 2.9 *Suppose that in model (12) $F_+F_-' = O$, assumption $(A2_2)$ holds and $\mathbf{E}\left[|X_0|^2\right] < \infty$. Set $A = B - \bar{P}_\infty H'H$. If $\lambda_1(A+A') < 0$ (i.e. the matrix $A + A'$ is asymptotically stable) then*

$$\mathbf{E}\left[\left|\hat{X}_t - \bar{X}_t\right|^2\right] \leq \frac{\mathrm{tr}\ (H'H)\lambda_1^2(F_-F_-')}{|\lambda_1(A+A')|^3}\,.$$

Proof We have

$$\mathrm{tr}\ (\Delta) = \int_0^\infty \mathrm{tr}\ \left(e^{At}DH'HDe^{A't}\right)\, dt\,.$$

Making use of Lemma 2.2 and Lemma 2.3, we get

$$\begin{aligned} \mathrm{tr}\ (\Delta) &\leq \mathrm{tr}\ (DH'HD) \int_0^\infty \lambda_1\left(e^{At}e^{A't}\right)\, dt \\ &\leq \mathrm{tr}\ (H'H)\lambda_1^2(D) \int_0^\infty e^{\lambda_1(A+A')t}\, dt\,. \end{aligned}$$

Under the assumption $\lambda_1(A+A') < 0$, we obtain

$$\mathrm{tr}\ (\Delta) \leq \frac{\mathrm{tr}\ (H'H)\lambda_1^2(D)}{-\lambda_1(A+A')}\,. \tag{22}$$

Let N be the solution of

$$O = AN + NA' + (FF' - \bar{F}\bar{F}')\,. \tag{23}$$

From equations (21) and (23), we obtain

$$N - D = \int_0^{+\infty} e^{At} D H^{'} H D e^{A^{'} t} \, dt \geq 0 \, .$$

Since the matrices D and N are symmetric and nonnegative definite,

$$\lambda_1^2(D) \leq \lambda_1^2(N) \, .$$

Under the assumption $\lambda_1(A+A^{'}) < 0$, the following bound is given by Kamaroff in [5]

$$\lambda_1(N) \leq \frac{\lambda_1(FF^{'} - \bar{F}\bar{F}^{'})}{-\lambda_1(A+A^{'})} = \frac{\lambda_1(F_- F_-^{'})}{-\lambda_1(A+A^{'})} \, .$$

Therefore we get

$$\mathrm{tr}\ (\Delta) \leq \frac{\mathrm{tr}\ (H^{'}H)\lambda_1^2(F_- F_-^{'})}{-\lambda_1^3(A+A^{'})} \, .$$

$\square$

Remark 2.10 Note that this bound is coherent with the result obtained in the previous section in the case $F_- = O$. In the case where $H^{'}H > 0$, we can get a tighter upper bound derived from the spectral norm bounds of algebraic matrix Riccati equations (see e.g. [6]).

3 The semi–linear case

3.1 The model

Let $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ be a filtered probability space. We consider the filtering problem where the state process $Z = (M, X)$ and the observation process Y are defined as follows:

$$\begin{aligned} &M = \{M_t \, , \, t \geq 0\} \textit{is a } \mathrm{I\!R}^k\textit{-valued Markov process progressively} \\ &\textit{measurable with respect to } \{\mathcal{F}_t\}, \end{aligned} \qquad (24)$$

$$dX_t = BX_t\,dt + b_M(M_t)\,dt + F\,dV_t\,, \tag{25}$$

$$dY_t = HX_t\,dt + h_M(M_t)\,dt + \,dW_t\,, \quad Y_0 = 0\,, \tag{26}$$

where $B \in \mathbb{R}^{d\times d}$, $F \in \mathbb{R}^{d\times q}$, $H \in \mathbb{R}^{p\times d}$, and the functions b_M and h_M are from $\mathbb{R}^k$ into $\mathbb{R}^d$ and $\mathbb{R}^p$ respectively. The processes V and W are $(\mathcal{F}_t)$–standard Brownian motions taking values in $\mathbb{R}^q$ and $\mathbb{R}^p$, respectively. We suppose that M, V and W are mutually independent. X_0 is a $\mathcal{F}_0$–measurable random vector independent of (M,V,W).

We assume that :

$(A3_1)$ $\mathbf{E}(|X_0|^2) < \infty$;

$(A3_2)$ $P\left(\sup_{[0,T]}|M_t| < \infty\right) = 1$ for each $T > 0$ and the process M converges towards 0 in the L^2–sense;

$(A3_3)$ The functions $b_M : \mathbb{R}^k \to \mathbb{R}^d$ and $h_M : \mathbb{R}^k \to \mathbb{R}^p$ are Lipschitzian i.e. there exists a constant $C > 0$ such that for all x_1, $x_2 \in \mathbb{R}^k$,

$$|b_M(x_1) - b_M(x_2)| + |h_M(x_1) - h_M(x_2)| \le C|x_1 - x_2|\,, \tag{27}$$

and $b_M(0) = 0$ and $h_M(0) = 0$.

$(A3_4)$ (B,H) is a detectable pair and (B,F) is a stabilizable pair.

Our aim is to estimate $Z_t = (M'_t, X'_t)'$ given the observation up to time t. Since $E(M_t^2)$ converges towards 0 as t goes to infinity, 0 is an "asymptotically efficient" approximation for M_t. Therefore we are only interested in the estimation of X_t. This problem can be interpreted as a linear filtering problem with coloured noises. The optimal filter for X_t is given by

$$\Pi_t^X(\varphi) = \mathbf{E}(\varphi(X_t)\,|\,\mathcal{Y}_t), \text{ for all } \varphi \in \mathcal{C}_b(\mathbb{R}^d)\,.$$

3.2 The approximate filter

In this section, we propose a finite dimensional approximate filter for X_t which approaches the optimal filter as time goes to infinity.

To construct this approximate filter, we do as if M_t was identically equal to 0, for all $t \ge 0$ in the system (25)–(26) i.e. as if the system was

$$\begin{cases} dX_t &= BX_t\,dt + F\,dV_t\,, \\ dY_t &= HX_t\,dt + dW_t\,,\ Y_0 = 0\,. \end{cases}$$

Hence, defining for some $x \in \mathbb{R}^d$ and $R \geq 0$, $(\bar{X}_t^{x,R}, \bar{Q}_t^R)$ as the solutions to the following Kalman type equations,

$$d\bar{X}_t^{x,R} = B\bar{X}_t^{x,R}\,dt + \bar{Q}_t^R H'(dY_t - H\bar{X}_t^{x,R}\,dt)\,,\ \bar{X}_0^{x,R} = x\,, \tag{28}$$

$$\dot{\bar{Q}}_t^R = B\bar{Q}_t^R + \bar{Q}_t^R B' + FF' - \bar{Q}_t^R H' H\bar{Q}_t^R\,,\ \bar{Q}_0^R = R\,, \tag{29}$$

the approximate filter is taken as the Gaussian law with mean $\bar{X}_t^{x,R}$ and with covariance matrix $\bar{Q}_t^R$.

Proposition 3.1 *Suppose that the assumptions $(A3_1) - (A3_4)$ hold. Then*

$$\lim_{t \to \infty} \mathbf{E}\left[\left|\hat{X}_t - \bar{X}_t^{x,R}\right|^2\right] = 0\,, \tag{30}$$

for any $x \in \mathbb{R}^d$ and $R \geq 0$. Moreover, if Π_t^X denotes the conditional law of X_t given $\mathcal{Y}_t$, then

$$\lim_{t \to \infty} \mathbf{E}\left[\left(\Pi_t^X(\varphi) - \mathcal{N}(\bar{X}_t^{x,R}, \bar{Q}_t^R)(\varphi)\right)^2\right] = 0\,, \tag{31}$$

for every bounded, uniformly continuous function $\varphi : \mathbb{R}^d \to \mathbb{R}$.

