On the stability and the exponential concentration of Extended Kalman-Bucy filters

Pierre del Moral, Aline Kurtzmann, Julian Tugaut

On the Stability and the Exponential Concentration of Extended Kalman-Bucy filters

P. Del Moral, A. Kurtzmann, J. Tugaut

June 27, 2016

Abstract

The exponential stability and the concentration properties of a class of extended Kalman-Bucy filters are analyzed. New estimation concentration inequalities around partially observed signals are derived in terms of the stability properties of the filters. These non asymptotic exponential inequalities allow to design confidence interval type estimates in terms of the filter forgetting properties with respect to erroneous initial conditions. For uniformly stable signals, we also provide explicit non-asymptotic estimates for the exponential forgetting rate of the filters and the associated stochastic Riccati equations w.r.t. Frobenius norms. These non asymptotic exponential concentration and quantitative stability estimates seem to be the first results of this type for this class of nonlinear filters. Our techniques combine χ-square concentration inequalities and Laplace estimates with spectral and random matrices theory, and the non asymptotic stability theory of quadratic type stochastic processes.

Keywords : Concentration inequalities, non asymptotic exponential stability, Lyapunov exponents, extended Kalman-Bucy filter, Riccati equation.

Mathematics Subject Classification : 93C55, 93D20, 93E11, 60M20, 60G25.

1 Introduction

The linear-Gaussian stochastic filtering problem has been solved in the beginning of the 1960s by Kalman and Bucy in their seminal articles [4, 5, 9]. Since this period, Kalman-Bucy filters have became one of the most powerful estimation algorithm in applied probability, statistical inference, information theory and engineering sciences. The Kalman-Bucy filter is designed to estimate in an optimal way (minimum variance) the internal states of linear-Gaussian time series from a sequence of partial and noisy measurements. The range of applications goes from tracking, navigation and control to computer vision, econometrics, statistics, finance, and many others. For linear-Gaussian filtering problems, the conditional distribution of the internal states of the signal given the observations up to a give time horizon are Gaussian. The Kalman-Bucy filters and the associated Riccati equation coincide with the evolution of the conditional averages and the conditional covariances error matrices of these conditions Gaussian distributions.

Using natural local linearization techniques Kalman-Bucy filters are also currently used to solve nonlinear and/or non Gaussian signal observation filtering problems. The resulting

Extended Kalman-Bucy filter (*abbreviated EKF*) often yields powerful and computational efficient estimators. Nevertheless it is well known that it fails to be optimal with respect to the minimum variance criteria. For a more thorough discussion on the origins and the applications of these observer type filtering techniques we refer to the articles [13, 19, 20] and the book by D. Simon [18].

There is a vast literature on the applications and the performance of extended Kalman filter, most on discrete time filtering problems, but very few on the stability properties, none on the exponential concentration properties.

In the last two decades, the convergence properties of the EKF have been mainly developed into three different but somehow related directions:

The first commonly used approach is to analyze the long time behavior of the estimation error between the filter and the partially observed signal. To bypass the fluctuations induced by the signal noise and the observation perturbations, one natural strategy is to design judicious deterministic observers as the asymptotic limit of the EKF when the observation and the sensor noise tends to zero. As underlined in [3], in deterministic setting the original covariance matrices of the stochastic signal and the one of the observation perturbations are interpreted as design/tuning type parameters associated with the confidence type matrices of the trusted model and the confidence matrix of associated with the measurements.

For a more detailed discussion on deterministic type observers as the limit of filters when the sensor and the observation noise tends to zero we refer the reader to the seminal article [1] and the more recent study [3]. Several articles proposed a series of observability and controllability conditions under which the estimation error of the corresponding discrete time observer converges to zero [1, 2, 19, 20]. These regularity conditions allows to control the maximal and the minimal eigenvalues of the solution of the Riccati equations (and its inverse).

One of the drawbacks of this approach is that it gives no precise information on the stochastic EKF but on the limiting noise free-type deterministic observer. On the other hand, up to our knowledge there doesn't exist any uniform result that allow to quantify the difference between the filter and its asymptotic limit with respect the time parameter. Another drawback is that the initial estimation errors need to be rather small and the signal model close to linear.

In general practical and stochastic situations, mean square errors doesn't converge to zero as the time parameter tends to ∞. The reasons are two folds: Firstly, the observation noise of the sensors cannot be totally cancelled. On the other hand the internal signal states are usually only partially observed, and some components may not be fully observable.

A second closely related strategy is to design a Lyapunov function to ensure the stochastic stability of the EKF. Here again these Lyapunov functions are expressed in terms of the inverse of the Riccati equation. These stability properties ensure that the mean square estimation error is uniformly bounded w.r.t. the time horizon [1, 10, 14, 15]. The regularity conditions are also based on a series of local observability and controllability conditions. As any variance type estimate, these mean square error control are somehow difficult to use in practical situations with rather crude confidence interval estimates.

The third and more recent approach is based on the contraction theory developed by W. Lohmiller and J.J.E. Slotine in the seminal articles [11, 12], and further developed in [3]. This approach is also designed to study deterministic type observers. The idea is to control the estimation error between a couple of close EKF trajectories in a given region w.r.t. the metric induced by the quadratic form associated with the inverse of the solution of the Riccati equation. This approach considers the partially observed signal as a deterministic system and requires the filter to start in a basin of attraction of the true state. In summary, these techniques show that the observer induced by the EKF converges locally exponentially to the state of the signal when the quadratic form induced by the inverse of the Riccati equation is sufficiently regular and under appropriate observability and controllability conditions.

The objective of this article is to complement these three approaches with a novel stochastic analysis based on exponential concentration inequalities and uniform χ-square type estimates for stochastic quadratic type processes.

Our regularity conditions are somehow stronger than the ones discussed in the above referenced articles but they don't rely on some observability and controllability conditions, nor on suitable local initial conditions nearby the true signal state. Last but not least our methodology applies to stochastic filtering problems, not to deterministic type observers.

In our framework the signal process is assumed to be uniformly and exponentially stable, and the sensor function is linear. In this apparently simple nonlinear filtering problem the quantitative analysis of the EKF exponential stability is based on sophisticated probabilistic tools. The complexity of these stochastic processes can be measured by the fact that the EKF is a nonlinear diffusion process equipped with a diffusion correlation matrix satisfying a coupled nonlinear and stochastic Riccati equation. Filtering problems with linear signals and nonlinear sensor functions are somehow simpler to analyze since the stability of the signal is directly transfer to the one of the Riccati equation.

This study has been motivated by one of our recent research project on the refined convergence analysis Ensemble type Kalman-Bucy filters. To derive some useful uniform convergence results with respect to the time horizon we shown in [6] that the signal process needs to be uniformly stable. This rather strong condition cannot be relaxed even for linear Gaussian filtering models. We plan to extend these results for nonlinear filtering models based on the non asymptotic estimates presented in this article.

In this context we present new exponential concentration inequalities to quantify the stochastic stability of the EKF. They allow to derive confidence intervals for the deviations of the stochastic flow of the EKF around the internal states of the partially observed signal. These estimates also show that the fluctuations induced by any erroneous initial condition tends to zero as the time horizon tends to $+\infty$.

Our second objective is to develop a non asymptotic quantitative analysis of the stability properties of the EFK. In contrast to the linear-Gaussian case discussed in [6], the Riccati equation associated with the EFK depends on the states of the filter. The resulting system is a nonlinear stochastic process evolving in multidimensional inner product spaces. To analyze these complex models we develop a stability theory of quadratic type stochastic processes. Our main contribution is a non asymptotic $\mathbb{L}_p$-exponential stability theorem. This theorem shows that the $\mathbb{L}_p$-distance between two solutions of the EKF and the stochastic Riccati equation with possibly different initial conditions converge to zero as the time horizon tends to ∞. We also provide a non asymptotic estimate of the exponential decay rate.

The rest of the article is organized as follows:

In the next two sections, section 1.1 and section 1.2, we present the nonlinear filtering models discussed in the article and we state the main results developed in this work. Section 2 is concerned with the stability properties of quadratic type processes. This section present the main technical results used in the further development of the article. Most of the technical proofs are provided in the appendix. Section 3 is dedicated to the stochastic stability properties of the signal and the EKF. The end of the article is mainly concerned with the proofs of the two main theorems presented in section 1.2.

1.1 Description of the models

Consider a time homogeneous nonlinear filtering problem of the following form

$$\begin{cases} dX_t &= A(X_t)\,dt \;+\; R_1^{1/2}\,dW_t \\ dY_t &= BX_t\,dt \;+\; R_2^{1/2}\,dV_t \end{cases} \qquad \text{and we set } \mathcal{F}_t = \sigma\left(Y_s,\ s \leqslant t\right).$$

In the above display, (W_t, V_t) is an $(r_1 + r_2)$-dimensional Brownian motion, X_0 is a r_1-valued Gaussian random vector with mean and covariance matrix $(\mathbb{E}(X_0), P_0)$ (independent of (W_t, V_t)), the symmetric matrices $R_1^{1/2}$ and $R_2^{1/2}$ are invertible, B is an $(r_2 \times r_1)$-matrix, and $Y_0 = 0$. The drift of the signal is differentiable vector valued function $A\ :\ x \in \mathbb{R}^{r_1} \mapsto A(x) \in \mathbb{R}^{r_1}$ with a Jacobian denoted by $\partial A\ :\ x \in \mathbb{R}^{r_1} \mapsto A(x) \in \mathbb{R}^{(r_1 \times r_1)}$. In the further development of the article we assume that the Jacobian matrix of A satisfies the following regularity conditions:

$$\begin{cases} -\lambda_{\partial A} &:= \sup_{x\in\mathbb{R}^{r_1}} \rho(\partial A(x) + \partial A(x)') < 0 \\[2ex] \|\partial A(x) - \partial A(y)\| &\leqslant \kappa_{\partial A}\ \|x - y\| \quad \text{for some } \kappa_{\partial A} < \infty. \end{cases} \tag{1}$$

where $\rho(P) := \lambda_{max}(P)$ stands for the maximal eigenvalue of a symmetric matrix P. In the above display $\|\partial A(x) - \partial A(y)\|$ stands for the $\mathbb{L}_2$-norm of the matrix operator $(\partial A(x) - \partial A(y))$, and $\|x - y\|$ the Euclidian distance between x and y.

A Taylor first order expansion shows that

$$(1) \Longrightarrow \langle x - y, A(x) - A(y)\rangle \leqslant -\lambda_A\ \|x - y\|^2 \quad \text{with } \lambda_A \geqslant \lambda_{\partial A}/2 > 0. \tag{2}$$

The Extended Kalman-Bucy filter is defined by the evolution equations

$$\begin{aligned} d\widehat{X}_t &= A(\widehat{X}_t)dt + P_t B' R_2^{-1} \left[dY_t - B\widehat{X}_t dt\right] \quad \text{with} \quad \widehat{X}_0 = \mathbb{E}(X_0) \\ \partial_t P_t &= \partial A(\widehat{X}_t) P_t + P_t \partial A(\widehat{X}_t)' + R_1 - P_t\ S P_t \quad \text{with} \quad S := B' R_2^{-1} B \end{aligned} \tag{3}$$

where B' stands for the transpose of the matrix B. For nonlinear signal processes the random matrices P_t cannot be interpreted as the error covariance matrices. Nevertheless, rewriting the EKF in terms of the signal process we have

$$d(X_t - \widehat{X}_t) = [(A(X_t) - A(\widehat{X}_t)) - P_t S(X_t - \widehat{X}_t)]\ dt + R_1^{-1/2} dW_t - P_t B' R_2^{-1/2} dV_t$$

Replacing $(A(X_t) - A(\widehat{X}_t))$ by the first order approximation $\partial A(\widehat{X}_t)(X_t - \widehat{X}_t)$ we define a process

$$d\widetilde{X}_t := [\partial A(\widehat{X}_t) - P_t S]\widetilde{X}_t\ dt + R_1^{-1/2} dW_t - P_t B' R_2^{-1/2} dV_t$$

It is a simple exercise to check that the solution of the Riccati equation (3) coincides with the $\mathcal{F}_t$-conditional covariance matrices of $\widetilde{X}_t$; that is, for any $t \geqslant 0$ we have $P_t = \mathbb{E}\left(\widetilde{X}_t \widetilde{X}_t' \mid \mathcal{F}_t\right)$.