First let us show the following lemma

Lemma 3.2 *If $\lim_{t \to \infty} x_t - \bar{x}_t = 0$ and if $\lim_{t \to \infty} Q_t = Q_\infty = \lim_{t \to \infty} \bar{Q}_t$, then*

$$\lim_{t \to \infty} \left[\mathcal{N}(x_t, Q_t)(\varphi) - \mathcal{N}(\bar{x}_t, \bar{Q}_t)(\varphi)\right] = 0\,,$$

for every bounded uniformely continuous function $\varphi : \mathbb{R}^d \to \mathbb{R}$.

Proof Clearly,

$$\lim_{t \to \infty} \left[\mathcal{N}(0, Q_t)(\varphi) - \mathcal{N}(0, \bar{Q}_t)(\varphi)\right] = \lim_{t \to \infty} [\mathcal{N}(0, Q_t)(\varphi) - \mathcal{N}(0, Q_\infty)(\varphi)]$$

$$+ \lim_{t \to \infty} \left[\mathcal{N}(0, \bar{Q}_t)(\varphi) - \mathcal{N}(0, Q_\infty)(\varphi)\right] = 0\,.$$

Then on some probability space there exist random processes X and $\bar{X}$ taking values in $\mathbb{R}^d$, such that X_t has law $\mathcal{N}(0, Q_t)$ and $\bar{X}_t$ has law $\mathcal{N}(0, \bar{Q}_t)$ for all $t \geq 0$, and moreover $|X_t - \bar{X}_t| \to 0$ almost surely. Let $Y_t = X_t + x_t$ and $\bar{Y}_t = \bar{X}_t + \bar{x}_t$. Since $\lim_{t \to \infty} x_t - \bar{x}_t = 0$, we obtain $|Y_t - \bar{Y}_t| \to 0$ almost surely. Given $\eta > 0$,

$$|\mathcal{N}(x_t, Q_t)(\varphi) - \mathcal{N}(\bar{x}_t, \bar{Q}_t)(\varphi)| = |\mathbf{E}\left(\varphi(Y_t) - \varphi(\bar{Y}_t)\right)|$$

$$\leq \sup_{|y-y'| \leq \eta} |\varphi(y) - \varphi(y')| + 2||\varphi||_\infty P(\{|Y_t - \bar{Y}_t| \geq \eta\})\,.$$

By first letting $t \to \infty$ and then using uniform continuity, we complete the proof. □

Proof of Proposition 3.1 The proof is similar step by step to the proof of Theorem 2.6 in [10]. The idea is to decompose the signal X and to introduce a new probability measure under which a new filtering problem is obtained. We write

$$X_t = \check{M}_t + \check{X}_t\,,$$

where

$$\begin{aligned} \check{M}_t &= e^{Bt} X_0 + \int_0^t e^{B(t-s)} b_M(M_s)\, ds\,, \\ \check{X}_t &= \int_0^t e^{B(t-s)} F\, dV_s. \end{aligned}$$

Then the observation equation can be rewritten

$$dY_t = H\check{X}_t\, dt + H\check{M}_t\, dt + h_M(M_t)\, dt + dW_t\,.$$

Under $(A3_1)$–$(A3_3)$, since (X_0, M), and hence also $(M, \check{M})$, is independent of W,

$$\mathbf{E}\left(\exp\left\{\int_0^T -(H\check{M}_s + h_M(M_s))'\, dW_s - \frac{1}{2}\int_0^T |H\check{M}_s + h_M(M_s)|^2\, ds\right\}\right) = 1\,,$$

and a new probability measure $\bar{P}$ can be defined by

$$\frac{d\bar{P}}{dP} = \exp\left\{\int_0^T -(H\check{M}_s + h_M(M_s))'\, dW_s - \frac{1}{2}\int_0^T |H\check{M}_s + h_M(M_s)|^2\, ds\right\}\,.$$

On $(\Omega, \bar{P}, \mathcal{F}_T)$, the process $\bar{W} = \{\bar{W}_t\,,\, 0 \leq t \leq T\}$ defined by

$$\bar{W}_t \triangleq \int_0^t (H\check{M}_s + h_M(M_s))\, ds + W_t\,,\ 0 \leq t \leq T\,,$$

is a Brownian motion, and $\{(X_0, M_t)\,,\, t \in [0,T]\}$ is independent of $\{(V_t, \bar{W}_t)\,,\, t \in [0,T]\}$. Moreover the law of $\{(X_0, M_t)\,,\, t \in [0,T]\}$ is the same under P and $\bar{P}$.

Define

$$K_t \triangleq \int_0^t (H\check{M}_s + h_M(M_s))' d\bar{W}_s\ ,\ < K >_t \triangleq \int_0^t |H\check{M}_s + h_M(M_s)|^2\, ds\,,$$

and

$$L_t \triangleq \exp\left\{K_t - \frac{1}{2} < K >_t\right\}\,.$$

Note that $L_T^{-1} = \frac{d\bar{P}}{dP}$. Then, by the Bayes formula, for any nonnegative measurable function $\varphi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^+$, we have

$$\mathbf{E}\left(\varphi(\check{M}_t, \check{X}_t) \mid \mathcal{Y}_t\right) = \frac{\bar{\mathbf{E}}\left(\varphi(\check{M}_t, \check{X}_t) L_t \mid \mathcal{Y}_t\right)}{\bar{\mathbf{E}}(L_t \mid \mathcal{Y}_t)}\,. \tag{32}$$

In the sequel, we shall show that both the numerator and the denominator may be expressed by an integral involving the Kalman filter for the process $(\check{X}_t, K_t)$ given $\mathcal{Y}_t$.

Consider the system

$$\left\{\begin{array}{rcl} d\check{X}_t &=& B\check{X}_t\,dt + F\,dV_t\,; \quad \check{X}_0 = 0\,, \\ dK_t &=& (H\check{M}_t + h_M(M_t))'\,d\bar{W}_t\,; \quad K_0 = 0\,, \\ dY_t &=& H\check{X}_t\,dt + d\bar{W}_t\,, \qquad Y_0 = 0\,, \end{array}\right. \tag{33}$$

where $\check{X}$ is a $\mathbb{R}^d$–valued process, K is a real–valued process and Y is the $\mathbb{R}^p$–valued observation process.

Note that $\{(X_0, M_t)\,,\, t \in [0,T]\}$ and hence also $\{(M_t, \check{M}_t)\,,\, t \in [0,T]\}$, is independent of $\{(\bar{W}_t, Y_t)\,,\, t \in [0,T]\}$ under $\bar{P}$. This system can be rewritten

$$\left\{\begin{array}{rl} d\begin{pmatrix} \check{X}_t \\ K_t \end{pmatrix} &= \begin{pmatrix} B & O \\ -(H\check{M}_t + h_M(M_t))'H & O \end{pmatrix}\begin{pmatrix} \check{X}_t \\ K_t \end{pmatrix} dt \\ & \quad + \begin{pmatrix} O \\ (H\check{M}_t + h_M(M_t))' \end{pmatrix} dY_t + \begin{pmatrix} F \\ O \end{pmatrix} dV_t\,, \\ dY_t &= (H \mid O)\begin{pmatrix} \check{X}_t \\ K_t \end{pmatrix} dt + d\bar{W}_t\,. \end{array}\right.$$

Under $\bar{P}$, the conditional law of $(\check{X}_t, K_t)$ given $\mathcal{Y}_t$ is the Gaussian law $\mathcal{N}((\bar{X}_t, \bar{K}_t), C_t)$ where $(\bar{X}_t, \bar{K}_t)$ is the solution to

$$d\begin{pmatrix} \bar{X}_t \\ \bar{K}_t \end{pmatrix} = \begin{pmatrix} B & O \\ -(H\check{M}_t + h_M(M_t))'H & O \end{pmatrix}\begin{pmatrix} \bar{X}_t \\ \bar{K}_t \end{pmatrix} dt \tag{34}$$

$$+ \begin{pmatrix} O \\ (H\check{M}_t + h_M(M_t))' \end{pmatrix} dY_t + C_t \begin{pmatrix} H' \\ O \end{pmatrix}\left[dY_t - (H \mid O)\begin{pmatrix} \bar{X}_t \\ \bar{K}_t \end{pmatrix} dt\right],$$

with $(\bar{X}_0, \bar{K}_0) = (0,0)$ and $C_t \in \mathbb{R}^{(d+1)\times(d+1)}$ is the nonnegative definite solution to

$$\dot{C}_t = \begin{pmatrix} FF' & O \\ O & O \end{pmatrix} - C_t \begin{pmatrix} H'H & O \\ O & O \end{pmatrix} C_t \tag{35}$$

$$+ \begin{pmatrix} B & O \\ -(H\check{M}_t + h_M(M_t))'H & O \end{pmatrix} C_t + C_t \begin{pmatrix} B & O \\ -(H\check{M}_t + h_M(M_t))'H & O \end{pmatrix}'$$

with $C_0 = O$.