1.2 Statement of the main results

We let $\phi_t(x) = X_t$ and $\varphi_t(x) := x_t$ be the stochastic and the deterministic flows of the stochastic and the deterministic systems

$$\left\{\begin{array}{rcl} dX_t & = & A(X_t)\,dt + R_1^{1/2} dW_t \\ \partial_t x_t & = & A(x_t) \end{array}\right. \qquad \text{starting at } x_0 = \varphi_0(x) = x = X_0 = \phi_0(x).$$

We also let $\overline{\Phi}_t = (\Phi_t, \Psi_t)$ be the stochastic flow associated with the EKF and the Riccati stochastic differential equations; that is

$$\overline{\Phi}_t(\widehat{X}_0, P_0) = \left(\Phi_t(\widehat{X}_0, P_0), \Psi_t(\widehat{X}_0, P_0)\right) := \left(\widehat{X}_t, P_t\right)$$

Given $(r_1 \times r_2)$ matrices P, Q we define the Frobenius inner product

$$\langle P, Q\rangle = \mathrm{tr}(P'Q) \quad \text{and the associated norm} \quad \|P\|_F^2 = \mathrm{tr}(P'P)$$

where $\mathrm{tr}(C)$ stands for the trace of a given matrix. We also equip the product space $\mathbb{R}^{r_1} \times \mathbb{R}^{r_1 \times r_1}$ with the inner product

$$\langle (x_1, P_1), (x_2, P_2)\rangle := \langle x_1, x_2\rangle + \langle P_1, P_2\rangle \quad \text{and the norm} \quad \|(x, P)\|^2 := \langle (x, P), (x, P)\rangle$$

We recall the χ-square Laplace estimate

$$\mathbb{E}\left(\exp\left[\|X_0 - \widehat{X}_0\|^2/\chi(P_0)\right]\right) \leqslant e \qquad \text{with} \quad \chi(P_0) := 4 r_1 \rho(P_0) \tag{4}$$

The proof of (4) and more refined estimates are housed in the appendix. We have the rather crude almost sure estimate

$$\begin{array}{rcl} \partial_t \mathrm{tr}\,(P_t) & = & \mathrm{tr}\left((\partial A(\widehat{X}_t) + \partial A(\widehat{X}_t)')P_t\right) + \mathrm{tr}(R_1) - \mathrm{tr}(S P_t^2) \\ & \leqslant & -\lambda_{\partial A}\,\mathrm{tr}\,(P_t) + \mathrm{tr}(R_1) \end{array}$$

This readily yields the upper bound

$$\mathrm{tr}(P_t) \leqslant \tau_t(P) := e^{-\lambda_{\partial A} t}\,\mathrm{tr}\,(P_0) + \mathrm{tr}(R_1)/\lambda_{\partial A} \quad \Rightarrow \quad \sup_{t \geqslant 0} \mathrm{tr}(P_t) \leqslant \mathrm{tr}\,(P_0) + \mathrm{tr}(R_1)/\lambda_{\partial A} \tag{5}$$

Most of the analysis developed in the article relies on the following quantities:

$$\sigma_{\partial A}^2 := 1 + 2\,\pi_{\partial A} \quad \text{with} \quad \pi_{\partial A}(t) := \tau_t^2(P)\,\rho(S)\,\mathrm{tr}(R_1)^{-1} \longrightarrow_{t \to \infty} \pi_{\partial A} := \frac{\rho(S)}{\lambda_{\partial A}}\,\frac{\mathrm{tr}(R_1)}{\lambda_{\partial A}} \tag{6}$$

The quantity $\rho(S)$ is connected to the sensor matrix B and to the inverse of the covariance matrix of the observation perturbations. We also have the rather crude estimate

$$\rho(S) \leqslant \mathrm{tr}(S) = \|\partial R_2^{-1/2} B\|_F^2 \leqslant \|R_2^{-1/2}\|_F^2\,\|B\|_F^2 \leqslant \mathrm{tr}(R_2^{-1})\,\|B\|_F^2$$

Our first main result concerns the stochastic stability of the EKF and it is described in terms of the function

$$\delta \in [0, \infty[\mapsto \varpi(\delta) := \frac{e^2}{\sqrt{2}}\left[\frac{1}{2} + \left(\delta + \sqrt{\delta}\right)\right]$$

More precisely we have the following exponential concentration theorem.

Theorem 1.1. *For any initial states* $(x, \widehat{x}, p) \in \mathbb{R}^{r_1+r_1+(r_1\times r_1)}$ *and any time horizon* $t \in [0,\infty[$, *and any* $\delta \geqslant 0$ *the probabilities of the following events are greater than* $1 - e^{-\delta}$:

$$\|\phi_t(x) - \varphi_t(x)\|^2 \leqslant \varpi(\delta) \frac{\mathrm{tr}(R_1)}{\lambda_A} \tag{7}$$

$$\begin{aligned}\|\phi_t(x) - \Phi_t(\widehat{x}, p)\|^2 \leqslant\ & 4\ \varpi(\delta) \frac{\mathrm{tr}(R_1)}{\lambda_A}\ \sigma^2_{\partial A} \\ & +2e^{-\lambda_{\partial A}t}\ \|x - \widehat{x}\|^2 + 8\varpi(\delta) \frac{|e^{-\lambda_A t} - e^{-\lambda_{\partial A} t}|}{|\lambda_A - \lambda_{\partial A}|}\ \rho(S)\ \mathrm{tr}(p)^2\end{aligned} \tag{8}$$

The proofs of the concentration inequalities (7) and (8) are provided respectively in section 3.1 and section 3.2. See also theorem 3.1 and theorem 3.2 for related Laplace χ-square estimates of time average distances.

The role of each quantity in (7) and (8) is clear. The size of the events the "confidence" are proportional to the signal or the observation perturbations, and inversely proportional to the stability rate of the systems. More interestingly, formula (8) shows that the impact of the initial conditions is exponentially small when the time horizon increases.

Our next objective is to better understand the stability properties of the EKF and the corresponding stochastic Riccati equation. To this end, it is convenient to strengthen our regularity conditions. We further assume that

$$\lambda_{\partial A} > \sqrt{2\kappa_{\partial A}\mathrm{tr}(R_1)} \vee (4\rho(S)) \tag{9}$$

and for some $\alpha > 1$

$$4e\alpha \sqrt{\frac{\rho(S)}{\lambda_{\partial A}}\ \frac{\mathrm{tr}(R_1)}{\lambda_A}}\ \left[1 + 2\ \frac{\mathrm{tr}(R_1)}{\lambda_{\partial A}} \frac{\rho(S)}{\lambda_{\partial A}}\right] < 1 \tag{10}$$

In contrast with the linear-Gaussian case, the Riccati equation (3) depends on the internal states of the EKF. As a result its stability properties are characterized by a stochastic Lyapunov exponent that depends on the random trajectories of the filter as well as on the signal-observation processes. Condition (10) is a technical condition that allows to control uniformly the fluctuations of these stochastic exponents with respect to the time horizon. By (9) this condition is met as soon as

$$\alpha e\ \mathrm{tr}(R_1)\ \left[1 + \frac{1}{8}\ \frac{\mathrm{tr}(R_1)}{\rho(S)}\right] < \lambda_A/2$$

Loosely speaking, when the signal is not sufficiently stable the erroneous initial conditions of EKF may be too sensitive to small perturbations of the sensor. When the exponential decay to equilibrium of the signal is stronger than these spectral instabilities the EKF and the corresponding stochastic Riccati equations are stable and forgets any erroneous initial conditions.

We set

$$\Delta_t := \|\overline{\Phi}_t\left(\widehat{X}_0, P_0\right) - \overline{\Phi}_t\left(\check{X}_0, \check{P}_0\right)\|^2$$

and

$$\begin{aligned}\Lambda/\lambda_{\partial A} &:= 1-2\,\frac{\kappa_{\partial A}}{\lambda_{\partial A}}\,\frac{\operatorname{tr}(R_1)}{\lambda_{\partial A}}-\sqrt{\frac{\rho(S)}{\lambda_{\partial A}}}\left[1-\frac{3}{4}\sqrt{\frac{\rho(S)}{\lambda_{\partial A}}}\right]\\ &\geqslant \frac{1}{2}-2\,\frac{\kappa_{\partial A}}{\lambda_{\partial A}}\,\frac{\operatorname{tr}(R_1)}{\lambda_{\partial A}}>0\end{aligned}$$

We are now in position to state our second main result.

Theorem 1.2. *When $\lambda_{\partial A}>0$ we have the uniform estimates*

$$\forall n\geqslant 1\qquad \sup_{t\geqslant 0}\mathbb{E}(\Delta_t^n)<\infty$$

Assume conditions (9) and (10) are satisfied for some $\alpha>1$. In this situation, for any $\epsilon\in]0,1]$ there exists some time horizon s such that for any $t\geqslant s$ we have the almost sure contraction estimate

$$\delta:=\frac{1}{2}\sqrt{\frac{\lambda_{\partial A}}{\rho(S)}}\ (\ >1)\implies \mathbb{E}\left(\Delta_t^{\delta/2}\mid\mathcal{F}_s\right)^{2/\delta}\leqslant \mathcal{Z}_s\ \exp\left[-(1-\epsilon)\,\Lambda(t-s)\right]\,\Delta_s$$

for some random process $\mathcal{Z}_t$ s.t. $\sup_{t\geqslant 0}\mathbb{E}\left(\mathcal{Z}_t^{\alpha\delta}\right)<\infty$.

Theorem 1.2 readily implies the stability δ-moment Lyapunov exponent estimates

$$\liminf_{t\to\infty}-\frac{1}{\delta t}\log\mathbb{E}(\Delta_t^{\delta})\geqslant\Lambda$$

In addition we have the non asymptotic estimates

$$\mathbb{E}\left(\Delta_t^{\delta/2}\right)^{2/\delta}\leqslant\ \nu(\alpha)\ \exp\left\{-(1-\epsilon)\,\Lambda\ (t-s)\right\}\,\mathbb{E}\left[\Delta_s^{\delta\alpha/(2\alpha-1)}\right]^{(2\alpha-1)/(\delta\alpha)}$$

with

$$\nu(\alpha):=\sup_{t\geqslant 0}\mathbb{E}\left(\mathcal{Z}_t^{\alpha\delta}\right)^{1/(\delta\alpha)}<\infty$$

We end this section with some comments on our regularity conditions. Notice that Λ doesn't depends on the parameter δ nor on $\rho(S)$. As mentioned above, we believe that these technical conditions can somehow be relaxed. These conditions are stronger than the ones discussed in [6] for linear-Gaussian models. In contrast with the linear case, the Riccati equation in nonlinear settings is a stochastic process in matrix spaces. For this class of models, these technical conditions are used to control the fluctuations of the stochastic Riccati equation entering into the EKF.

2 Stability properties of quadratic type processes

Let $(\mathcal{U}_t,\mathcal{V}_t,\mathcal{W}_t,\mathcal{Y}_t)$ be some non negative processes defined on some probability space $(\Omega,\mathcal{F},\mathbb{P})$ equipped with a filtration $\mathcal{F}=(\mathcal{F}_t)_{t\geqslant 0}$ of σ-fields. Also let $(\mathcal{Z}_t,\mathcal{Z}_t^+)$ be some processes and $\mathcal{M}_t$ be some continuous $\mathcal{F}_t$-martingale. We use the notation

$$d\mathcal{Y}_t\leqslant\mathcal{Z}_t^+\ dt+d\mathcal{M}_t\Longleftrightarrow\left(d\mathcal{Y}_t=\mathcal{Z}_t\ dt+d\mathcal{M}_t\quad\text{with}\quad\mathcal{Z}_t\leqslant\mathcal{Z}_t^+\right)\tag{11}$$

Let us mention some useful properties of the above stochastic inequalities.