Let

$$C_t \triangleq \begin{pmatrix} \bar{Q}_t & S_t \\ S_t^{'} & T_t \end{pmatrix} .$$

Therefore we have

$$\begin{aligned} \dot{\bar{Q}}_t &= B\bar{Q}_t + \bar{Q}_t B^{'} + FF^{'} - \bar{Q}_t H^{'} H \bar{Q}_t \,,\ \bar{Q}_0 = O \,, & (36) \\ \dot{S}_t &= (B - \bar{Q}_t H^{'} H) S_t - \bar{Q}_t H^{'} (h_M(M_t) + H\check{M}_t) \,,\ S_0 = 0 \,, & (37) \\ \dot{T}_t &= -2(h_M(M_t) + H\check{M}_t)^{'} H S_t - S_t^{'} H^{'} H S_t \,,\ T_0 = 0 \,, & (38) \end{aligned}$$

where $\bar{Q}_t \in \mathbb{R}^{d\times d}$, $S_t \in \mathbb{R}^d$ and $T_t \in \mathbb{R}$. We note that clearly $\bar{X}_t = \bar{X}_t^{0,O}$ and $\bar{Q}_t = \bar{Q}_t^O$. Recall that the law of $\{(X_0, M_t)\,,\, t \in [0,T]\}$ is the same under P and $\bar{P}$ and that $\{(X_0, M_t))\,,\, t \in [0,T]\}$ and $\{(\bar{W}_t, Y_t))\,,\, t \in [0,T]\}$ are $\bar{P}$–independent. Since $< K >_t$ and $\check{M}_t$ depend on $\{X_0, M_s\,,\, 0 \le s \le t\}$ only, the numerator of the expression (32) can be written

$$\begin{aligned} & \bar{\mathbf{E}}\left(\varphi(\check{M}_t, \check{X}_t) \exp\left\{K_t - \tfrac{1}{2} < K >_t\right\} \mid \mathcal{Y}_t\right) \\ = \ & \int_{\mathbb{R}^d \times \mathcal{C}(\mathbb{R}_+, \mathbb{R}^k)} e^{-\frac{1}{2}<K>_t(x,m)} I_t(x,m)\, d\pi_0(x) dP_M(m) \,, \end{aligned} \qquad (39)$$

where

$$I_t(x,m) = \int_{\mathbb{R}^d \times \mathbb{R}} \varphi(\check{M}_t(x,m), r_1) e^{r_2} dn_t(r_1, r_2) \,,$$

and n_t is the conditional law of $(\check{X}_t, K_t)$ given $\mathcal{Y}_t$ i.e. n_t is the Gaussian law with mean $(\bar{X}_t, \bar{K}_t)$ and covariance C_t defined above.

Note that

$$\begin{aligned} & \int_{\mathbb{R}^d \times \mathbb{R}} e^{i<\lambda, r_1> + i\mu r_2 + r_2}\, dn_t(r_1, r_2) \\ = \ & e^{i<\lambda, \bar{Z}_t> + i(\mu - i)\bar{K}_t} \exp\left\{-\frac{1}{2}\left\langle C_t \begin{pmatrix} \lambda \\ \mu - i \end{pmatrix}, \begin{pmatrix} \lambda \\ \mu - i \end{pmatrix}\right\rangle\right\} \qquad (40) \\ = \ & e^{\frac{1}{2}T_t + \bar{K}_t} \exp\left\{i < \lambda, \bar{Z}_t + S_t > + i\mu(\bar{K}_t + T_t) - \frac{1}{2}\left\langle C_t \begin{pmatrix} \lambda \\ \mu \end{pmatrix}, \begin{pmatrix} \lambda \\ \mu \end{pmatrix}\right\rangle\right\} , \end{aligned}$$

with $\mu \in \mathbb{R}^d$ and $\lambda \in \mathbb{R}$. Let

$$d\bar{n}_t(r_1, r_2) \triangleq e^{-\frac{1}{2}T_t - \bar{K}_t} e^{r_2} dn_t(r_1, r_2) \,.$$

Following the above equalities, we get that the characteristic functional of $\bar{n}_t$ is given by

$$\int_{\mathbb{R}^d\times\mathbb{R}} e^{i<\lambda,r_1>+i\mu r_2}\, d\bar{n}_t(r_1,r_2)$$
$$= \exp\left\{i\,<\lambda,\bar{Z}_t+S_t> + i\,\mu(\bar{K}_t+T_t) - \frac{1}{2}\left\langle C_t\begin{pmatrix}\lambda\\ \mu\end{pmatrix},\begin{pmatrix}\lambda\\ \mu\end{pmatrix}\right\rangle\right\}.$$

Therefore $\bar{n}_t$ is the Gaussian law with mean

$$\begin{pmatrix}\bar{Z}_t+S_t\\ \bar{K}_t+T_t\end{pmatrix},$$

and covariance C_t. Then by (32) and (39), we obtain

$$\mathbf{E}[\varphi(X_t)\,|\,\mathcal{Y}_t] = \mathbf{E}[\varphi(\check{M}_t+\check{X}_t)\,|\,\mathcal{Y}_t] =$$
$$\frac{\int_{\mathbb{R}^d\times\mathcal{C}(\mathbb{R}_+,\mathbb{R}^k)} e^{\left(-\frac{1}{2}<K>_t+\frac{1}{2}T_t+\bar{K}_t\right)(x,m)} J_t(x,m)d\pi_0(x)dP_M(m)}{\int_{\mathbb{R}^d\times\mathcal{C}(\mathbb{R}_+,\mathbb{R}^k)} e^{\left(-\frac{1}{2}<K>_t+\frac{1}{2}T_t+\bar{K}_t\right)(x,m)} d\pi_0(x)dP_M(m)}, \tag{41}$$

where

$$J_t(x,m) = \int_{\mathbb{R}^d\times\mathbb{R}} \varphi(\check{M}_t(x,m)+r_1)d\bar{n}_t(r_1,r_2).$$

Since under assumptions $(A3_1)$–$(A3_3)$, $\check{M}_t$ and $\check{X}_t$ are integrable, the equality (41) is satisfied in the case $\varphi(x)=x$. In this case,

$$\begin{aligned}\int_{\mathbb{R}^d\times\mathbb{R}}\varphi\left(\check{M}_t(x,m)+r_1\right)d\bar{n}_t(r_1,r_2) &= \check{M}_t(x,m)+\int_{\mathbb{R}^d} r_1 d\bar{n}^1_t(r_1)\\ &= \check{M}_t(x,m)+\bar{X}_t+S_t(x,m),\end{aligned} \tag{42}$$

where $\bar{n}^1_t$ is the marginal of $\bar{n}_t$ on $\mathbb{R}^d$ i.e $\bar{n}^1_t$ is a Gaussian law with mean $\bar{X}_t+S_t$ and covariance $\bar{Q}_t$.