Let $(\overline{\mathcal{Y}}_t, \overline{\mathcal{Z}}_t^+, \overline{\mathcal{Z}}_t, \overline{\mathcal{M}}_t)$ be another collection of processes satisfying the above inequalities. In this case it is readily checked that

$$d(\mathcal{Y}_t + \overline{\mathcal{Y}}_t) \leqslant (\mathcal{Z}_t^+ + \overline{\mathcal{Z}}_t^+)\ dt + d(\mathcal{M}_t + \overline{\mathcal{M}}_t)$$

and

$$d(\mathcal{Y}_t\overline{\mathcal{Y}}_t) \leqslant \left[\overline{\mathcal{Z}}_t^+\mathcal{Y}_t + \mathcal{Z}_t^+\overline{\mathcal{Y}}_t + \partial_t\langle\mathcal{M},\overline{\mathcal{M}}\rangle_t\right]\ dt + \mathcal{Y}_t\ d\overline{\mathcal{M}}_t + \overline{\mathcal{Y}}_t\ d\mathcal{M}_t$$

Let $(\mathcal{H}, \langle ., .\rangle)$ be some inner product space, and let $\mathcal{A}_t\ :\ x \in \mathcal{H} \mapsto \mathcal{A}_t(x) \in \mathcal{H}$ be a linear operator-valued stochastic process with finite logarithmic norm $\rho(\mathcal{A}_t) < \infty$. Consider an $\mathcal{H}$-valued stochastic process $\mathcal{X}_t$ such that

$$d\|\mathcal{X}_t\|^2 \leqslant [\langle \mathcal{X}_t, \mathcal{A}_t\mathcal{X}_t\rangle + \mathcal{U}_t]\ dt + d\mathcal{M}_t \tag{12}$$

for some continuous $\mathcal{F}_t$-martingale $\mathcal{M}_t$ with angle bracket satisfying the following property

$$\partial_t\langle\mathcal{M}\rangle_t \leqslant \mathcal{V}_t\ \|\mathcal{X}_t\|^2 + \mathcal{W}_t\ \|\mathcal{X}_t\|^4$$

This section is concerned with the long time quantitative behavior of the above quadratic type processes. The main difficulty here comes from the fact that $\mathcal{A}_t$ is a stochastic flow of operators. As a result we cannot apply conventional Lyapunov techniques based on Dynkin's formula, supermartingale theory and/or more conventional Gronwall type estimates.

Next theorem provides a way to estimate these processes in terms of geometric type processes and exponential martingales.

Theorem 2.1. *When $\mathcal{U}_t = 0 = \mathcal{V}_t$ we have the almost sure estimate*

$$\|\mathcal{X}_t\|^2 \leqslant \|\mathcal{X}_0\|^2\ \exp\left(\int_0^t \rho(\mathcal{A}_s)\ ds\right)\ \exp\left(\int_0^t \sqrt{\mathcal{W}_s}\ d\mathcal{N}_s - \frac{1}{2}\ \int_0^t \mathcal{W}_s\ ds\right)$$

with a martingale $\mathcal{N}_t$ s.t. $\partial_t\langle\mathcal{N}\rangle_t \leqslant 1$. More generally, for any $n \geqslant 1$ we have

$$\begin{aligned}\mathbb{E}\left(\|\mathcal{X}_t\|^n\ |\ \mathcal{F}_0\right)^{2/n} \leqslant \mathbb{E}\left[\exp\left(n\ \int_0^t \left\{\rho(\mathcal{A}_s)\ + \frac{(n-1)}{2}\ \mathcal{W}_s\right\}\ ds\ |\ \mathcal{F}_0\right)\right]^{1/n} \\ \times \left\{\|\mathcal{X}_0\|^2 + \int_0^t\ \left(\mathbb{E}\left[\overline{\mathcal{U}}_s^n\ |\ \mathcal{F}_0\right]^{1/n} + \frac{(n-1)}{2}\ \mathbb{E}\left[\overline{\mathcal{V}}_s^n\ |\ \mathcal{F}_0\right]^{1/n}\right)\ ds\right\}\end{aligned} \tag{13}$$

with the rescaled processes

$$\overline{\mathcal{U}}_t/\mathcal{U}_t := \exp\left(-\int_0^t [\rho(\mathcal{A}_s) + (n-1)\mathcal{W}_s]\, ds\right) := \overline{\mathcal{V}}_t/\mathcal{V}_t$$

The proof of this theorem is rather technical thus it is housed in section 5.2 in the appendix.

Corollary 2.2. *When $\mathcal{U}_t = 0 = \mathcal{V}_t$ we have*

$$\mathbb{E}\left(\|\mathcal{X}_t\|^n\ |\ \mathcal{F}_0\right) \leqslant \mathbb{E}\left(\exp\left(\int_0^t \left[n\ \rho(\mathcal{A}_s)\ + \frac{n(n-1)}{2}\ \mathcal{W}_s\right]\ ds\right)\ |\ \mathcal{F}_0\right)^{1/2}\ \|\mathcal{X}_0\|^n \tag{14}$$