From (41) and (42) we get

$$\hat{X}_t = \bar{X}_t + \mathbf{E}[\check{M}_t+S_t\,|\,\mathcal{Y}_t]. \tag{43}$$

Let us consider $\mathbf{E}[\check{M}_t+S_t\,|\,\mathcal{Y}_t]$. From (34) and (37), we obtain

$$d(\check{M}_t+S_t) = (B-\bar{Q}_tH'H)(\check{M}_t+S_t)\,dt + [b_M(M_t)-\bar{Q}_tH'h_M(M_t)]\,dt, \tag{44}$$

with $\check{M}_0 + S_0 = X_0$. Thus

$$\check{M}_t + S_t = \Phi_t X_0 + \Phi_t \int_0^t \Phi_s^{-1}(b_M(M_s) - \bar{Q}_s H' h_M(M_s))\, ds\,, \tag{45}$$

where $\Phi_t \in \mathbb{R}^{d\times d}$ is the transition matrix associated with $B - \bar{Q}_t H' H$. Under the assumption $(A3_4)$ ((B, H) detectable and (B, F) stabilizable), according to Lemma 2.4 there exists a unique $\bar{Q}_\infty \geq 0$ solution to the ARE

$$0 = B\bar{Q}_\infty + \bar{Q}_\infty B' + F'F - \bar{Q}_\infty H' H \bar{Q}_\infty\,,$$

such that $\lim_{t\,\to\,\infty} \bar{Q}_t = \bar{Q}_\infty$ and the matrix $B - \bar{Q}_\infty H' H$ is asymptotically stable. Moreover, for every $0 < \sigma < -\mathrm{Re}\lambda_1(B - \bar{Q}_\infty H' H)$, there exist a constant $C_\sigma > 0$ and a $t_\sigma < \infty$ such that

$$||\Phi_t \Phi_s^{-1}|| \leq C_\sigma e^{-\sigma(t-s)}\,,\ \text{for } t_\sigma \leq s < t\,, \tag{46}$$

and

$$||\bar{Q}_t - \bar{Q}_\infty|| \leq C_\sigma e^{-\sigma t}\,,\ \text{for all } t \geq 0\,. \tag{47}$$

Then the equation (45) can be rewritten

$$\begin{array}{rcl} \check{M}_t + S_t & = & \Phi_t \Phi_{t_\sigma}^{-1}(\check{M}_{t_\sigma} + S_{t_\sigma}) \\ & + & \Phi_t \int_{t_\sigma}^t \Phi_s^{-1}(b_M(M_s) - \bar{Q}_\infty H' h_M(M_s))\, ds \\ & + & \Phi_t \int_{t_\sigma}^t \Phi_s^{-1}(\bar{Q}_\infty - \bar{Q}_s) H' h_M(M_s)\, ds\,. \end{array} \tag{48}$$

Hence

$$\mathbf{E}[|\check{M}_t + S_t|^2] \leq 4\mathbf{E}[|\Phi_t \Phi_{t_\sigma}^{-1}(\check{M}_{t_\sigma} + S_{t_\sigma})|^2]$$

$$+ \quad 4C_\sigma \int_{t_\sigma}^t e^{-2\sigma(t-s)} \mathbf{E}[|M_s|^2]\, ds + 4C_\sigma \int_{t_\sigma}^t e^{-2\sigma(t-s)} e^{-2\sigma s} \mathbf{E}[|M_s|^2]\, ds\,.$$

Then there exist a constant $C_\sigma > 0$ and a $t_\sigma < \infty$ such that

$$\mathbf{E}[|\check{M}_t + S_t|^2] \leq C_\sigma \left[e^{-2\sigma(t-t_\sigma)} + \int_{t_\sigma}^t e^{-2\sigma(t-s)} \mathbf{E}[|M_s|]^2\, ds \right]\,. \tag{49}$$

Under the assumption $(A3_2)$, making use of the Tœplitz Lemma we obtain

$$\lim_{t\,\to\,\infty} \mathbf{E}[|\check{M}_t + S_t|^2] = 0\,. \tag{50}$$

Since $\hat{X}_t - \bar{X}_t = \mathbf{E}(\check{M}_t + S_t \,|\, \mathcal{Y}_t)$, by Jensen's inequality, it follows that

$$\lim_{t \to \infty} \mathbf{E}[|\hat{X}_t - \bar{X}_t|^2] = 0\,. \tag{51}$$

To complete the proof of (30), we have to show that $\bar{X}_t - \bar{X}_t^{x,R}$ converges towards 0 in L^2–sense for any $x \in \mathbb{R}^d$ and $R \geq 0$. More precisely, we are going to show that for any $x \in \mathbb{R}^d$, $R \geq 0$ and for any $0 < \sigma < -\mathrm{Re}\lambda_1(B - \bar{Q}_\infty H'H)$ there exist a constant C_σ and a $t_\sigma < \infty$ such that

$$\mathbf{E}\left[\left|\bar{X}_t - \bar{X}_t^{x,R}\right|^2\right] \leq C_\sigma e^{-2\sigma t}\,,\ \forall t > t_\sigma\,. \tag{52}$$

Recall that $\bar{Q} = \bar{Q}^O$ and $\bar{X}_t = \bar{X}_t^{0,O}$. From equations (34) and (28) we obtain

$$\left\{\begin{array}{rcl} d(\bar{X}_t - \bar{X}_t^{x,R}) & = & (B - \bar{Q}_t^R H'H)(\bar{X}_t - \bar{X}_t^{x,R})\,dt + (\bar{Q}_t - \bar{Q}_t^R)H'\,d\nu_t \\ & & +(\bar{Q}_t - \bar{Q}_t^R)H'H(\hat{X}_t - \bar{X}_t)\,dt \\ & & +(\bar{Q}_t - \bar{Q}_t^R)H'h_M\widehat{(M_t)}\,dt\,, \\ \bar{X}_0 - \bar{X}_0^{x,R} & = & -x\,, \end{array}\right.$$

where $\nu_t = Y_t - \int_0^t (H\hat{X}_s + h_M\widehat{(M_s)})\,ds$ defines the inovation process. Then for a given $t_\sigma > 0$, we have

$$\begin{array}{rcl} \bar{X}_t - \bar{X}_t^{x,R} & = & \Phi_t^R(\Phi_{t_\sigma}^R)^{-1}(\bar{X}_{t_\sigma} - \bar{X}_{t_\sigma}^{x,R}) + \Phi_t^R \displaystyle\int_{t_\sigma}^t (\Phi_s^R)^{-1}(\bar{Q}_s - \bar{Q}_s^R)H'\,d\nu_s \\ & & +\Phi_t^R \displaystyle\int_{t_\sigma}^t (\Phi_s^R)^{-1}(\bar{Q}_s - \bar{Q}_s^R)H'H(\hat{X}_s - \bar{X}_s) \\ & & +\Phi_t^R \displaystyle\int_{t_\sigma}^t (\Phi_s^R)^{-1}(\bar{Q}_s - \bar{Q}_s^R)H'h_M\widehat{(M_s)}\,ds\,, \end{array}$$

where Φ_t^R is the transition matrix associated with $B - \bar{Q}_t^R H'H$.

To prove (52) it suffices to establish the bound for each of the terms in the above right-hand side. The first two terms are treated as in (48) respectively, the bound for the last two terms is easily obtained. This completes the proof of (30).