When $\rho(\mathcal{A}_t) \leqslant -a_t$ and $\mathcal{W}_t \leqslant w_t$ for some constants a_t, w_t, and $\mathcal{X}_0 = 0$ we have

$$\mathbb{E}\left(\|\mathcal{X}_t\|^n\right)^{2/n} \leqslant \int_0^t \exp\left(-\left[\int_s^t \lambda_n(a_u, w_u)\ du + \frac{n-1}{2}\int_0^s w_u\ du\right]\right) \times \left[\mathbb{E}\left(\mathcal{U}_s^n\right)^{1/n} + \frac{(n-1)}{2}\ \mathbb{E}\left(\mathcal{V}_s^n\right)^{1/n}\right]\ ds \tag{15}$$

with

$$\lambda_n(a_s, w_s) := a_s - \frac{n-1}{2}\ w_s$$

Proof. The first assertion is a direct consequence of the estimates stated in theorem 2.1. Replacing $\mathcal{W}_t$ and $\rho(\mathcal{A}_t)$ by w_t and $(-a_t)$ from the start in the proof of theorem 2.1 we find that

$$\begin{aligned}\mathbb{E}\left(\|\mathcal{X}_t\|^n \mid \mathcal{F}_0\right)^{2/n} \leqslant\ & \exp\left(-\int_0^t \lambda_n(a_s, w_s)\ ds\right)\|\mathcal{X}_0\|^2 \\ & + \int_0^t \exp\left(-\left[\int_s^t \lambda_n(a_u, w_u)\ du + \frac{n-1}{2}\int_0^s w_u\ du\right]\right) \\ & \times \left[\mathbb{E}\left(\mathcal{U}_s^n \mid \mathcal{F}_0\right)^{1/n} + \frac{(n-1)}{2}\ \mathbb{E}\left(\mathcal{V}_s^n \mid \mathcal{F}_0\right)^{1/n}\right]\ ds\end{aligned}$$

In the above display we have used the fact that

$$\begin{aligned}&\int_s^t \left(-a_u + \frac{n-1}{2}\ w_u\right)\ du + \int_0^s \left(-a_u + \frac{n-1}{2}\ w_u\right)\ du + \int_0^s \left(a_u - (n-1)\ w_u\right)\ du \\ &\leqslant -\int_s^t \left(a_u - \frac{n-1}{2}\ w_u\right)\ du - \frac{n-1}{2}\int_0^s w_u\ du\end{aligned}$$

This ends the proof of the corollary.

■

Proposition 2.3. *Assume that $\rho(\mathcal{A}_t) \leqslant -a$ for some parameter $a > 0$, and $\mathcal{X}_0 = 0 = \mathcal{W}_t$. Also assume that for any $n \geqslant 1$ and any $t \geqslant 0$ we have*

$$\mathbb{E}\left(\mathcal{U}_t^n\right)^{1/n} \leqslant u_t \quad \text{and} \quad \mathbb{E}\left(\mathcal{V}_t^n\right)^{1/n} \leqslant v_t$$

for some functions $u_t, v_t \geqslant 0$. In this situation, for any $\epsilon \in]0,1]$ we have the uniform estimates

$$\sup_{t\geqslant 0}\mathbb{E}\left(\exp\left[\frac{(1-\epsilon)}{e}\ \frac{1}{2v_t(a)}\ \|\mathcal{X}_t\|^2\right]\right) \leqslant \frac{1}{2}\ \exp\left(\frac{1-\epsilon}{e}\frac{u_t(a)}{v_t(a)}\right) + \frac{e}{2\sqrt{2}}\ \frac{1}{\sqrt{\epsilon}} \tag{16}$$

for any functions $(u_t(a), v_t(a))$ such that

$$\int_0^t e^{-a(t-s)}\ u_s\ ds \leqslant u_t(a) \quad \text{and} \quad \int_0^t e^{-a(t-s)}\ v_s\ ds \leqslant v_t(a)$$

In addition, when $v_t = v$ for any $\epsilon \in [0,1]$ we have

$$\mathbb{E}\left(\exp\left[\frac{a^2}{4v}\ \epsilon\ \int_0^t\ \|\mathcal{X}_s\|^2\ ds\right]\right)\ \leqslant\ \mathbb{E}\left(\exp\left[\frac{a}{v}\ \frac{\epsilon}{1+\sqrt{1-\epsilon}}\ \int_0^t \mathcal{U}_s\ ds\right]\right)^{1/2} \tag{17}$$

The proof of the proposition is provided in the appendix, section 5.3.

We end this section with some comments on the estimate (17). Let us suppose that

$$d\|\mathcal{X}_t\|^2 = \left[-a\ \|\mathcal{X}_t\|^2 + u\right]\ dt + d\mathcal{M}_t$$

for some $u \geqslant 0$ (with $\mathcal{X}_0 = 0$). In this case, by Jensen's inequality we have

$$a\ \mathbb{E}(\|\mathcal{X}_t\|^2) = u\ \left(1-e^{-at}\right)$$

$$\Rightarrow \mathbb{E}\left(\exp\left[\frac{a^2}{v}\ \epsilon\ \int_0^t\ \|\mathcal{X}_s\|^2\ ds\right]\right) \geqslant \exp\left[\frac{a}{v}\ \epsilon\ u\ t\ \left(1-\frac{1}{at}\left[1-e^{-at}\right]\right)\right] \geqslant \exp\left[\frac{a}{v}\ \epsilon\ u\ t\ \left(1-\frac{1}{at}\right)\right]$$

The r.h.s. of (17) gives the estimate

$$\mathbb{E}\left(\exp\left[\frac{a^2}{v}\ \epsilon\ \int_0^t\ \|\mathcal{X}_s\|^2\ ds\right]\right) \leqslant \exp\left[\frac{a}{v}\ \epsilon\ u\ t\ \frac{1}{1+\sqrt{1-\epsilon}}\right]$$

The above estimates coincides for any $\epsilon \in [0,1]$ and any $u \geqslant 0$ as soon as

$$t \geqslant \frac{1}{a}\ \left(1+\frac{1}{\sqrt{1-\epsilon}}\right)$$

3 Stochastic stability properties

3.1 The signal process

This section is mainly concerned with the stochastic stability properties of the signal process. One natural way to derive some useful concentration inequalities is to compare the flow of the stochastic process with the one of the noise free deterministic system discussed in the beginning of section 1.2.