Finally, let us show (31). From (52), by applying the Borel-Cantelli lemma to $\bar{X}_n - \bar{X}_n^{x,R}$ and to $\sup_{n\leq t\leq n+1} |\bar{X}_t - \bar{X}_t^{x,R} - (\bar{X}_n - \bar{X}_n^{x,R})|$, we can easily obtain that $\bar{X}_t - \bar{X}_t^{x,R}$ converges almost surely towards 0. Since

$$\lim_{t \to \infty} \bar{Q}_t = \bar{Q}_\infty = \lim_{t \to \infty} \bar{Q}_t^R ,$$

according to Lemma 3.2 and the dominated convergence Theorem, we obtain

$$\lim_{t \to \infty} \mathbf{E}\left[\left(\mathcal{N}(\bar{X}_t, \bar{Q}_t)(\varphi) - \mathcal{N}(\bar{X}_t^{x,R}, \bar{Q}_t^R)(\varphi)\right)^2\right] = 0 .$$

Therefore it remains to show that

$$\lim_{t \to \infty} \mathbf{E}\left[\left(\Pi_t(\varphi) - \mathcal{N}(\bar{X}_t, \bar{Q}_t)(\varphi)\right)^2\right] = 0 , \tag{53}$$

for any bounded, uniformely continuous function $\varphi : \mathbb{R}^d \to \mathbb{R}$. Observe that

$$\Pi_t(\varphi) - \mathcal{N}(\bar{X}_t, \bar{Q}_t)(\varphi) = \mathbf{E}(\varphi(X_t) \,|\, \mathcal{Y}_t) - \mathbf{E}(\varphi(\bar{X}_t + U_t)) ,$$

where for each $t \geq 0$, U_t is a Gaussian vector with zero–mean and covariance matrix $\bar{Q}_t$.

From (41), we have

$$\Pi_t(\varphi) - \mathcal{N}(\bar{X}_t, \bar{Q}_t)(\varphi) = \frac{\int_{\mathbb{R}^d \times \mathcal{C}(\mathbb{R}_+, \mathbb{R}^k)} e^{\left(-\frac{1}{2}<K>_t + \frac{1}{2}T_t + \bar{K}_t\right)(x,m)} \Delta_t(x,m) d\pi_0(x) dP_M(m)}{\int_{\mathbb{R}^d \times \mathcal{C}(\mathbb{R}_+, \mathbb{R}^k)} e^{\left(-\frac{1}{2}<K>_t + \frac{1}{2}T_t + \bar{K}_t\right)(x,m)} d\pi_0(x) dP_M(m)} , \tag{54}$$

where

$$\Delta_t(x,m) = \mathbf{E}\left(\varphi(\bar{X}_t + (S_t + \check{M}_t)(x,m) + U_t)\right) - \mathbf{E}\left(\varphi(\bar{X}_t + U_t)\right) .$$

Let $\varepsilon > 0$. Choose $\eta > 0$ such that for all y and $y' \in \mathbb{R}^d$, if $|y - y'| \leq \eta$ then

$$|\varphi(y) - \varphi(y')| < \frac{\varepsilon}{2} .$$

Decompose the integral in the numerator of (54) into the sum of the integral over the region $\{|(\check{M}_t+S_t)(x,m)| < \eta\}$ and the integral over the region $\{|(\check{M}_t+S_t)(x,m)| \geq \eta\}$. We obtain

$$\begin{aligned} |\Pi_t(\varphi) - \mathcal{N}(\bar{X}_t, \bar{Q}_t)(\varphi)| \quad &\leq \quad \sup_{|y-y'|<\eta} |\varphi(y) - \varphi(y')| \,, \\ &\qquad +2||\varphi||_\infty \mathbf{E}\left(\mathbf{1}_{\{|(\check{M}_t+S_t)|\geq\eta\}} \,|\, \mathcal{Y}_t\right) \,, \end{aligned}$$

and therefore

$$\mathbf{E}\left(|\Pi_t(\varphi) - \mathcal{N}(\bar{X}_t, \bar{Q}_t)(\varphi)|^2\right) \leq \frac{\varepsilon}{2} + \frac{2||\varphi||_\infty}{\eta^2}\mathbf{E}\left[\left|\check{M}_t + S_t\right|^2\right] \,.$$

Since $\lim_{t\,\to\,\infty} \mathbf{E}\left[\left|\check{M}_t + S_t\right|^2\right] = 0$, there is a $T > 0$ such that for all $t > T$

$$\mathbf{E}\left[\left|\check{M}_t + S_t\right|^2\right] \leq \frac{\varepsilon}{2} \,,$$

and we obtain (53) which achieves the proof of (31).

□

Remark 3.3 One consequence of Proposition 3.1 is that the approximate filter is "asymptotically insensitive to perturbations of its initial condition". It is interesting to note that this result does not require asymptotic ergodic behavior of the signal. The statement extends to the case of a coloured noise in the signal and in the observation the one obtained by Ocone and Pardoux (see Theorem 2.6 [10]) for a white noise. Moreover it can be seen also as some extension of Proposition 2.8 above to the case of a nonlinear action of the stable signal component on the unstable one.

4 The nonlinear case

4.1 The result of Ocone–Pardoux [10]

First we present the result of Ocone–Pardoux [10] about the stability of the optimal filter with respect to its initial condition when the state process Z is ergodic.

Consider the filtering model specified by the signal–observation pair (Z, Y) defined on the probability space $(\Omega, \mathcal{F}, P)$ as follows:

$Z = \{Z_t\,,\, t \geq 0\}$ is a continuous, $\mathbb{R}^n$–valued Markov process with law P^{π_0} corresponding to the initial distribution π_0 and

$$Y_t = \int_0^t h(Z_s)\, ds \, + \, W_t\,,\ t \geq 0\,,$$

where W is a $\mathbb{R}^p$–valued Brownian motion independent of Z and $h : \mathbb{R}^n \to \mathbb{R}^p$ is bounded and continuous. Let $(S_t)_{t\geq 0}$ denote the semigroup on $\mathcal{C}_b(\mathbb{R}^d)$ associated with the Markov process Z,

$$S_t\varphi(z) = \mathbf{E}^z(\varphi(Z_t))\,, \quad \forall \varphi \in \mathcal{C}_b(\mathbb{R}^n)\,.$$

Assume that $(S_t)_{t\geq 0}$ is a strongly continuous positive and conservative contraction semigroup on $\mathcal{C}_b(\mathbb{R}^n)$. In [10], the following ergodicity properties are introduced for the signal semigroup.

The semigroup $(S_t)_{t\geq 0}$ admits a unique invariant measure μ and

(H_1) $\limsup_{t \to +\infty} \int_{\mathbb{R}^n} |S_t\varphi(z) - \mu(\varphi)|\, \mu(dz) = 0\,, \quad \forall \varphi \in \mathcal{C}_b(\mathbb{R}^n)$.

For a probability measure ν on $\mathbb{R}^n$, let νS_t (resp. R^ν) denote the law of the random variable Z_t (resp. the observation process Y) when the distribution of Z_0 is ν. One says that $(S_t)_{t\geq 0}$ forgets ν for μ if

$(H_2(\nu))$ $\nu S_t \to \mu$ weakly as $t \to +\infty$.

Then Theorem 3.2 in [10] stands as follows:

Theorem 4.1 *Assume*

(i) $(S_t)_{t\geq 0}$ *satisfies* $(H1)$*;*

(ii) $(H_2(\pi_0))$ *and* $(H_2(\bar{\pi}_0))$ *are both satisfied;*

(iii) R^{π_0} *is absolutely continuous with respect to* $R^{\bar{\pi}_0}$.

If Π_t *(resp.* $\bar{\Pi}_t$*) denotes the conditional distribution of* Z_t *given* $\mathcal{Y}_t$ *when the initial distribution of* Z *is* π_0 *(resp.* $\bar{\pi}_0$*), then*

$$\lim_{t \to +\infty} \mathbf{E}\left[\left(\Pi_t(\varphi) - \bar{\Pi}_t(\varphi)\right)^2\right] = 0\,,$$

for every bounded, continuous $\varphi : \mathbb{R}^p \to \mathbb{R}$.

Remark 4.2 This result has been established by Ocone and Pardoux [10] in a more general context where the state process Z is càdlàg and takes its values in a locally compact, complete separable metric space.

4.2 Problem formulation

In this section, the following nonlinear filtering problem is studied:

$$dM_t = f(M_t)\,dt\,, \tag{55}$$

$$dX_t = b(X_t)\,dt + b_M(M_t)\,dt + dV_t\,, \tag{56}$$

$$dY_t = h(X_t)\,dt + h_M(M_t)\,dt + dW_t\,, \quad Y_0 = 0\,, \tag{57}$$

where $Z = (M', X')'$ is the n–dimensional state process with $n = k + d$ and Y is the p–dimensional observation process. The processes V and W are two independent standard Brownian motions taking values in $\mathbb{R}^d$ and $\mathbb{R}^p$, respectively, and defined on the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ whose precise form will be given below. Z_0 is a $\mathcal{F}_0$–measurable random vector independent of (V, W) with the given law π_0. The functions f, b_M and h_M are from $\mathbb{R}^k$ into $\mathbb{R}^k$, $\mathbb{R}^d$ and $\mathbb{R}^p$ respectively. The functions b and h are from $\mathbb{R}^d$ into $\mathbb{R}^d$ and $\mathbb{R}^p$ respectively.

Let ϕ_t denote the deterministic flow of the equation (55); $\phi_t(m)$ is the state reached by the flow at time t, starting from $m \in \mathbb{R}^k$ at $t = 0$.

We assume that:

($A4_1$) $\mathbf{E}(\exp c_0 |M_0|^2) < +\infty$ for some suitable $c_0 > 0$ (see Remark 4.6 below).

($A4_2$) The flow ϕ_t is contracting with exponential rate, i.e. there exists a $\lambda > 0$ such that for all $t \geq 0$ and m_1, $m_2 \in \mathbb{R}^k$:

$$|\phi_t(m_1) - \phi_t(m_2)| \leq e^{-\lambda t} |m_1 - m_2|\,,$$

and $\phi_t(0) = 0$ for all $t \geq 0$.

$(A4_3)$ The functions f, b and b_M are Lipschitzian and h and h_M are continuous and bounded. Moreover

$$b_M(0) = 0 \text{ and } h_M(0) = 0\,.$$

It is convenient to work on the canonical probability space of the process (Z, Y). Let $\Omega_1 = \mathcal{C}(\mathbb{R}^+ ; \mathbb{R}^n)$ and $\Omega_2 = \mathcal{C}(\mathbb{R}^+ ; \mathbb{R}^p)$. Define $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F}$ its Borel field, $\mathcal{F}_t$ the canonical Borel filtration on Ω and $(Z_t, Y_t)(\omega_1, \omega_2) = (\omega_1(t), \omega_2(t))$. For a probability measure ν on $\mathbb{R}^n$, let Q^ν be the measure on Ω corresponding to the filtering model (55)–(57) when ν is the law of Z_0. Let P^ν (resp. R^ν) denote the marginal of Q^ν on Ω_1 (resp. on Ω_2). Moreover we use $\mathbf{E}^\nu(.)$ to denote expectation with respect to Q^ν. In particular $P = P^{\pi_0}$ and $\mathbf{E}(.) = \mathbf{E}^{\pi_0}(.)$.

Of course the conditional distribution of Z_t given $\mathcal{Y}_t$ under Q^{π_0} which we shall write as Π_t, is the optimal solution of the filtering problem in model (55)–(57):

$$\Pi_t(\varphi) = \mathbf{E}(\varphi(Z_t) \,|\, \mathcal{Y}_t)\,,\ \forall \varphi \in \mathcal{C}_b(\mathbb{R}^n)\,.$$

The evolution in time of the conditional distribution can be described by the Zakai equation. Introducing

$$\Gamma_t \triangleq \exp\left\{ \int_0^t (h(X_s) + h_M(M_s))' dY_s - \frac{1}{2}\int_0^t |h(X_s) + h_M(M_s)|^2 ds \right\}\,,$$

under our assumptions, it is standard that one may define a probability measure $\check{Q}^\nu$ wich is locally absolutely continuous with respect to the original probability measure Q^ν with local Radon-Nikodym derivative Γ_t. Moreover $\check{Q}^\nu$ and Q^ν are locally equivalent and under $\check{Q}^\nu$ the process Y is a standard Brownian motion independent of X. By the Bayes formula, we have

$$\mathbf{E}^\nu(\varphi(Z_t) \,|\, \mathcal{Y}_t) = \frac{\check{\mathbf{E}}^\nu(\varphi(Z_t)\Gamma_t \,|\, \mathcal{Y}_t)}{\check{\mathbf{E}}^\nu(\Gamma_t \,|\, \mathcal{Y}_t)}\,.$$

Therefore, taking $\nu = \pi_0$, to compute the optimal filter Π_t, it is enough to compute the unnormalised conditional measure $\{\sigma_t, t \geq 0\}$ defined by

$$\sigma_t(\varphi) \triangleq \check{\mathbf{E}}^\nu(\varphi(Z_t)\Gamma_t \,|\, \mathcal{Y}_t)\,,\ \forall \varphi \in \mathcal{C}_b(\mathbb{R}^n)\,.$$

The infinitesimal generator of the diffusion process Z is defined by

$$L\varphi(m,x) = f(m)'\nabla_m\varphi(m,x) + (b(x)+b_M(m))'\nabla_x\varphi(m,x) + \frac{1}{2}\,\mathrm{tr}\,\nabla^2_{xx}\varphi(m,x)\,,$$

for all $m \in \mathbb{R}^k$, $x \in \mathbb{R}^d$ and $\varphi \in \mathcal{C}^2_b(\mathbb{R}^n)$. Then $\sigma_t(\varphi)$ satisfies the Zakai equation (cf. [13] or e.g. [11])

$$\sigma_t(\varphi) = \sigma_0(\varphi) + \int_0^t \sigma_s(L\varphi)\,ds \; + \int_0^t \sigma_s(H\varphi)'\,dY_s\,, \tag{58}$$

where $H(m,x) = h_M(m) + h(x)$, for all $m \in \mathbb{R}^k$ and $x \in \mathbb{R}^d$. The spatial variable dimension of this equation is n, the dimension of Z and the initial condition is π_0.

In this section, in the case of signal ergodicity we propose a reduced–order filter which approaches the optimal filter as the time goes to infinity. The idea is to do as if the process M was equal to 0 in the system (55)–(57). In fact, $\bar{\pi}_0 = \delta_0 \otimes \eta$ where δ_0 is the Dirac measure at point $0 \in \mathbb{R}^k$ and η is the marginal of π_0 on $\mathbb{R}^d$ is taken as the initial condition of the Zakai equation (58) instead of π_0.

Clearly the conditionial law of Z_t given $\mathcal{Y}_t$ with initial law $\bar{\pi}_0$ is

$$\bar{\Pi}_t = \delta_0 \otimes \bar{\Pi}^X_t\,,$$

where $\bar{\Pi}^X_t$ is the conditional law of X_t given $\mathcal{Y}_t$ in the model

$$dX_t = b(X_t)\,dt + dV_t\,, \tag{59}$$

$$dY_t = h(X_t)\,dt + dW_t\,, \quad Y_0 = 0\,, \tag{60}$$

when the distribution of X_0 is η. Same as above, to know $\bar{\Pi}^X_t$ it is enough to compute the unnormalised conditional measure $\{\bar{\sigma}^X_t\,,\, t \geq 0\}$ which satisfies the following Zakai eqation for all $\varphi \in \mathcal{C}^2_b(\mathbb{R}^d)$:

$$\bar{\sigma}^X_t(\varphi) = \bar{\sigma}^X_0(\varphi) + \int_0^t \bar{\sigma}^X_s(\bar{L}^X\varphi)\,ds \; + \int_0^t \bar{\sigma}^X_s(h\varphi)'\,dY_s\,, \tag{61}$$

where $\bar{L}^X$ is defined by

$$\bar{L}^X\varphi(x) = b(x)'\nabla_x\varphi(x) + \frac{1}{2}\,\mathrm{tr}\,\nabla^2_{xx}\varphi(x),\ \forall x \in \mathbb{R}^d\,.$$

The spatial variable dimension of this equation is d, the dimension of X and the initial condition is η. The conditional law $\bar{\Pi}_t$ will be referred to as the reduced–order filter. The advantage of this filter is that the numerical computations of the Zakai equation are easier.