We start with a brief review on the long time behavior of the semigroup $\varphi_t(x)$. It is readily check that

$$\partial_t\|\varphi_t(x)-\varphi_t(y)\|^2\ \ \leqslant\ \ -2\lambda_A\|\varphi_t(x)-\varphi_t(y)\|^2 \Rightarrow \|\varphi_t(x)-\varphi_t(y)\| \leqslant e^{-\lambda_A t}\ \|x-y\|$$

This contraction property ensures the existence and the uniqueness of a fixed point

$$\forall t \geqslant 0 \quad \varphi_t(x_\star) := x_\star \Longleftrightarrow A(x_\star) = 0 \Longrightarrow \|\varphi_t(x)-x_\star\| \leqslant e^{-\lambda_A t}\ \|x-x_\star\|$$

We let $\delta\phi_t(x)$ be the Jacobian of the stochastic flow $\phi_t(x)$. We have the matrix valued equation

$$\partial_t\delta\phi_t(x) = \partial A(\phi_t(x))\ \delta\phi_t(x) \Rightarrow \delta\phi_t(x)\ u = \exp\left(\int_0^t \partial A(\phi_s(x))\ ds\right)\ u$$

for any $u \in \mathbb{R}^{r_1}$. This implies that

$$\|\delta\phi_t(x)\| := \sup_{\|u\|\leqslant 1} \|\delta\phi_t(x)\ u\| \leqslant \exp\left(-\lambda_{\partial A} t/2\right) \longrightarrow_{t\rightarrow\infty} 0$$

Using the formula

$$\phi_t(y) - \phi_t(x) = \int_0^1 \delta\phi_t(x + \epsilon(y-x))\ (y-x)\ d\epsilon$$

we easily check the almost sure exponential stability property

$$\|\phi_t(x) - \phi_t(y)\| \leqslant \exp\left(-\lambda_{\partial A} t/2\right) \|x - y\| \tag{18}$$

The same analysis applies to estimate the Jacobian $\delta\varphi_t(x)$ of the deterministic flow $\varphi_t(x)$. Using the estimate

$$\|\phi_t(X_0) - \phi_t\left(\mathbb{E}\left(X_0\right)\right)\| \leqslant e^{-\lambda_{\partial A} t/2}\ \|X_0 - \mathbb{E}\left(X_0\right)\|$$

we also have

$$\lambda_{\partial A} \int_0^t \|\phi_s(X_0) - \phi_s\left(\mathbb{E}\left(X_0\right)\right)\|^2\ ds \leqslant \|X_0 - \mathbb{E}\left(X_0\right)\|^2$$

from which we conclude that

$$(4) \Longrightarrow \mathbb{E}\left(\exp\left(\frac{\lambda_{\partial A}}{\chi(P_0)} \int_0^t \|\phi_s(X_0) - \phi_s\left(\mathbb{E}\left(X_0\right)\right)\|^2\ ds\right)\right) \leqslant e$$

Next proposition quantify the relative stochastic stability of the flows (φ_t, ϕ_t) in terms of $\mathbb{L}_n$-norms and χ-square uniform Laplace estimates.

Proposition 3.1. *For any $n \geqslant 1$ and any $x \in \mathbb{R}^{r_1}$ we have the uniform moment estimates*

$$\mathbb{E}\left(\|\phi_t(x) - \varphi_t(x)\|^{2n}\right)^{1/n} \leqslant (n - 1/2)\ \mathrm{tr}(R_1)/\lambda_A \tag{19}$$

In addition, for any $\epsilon \in]0,1]$ we have the uniform Laplace estimates

$$\sup_{t\geqslant 0} \mathbb{E}\left(\exp\left[\frac{(1-\epsilon)}{4e}\ \frac{\lambda_A}{\mathrm{tr}(R_1)}\ \|\phi_t(x) - \varphi_t(x)\|^2\right]\right) \leqslant \frac{e}{2\sqrt{2}} \frac{1}{\sqrt{\epsilon}} + \frac{1}{2} \exp\left[\frac{1-\epsilon}{4e}\right]$$

as well as

$$\mathbb{E}\left(\exp\left[\frac{\lambda_A^2}{4\mathrm{tr}(R_1)}\ \epsilon \int_0^t \|\phi_s(x) - \varphi_s(x)\|^2\ ds\right]\right) \leqslant \exp\left[\frac{\lambda_A}{2}\ \epsilon\ t\right]$$

Combining (19) with the concentration inequality (26) we prove that the probability of the events

$$\|\phi_t(x) - \varphi_t(x)\|^2 \leqslant \varpi(\delta)\ \mathrm{tr}(R_1)/\lambda_A$$

is greater than $1 - e^{-\delta}$, for any $\delta \geqslant 0$ and any initial states $x \in \mathbb{R}^{r_1}$. This ends the proof of (7).

Proof of proposition 3.1:

We have

$$d(X_t - x_t) = [A(X_t) - A(x_t)]\ dt + R_1^{1/2} dW_t$$

with $X_0 = x_0$, and therefore

$$\begin{array}{rcl} d\|X_t - x_t\|^2 & = & [2\langle A(X_t) - A(x_t), X_t - x_t, \rangle + \mathrm{tr}(R_1)]\ dt + dM_t \\ & \leqslant & \left[-2\lambda_A \|X_t - x_t\|^2 + \mathrm{tr}(R_1)\right]\ dt + dM_t \end{array}$$

with the martingale

$$dM_t := 2\langle X_t - x_t, R_1^{1/2} dW_t\rangle \Longrightarrow \partial_t \langle M\rangle_t = 4\ \mathrm{tr}\left(R_1(X_t - x_t)(X_t - x_t)'\right) \leqslant 4\ \mathrm{tr}(R_1)\ \|X_t - x_t\|^2$$

The end of the proof is now a direct consequence of (15) and proposition 2.3 applied to

$$\mathcal{X}_t = \|X_t - x_t\| \qquad \mathcal{A}_t x := -ax = -2\lambda_A x \qquad \mathcal{U}_t = u = \mathrm{tr}(R_1) \quad \text{and} \quad \mathcal{V}_t = v = 4\ \mathrm{tr}(R_1)$$

The proof of the proposition is now completed. ■

3.2 The Extended Kalman-Bucy filter

This section is mainly concerned with the stochastic stability and the concentration properties of the semigroup of the EKF stochastic process. As for the signal process discussed in section 3.1 these properties are related to $\mathbb{L}_p$-mean error estimates and related χ-square type Laplace inequalities. Our main results are described by the following theorem.