Since we have

$$\bar{\Pi}_t(\varphi) = \mathbf{E}^{\bar{\pi}_0}\left(\varphi(0, X_t) \mid \mathcal{Y}_t\right) = \mathbf{E}^{\bar{\pi}_0}\left(\varphi(Z_t) \mid \mathcal{Y}_t\right)\ , \forall \varphi \in \mathcal{C}_b(\mathbb{R}^n)\,,$$

the study of the asymptotic deviation of the reduced-order filter from the optimal filter yields to the study of the asymptotic stability of the optimal filter with respect to its initial condition as studied by Ocone and Pardoux in [10]. Indeed under Q^{π_0}, the difference $\Pi_t(\varphi) - \bar{\Pi}_t(\varphi)$ is the difference between the conditional law with "true" initial condition π_0 (the optimal filter) and the conditional law with erroneous initial condition $\bar{\pi}_0$ (the reduced–order filter).

Under some ergodicity assumptions we shall show that the assumptions of Theorem 4.1 above are verified and then this Theorem will be applied.

4.3 Asymptotic behaviour of the reduced–order filter

Let $(S_t^{\bar{X}})_{t \geq 0}$ denote the semigroup on $\mathcal{C}_b(\mathbb{R}^d)$ associated with the Markov process $\bar{X}$ defined by the stochastic differential equation

$$d\bar{X}_t = b(\bar{X}_t)dt + dB_t\,, \forall t \geq 0\,, \tag{62}$$

where B is a Brownian motion.

The following sufficient condition for such a process to be ergodic is due to Hasminski [3]:

$(A4_4)$ there exist two constants $r > 0$ and $\alpha > 0$ such that

$$x'b(x) \leq -\alpha|x|\ ,\ \ \forall |x| > r\,.$$

Under this condition, $(S_t^{\bar{X}})_{t \geq 0}$ admits a unique invariant measure $\mu^{\bar{X}}$ and for any probability measure ν on $\mathbb{R}^d$,

$$\nu S_t^{\bar{X}} \ \to\ \mu^{\bar{X}} \text{ weakly as } t \to +\infty\,, \tag{63}$$

where $\nu S_t^{\bar{X}}$ denotes the law of the random variable $\bar{X}_t$ when the distribution of $\bar{X}_0$ is ν.

Then we can prove the following statement

Proposition 4.3 *Assume that in the model (55)–(57), the conditions* $(A4_1)$ – $(A4_4)$ *hold. Then for every bounded continuous* $\varphi : \mathbb{R}^n \to \mathbb{R}$*,*

$$\lim_{t \to +\infty} \mathbf{E}\left[\left(\Pi_t(\varphi) - \bar{\Pi}_t(\varphi)\right)^2\right] = 0\,. \tag{64}$$

The proof consists to show that the conditions $(i)-(iii)$ of Theorem 4.1 are all satisfied. This is done through two lemmas. The first one says that condition (iii) is fulfilled.

Lemma 4.4 *Assume that conditions* $(A4_1) - (A4_3)$ *hold. Then* R^{π_0} *is absolutely continuous with respect to* $R^{\bar{\pi}_0}$ *.*

Proof First we introduce a new probability measure $\bar{Q}^{\pi_0}$ with a marginal $\bar{R}^{\pi_0}$ on Ω_2 which coincides with the marginal $R^{\bar{\pi}_0}$ of $Q^{\bar{\pi}_0}$ on Ω_2.

Define the processes $\bar{V}$ and $\bar{W}$ as follows:

$$\begin{aligned} \bar{V}_t &= V_t + \int_0^t b_M(M_s)\, ds\,, \\ \bar{W}_t &= W_t + \int_0^t h_M(M_s)\, ds\,, \end{aligned}$$

Let

$$K_t = \int_0^t b_M(M_s)'\, dV_s + \int_0^t h_M(M_s)'\, dW_s\,,$$

$< K >_t = \int_0^t (|b_M(M_s)|^2 + |h_M(M_s)|^2)\; ds$ and $L_t = \exp\left\{-K_t - \frac{1}{2} < K >_t\right\}$.

Under our assumptions, $\mathbf{E}(L_t) = 1$ and then $\{L_t\,,\, t \geq 0\}$ is a $(\mathcal{F}_t)$–martingale under Q^{π_0}. Then we may define the probability measure $\bar{Q}^{\pi_0}$ which is locally absolutely continuous with respect to Q^{π_0} with the local Radon–Nikodym derivative L_t. Moreover under $\bar{Q}^{\pi_0}$ we have

$$\begin{aligned} dX_t &= b(X_t)\, dt + d\bar{V}_t\,, \\ dY_t &= h(X_t)\, dt + \bar{W}_t\;,\; Y_0 = 0\,, \end{aligned}$$

where the processes $\bar{V}$ and $\bar{W}$ are standard Brownian motions taking values in $\mathbb{R}^d$ and $\mathbb{R}^p$ respectively and X_0 is independent of $(\bar{V}, \bar{W})$ and has law η the marginal of π_0 on $\mathbb{R}^d$. Since also under $Q^{\bar{\pi}_0}$ the same facts hold with (V, W) in place of $(\bar{V}, \bar{W})$, we can assert that $\bar{R}^{\pi_0} = R^{\bar{\pi}_0}$.

Now we shall prove that in fact $\bar{Q}^{\pi_0} << Q^{\pi_0}$ on $(\Omega, \mathcal{F})$. Thanks to Proposition III.3.5 in [4], it is enough to show that the density process L_t is a square–integrable martingale i.e.

$$\sup_t \mathbf{E}(L_t^2) < +\infty \, .$$

Recall that M_0 is independent of (V, W). Hence M is also independent of (V, W) and it is easy to check that

$$\mathbf{E}(L_t^2) = \mathbf{E}(\exp < K >_t) \, .$$

Now consider that in assumption $(A4_1)$ a suitable c_0 means that $c_0 \geq \sup(c_{b_M}, c_{h_M})^2/\lambda$ where λ is the convergence rate of the flow $\phi(t)$ appearing in $(A4_2)$ and c_{b_M} and c_{h_M} are the Lipschitz constants of b_M and h_M in $(A4_3)$ respectively. Then it is easy to verify that there is a constant $C > 0$ independent of t such that

$$\mathbf{E}(L_t^2) \leq C \, , \quad \forall t \geq 0 \, . \tag{65}$$

Therefore we can conclude that $\bar{Q}^{\pi_0} << Q^{\pi_0}$ on $(\Omega, \mathcal{F})$ and consequently $R^{\bar{\pi}_0} = \bar{R}^{\pi_0} << R^{\pi_0}$, which achieves the proof of the lemma. □

Finally the next lemma says that conditions $(i) - (ii)$ in Theorem 4.1 are satisfied.

Lemma 4.5 *Assume that conditions $(A4_1) - (A4_4)$ hold. Let $(S_t)_{t\geq 0}$ be the semigroup on $\mathcal{C}_b(\mathbb{R}^n)$ associated with the Markov process $Z = (M', X')'$ generated by equations (55)–(57). Then $(S_t)_{t\geq 0}$ is a strongly continous positive and conservative contraction semigroup which admits a unique invariant measure μ defined by $\mu = \delta_0 \otimes \mu^{\bar{X}}$ which satisfies (H_1). Moreover conditions $(H_2(\pi_0))$ and $(H_2(\bar{\pi}_0))$ are fulfilled.*

Proof Under our assumptions, $(S_t)_{t\geq 0}$ is a strongly continuous positive and conservative contraction semigroup on $\mathcal{C}_b(\mathbb{R}^n)$. Moreover, under the above ergodicity assumption $(A4_4)$, it is straightforward to show that the measure $\mu = \delta_0 \otimes \mu^{\bar{X}}$, where $\mu^{\bar{X}}$ is the unique invariant measure of $(S_t^{\bar{X}})_{t\geq 0}$, is an invariant measure of $(S_t)_{t\geq 0}$. The uniqueness of $\mu^{\bar{X}}$ implies the uniqueness of μ and from (63), (H_1) can be easily obtained. On the other hand, since the law of Z_t under $P^{\bar{\pi}_0}$ coincides with the law $\delta_0 \otimes (\eta S_t^{\bar{X}})$ on Ω_1 where $\bar{X}$ is defined by (62), condition $(H_2(\bar{\pi}_0))$ is directly obtained from (63). It remains to verify that $(H_2(\pi_0))$ is fulfilled. Since $(S_t)_{t\geq 0}$ admits a unique invariant measure it suffices to show that the family of laws $\{P_{Z_t}\,,\, t \geq 0\}$ where $P_{Z_t} = \pi_0 S_t$, is uniformly tight i.e. for any $\varepsilon > 0$, there is a compact set $K \subset \mathbb{R}^n$ such that

$$P(Z_t \in K) \geq 1 - \varepsilon\,,\ \forall t \geq 0\,.$$

Since from $(A4_2)$, M converges to 0 in probability, it suffices to prove that the family $\{P_{X_t}, t \geq 0\}$ is uniformly tight.

For any $R > 0$, we have

$$P(|X_t| > R) = \bar{\mathbf{E}}\left(\mathbf{1}_{|X_t|>R)} L_t^{-1}\right)\,,$$

where $\bar{\mathbf{E}}$ denotes expectation with respect to the probability measure $\bar{Q}^{\pi_0}$ defined in the proof of the Lemma 4.4.

Recall that from the above proof, under $\bar{Q}^{\pi_0}$ we have

$$dX_t = b(X_t)\,dt + d\bar{V}_t\,,$$

where $\bar{V}$ is a Brownian motion and X_0 has distribution η and is independent of $\bar{V}$. Therefore under $\bar{Q}^{\pi_0}$ the distribution of the random variable X_t coincides with $\eta S_t^{\bar{X}}$ where $\bar{X}$ is defined by (62). Hence by the Cauchy-Schwarz inequality for any $R > 0$ we get

$$\begin{aligned} P(|X_t| > R) &\leq \bar{Q}^{\pi_0}(|X_t| > R)^{\frac{1}{2}}(\bar{\mathbf{E}}(L_t^{-2}))^{\frac{1}{2}} \\ &= (\eta S_t^{\bar{X}}(\{x \in \mathbb{R}^d : |x| > R\})^{\frac{1}{2}}(\mathbf{E}(L_t^{-1}))^{\frac{1}{2}}\,. \end{aligned}$$

As in the proof of Lemma 4.4, we can easily obtain

$$\mathbf{E}(L_t^{-1}) = \mathbf{E}(\exp < K >_t) \leq C\,,\ \forall t \geq 0\,,$$

where the constant C is independent of t and is already obtained in (65).

Moreover, assumption $(A4_4)$ implies that the family of laws $\{\eta S_t^{\bar{X}}, t \geq 0\}$ is uniformly tight. Therefore for a given $\varepsilon > 0$, there is a $R > 0$ such that

$$\eta S_t^{\bar{X}}(\{x \in \mathbb{R}^d : |x| > R\}) \leq \frac{\varepsilon^2}{C},\ \forall t \geq 0\,.$$

Hence for this R it follows that

$$P(|X_t| > R) \leq \varepsilon\,.$$

This means that the family of laws $\{P_{X_t}, t \geq 0\}$ is uniformly tight and we can conclude that $(H_2(\pi_0))$ is fulfilled.

□

Remark 4.6 From the proof of Lemma 4.4 it appears that assumption $(A4_1)$ can be precisely reformulated as:

"$\mathbf{E}(\exp(c_0|M_0|^2) < +\infty$ for $c_0 = \sup(c_{b_M}, c_{h_M})^2/\lambda$ where λ is the convergence rate of the flow $\phi(t)$ appearing in $(A4_2)$ and c_{b_M} and c_{h_M} are the Lipschitz constants of b_M and h_M in $(A4_3)$ respectively".
It is worth to emphasize that the larger is λ, the smaller is the constant c_0.

Remark 4.7 With slight modifications in the proof of Proposition 4.3, one can show that the same asymptotic result holds for any initial law of the form $\delta_0 \otimes \beta$ in place of $\bar{\pi}_0$ provided that $\beta << \eta$ (we have $R^{\delta_0\otimes\beta} << R^{\pi_0}$). So for two different initial conditions of this form the asymptotic behaviour of the reduced-order filter is unchanged. In the particular case of the filtering model (59)–(60), this reduces to the fact that, under the assumptions "b Lipschitzian", "h continuous bounded", and $(A4_4)$, a filter initialized with an erroneous initial condition $\beta << \eta$ has the same asymptotic behaviour as the optimal filter (initialized with η). Actually that property in model (59)–(60) is an immediate consequence of Theorem 3.2 and Remark 3.3 in [10].

Acknowledgements: We are very grateful to Marina Kleptsyna for valuable discussions concerning ergodic properties of diffusion processes.

References

[1] V. E. Beněs and I. Karatzas : *Estimation and control for linear, partially observable systems with non–Gaussian initial distribution*, Stochastic Process. Appl., 14, 1983, pp. 233–248.

[2] W. A. Coppel : *Stability and asymptotic behavior of differential equations*, Boston, MA: Heath, 1965.

[3] R. Z. Hasminski : *Stochastic stability of differential equations*, Sijthoff & Noordhoff, Alphen aan den Rijn, The Netherlands and Rockville, Maryland ,USA, 1980.

[4] J. Jacod and A. N. Shiryaev : *Limit theorems for stochastic processes*, Springer–Verlag, 1987.

[5] N. Kamaroff : *Upper bounds for the eigenvalues of the solution of the Lyapunov matrix equation*, IEEE Trans. Automat. Contr., Vol. 35, No. 6, June 1990, pp. 737–739.

[6] T. Kang, B. S. Kim and J. G. Lee : *Spectral norm and trace bounds of algebraic matrix Riccati equations*, IEEE Trans. on Automat. Contr., Vol. 41, No. 12, 1996, pp. 1828–1830.

[7] H. Kwakernaak and R. Sivan : *Linear Optimal Control Systems*, Wiley-Interscience, New-York, 1972.

[8] A. M. Makowski : *Filtering formulae for partially observed linear systems with non–Gaussian initial conditions*, Stochastics, 16, 1986, pp. 1–24.

[9] T. Mori : *Comments on "A matrix inequality associated with bounds on solutions of algebraic Riccati and Lyapunov equation"*, IEEE Trans. Automat. Contr., Vol. AC-29, 1988, pp. 1088.

[10] D. Ocone and E. Pardoux : *Asymptotic stability of the optimal filter with respect to its initial condition*, SIAM J. Control and Optimization, Vol. 34, No. 1, 1996, pp. 226–243,

[11] E. Pardoux : *Filtrage non linéaire et équations aux dérivées partielles stochastiques associées*, In P.L. Hennequin, editor, Ecole d'été de Probabilités de Saint–Flour XIX – 1989, Vol. 1464 of Lecture Notes in Mathematics, pp. 67–163, Springer Verlag, Berlin, 1991.

[12] D.-T. Pham, J. Verron, M. C. Roubaud : *A Singular Evolutive Extended Kalman Filter for Data Assimilation in Oceanography* to appear in Journal of Marine Systems.

[13] M. Zakai : *On the optimal filtering of diffusions processes*, Zeit. Wahr. Verv. Geb., 11, 1969, pp. 230–243.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - B.P. 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Lorraine : Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - B.P. 101 - 54602 Villers lès Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot St Martin (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, B.P. 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399