Let $(\widehat{X}_t, P_t)$ be the solution of the evolution equations (3) starting at $(\widehat{X}_0, P_0)$.

Theorem 3.2. *For any $n \geqslant 1$ we have*

$$\mathbb{E}\left(\|\phi_t(\widehat{X}_0) - \widehat{X}_t\|^n\right)^{2/n} \leqslant (2n-1)\left\{\frac{\mathrm{tr}(R_1)}{\lambda_A}\ \frac{\sigma_{\partial A}^2}{2} + \frac{|e^{-\lambda_A t} - e^{-\lambda_{\partial A} t}|}{|\lambda_A - \lambda_{\partial A}|}\ \rho(S)\ \mathrm{tr}(P_0)^2\right\} \tag{20}$$

For any $\epsilon \in]0,1]$ and any P_0 there exists some time horizon $t_0(\epsilon, P_0)$ such that

$$\sup_{t \geqslant t_0(\epsilon, P_0)} \mathbb{E}\left(\exp\left[\frac{(1-\epsilon)}{4e\sigma_{\partial A}^2}\ \frac{\lambda_A}{\mathrm{tr}(R_1)}\ \ \|\phi_t(\widehat{X}_0) - \widehat{X}_t\|^2\right]\right) \leqslant \frac{1}{2}\ \exp\left(\frac{1-\epsilon}{4e}\right) + \frac{e}{2\sqrt{2}}\ \frac{1}{\sqrt{\epsilon}} \tag{21}$$

In addition for any $t \geqslant s \geqslant 0$ and any $\epsilon \in]0,1]$ we have

$$\mathbb{E}\left(\exp\left[\frac{\epsilon}{1 + \pi_{\partial A}(s)}\ \frac{\lambda_A^2}{4\mathrm{tr}(R_1)}\ \int_s^t \|\phi_{r-s}(\widehat{X}_s) - \widehat{X}_r\|^2\ ds\right]\right) \leqslant \exp\left[\frac{\lambda_A}{2}\ \epsilon\ (t-s)\right] \tag{22}$$

Before getting into the details of the proof of this theorem we mention that (8) is a direct consequence of (20) combined with (26) and (18). Indeed, applying (26) to

$$Z = \|\phi_t(\widehat{X}_0) - \widehat{X}_t\| \quad \text{and} \quad z^2 = 4\left\{\frac{\mathrm{tr}(R_1)}{\lambda_A}\ \frac{\sigma_{\partial A}^2}{2} + \frac{|e^{-\lambda_A t} - e^{-\lambda_{\partial A} t}|}{|\lambda_A - \lambda_{\partial A}|}\ \rho(S)\ \mathrm{tr}(P_0)^2\right\}$$

by (20) we readily check that the probability of the events

$$\begin{array}{l} \|\phi_t(x) - \Phi_t(\widehat{x}, p)\|^2 \\ \leqslant 2\exp\left(-\lambda_{\partial A} t\right)\ \|x - \widehat{x}\|^2 + 8\varpi(\delta)\ \left\{\frac{\mathrm{tr}(R_1)}{\lambda_A}\ \frac{\sigma_{\partial A}^2}{2} + \frac{|e^{-\lambda_A t} - e^{-\lambda_{\partial A} t}|}{|\lambda_A - \lambda_{\partial A}|}\ \rho(S)\ \mathrm{tr}(P_0)^2\right\} \end{array}$$

is greater than $1-e^{-\delta}$, for any $\delta \geqslant 0$ and any initial states $(x,\widehat{x},p) \in \mathbb{R}^{r_1+r_1+(r_1\times r_1)}$. In this connection, the Laplace estimates (22) readily implies that the probability of the events

$$\frac{1}{t-s}\int_s^t \|\phi_u(\widehat{X}_s)-\widehat{X}_u\|^2\, du \leqslant \left(\frac{1}{2}+\frac{\delta}{\lambda_A}\right)\ (1+\pi_{\partial A}(s))\ \frac{4\mathrm{tr}(R_1)}{\lambda_A}$$

is greater than $1-e^{-\delta}$, for any $\delta \geqslant 0$ and any time horizon t.

Now we come to the proof of the theorem.

Proof of theorem 3.2:
We set $\mathcal{X}_t := \phi_t(\widehat{X}_0)-\widehat{X}_t$. We have

$$d\|\mathcal{X}_t\|^2 \leqslant \left(2\ \langle A(\phi_t(\widehat{X}_0))-A(\widehat{X}_t),\mathcal{X}_t\rangle - 2\ \langle P_tS\mathcal{X}_t,\mathcal{X}_t\rangle + \mathrm{tr}(R_1)+\mathrm{tr}(SP_t^2)\right)\ dt + d\mathcal{M}_t$$

with the martingale

$$d\mathcal{M}_t := 2\mathcal{X}_t'\left(R_1^{-1/2}dW_t - P_tB'R_2^{-1/2}dV_t\right) \to \partial_t\langle\mathcal{M}_t\rangle_t$$

This yields the estimate

$$d\|\mathcal{X}_t\|^2 \leqslant \left[-2\lambda_A\ \|\mathcal{X}_t\|^2+\mathcal{U}_t\right]\ dt + d\mathcal{M}_t \quad \text{with} \quad \mathcal{U}_t = u_t := \left[\mathrm{tr}(R_1)+\tau_t^2(P)\ \rho(S)\right]$$

Also observe that

$$\partial_t\langle\mathcal{M}\rangle_t \quad \leqslant \quad 4\ \|\mathcal{X}_t\|^2\ \left(\mathrm{tr}(R_1)+\mathrm{tr}(SP_t^2)\right) \leqslant \mathcal{V}_t\ \|\mathcal{X}_t\|^2 \quad \text{with } \mathcal{V}_t = v_t := 4u_t.$$

On the other hand we have

$$2e^{-2\lambda_A t}\int_0^t e^{2(\lambda_A-\lambda_{\partial A})s}\ ds = \frac{|e^{-\lambda_A t}-e^{-\lambda_{\partial A}t}|}{|\lambda_A-\lambda_{\partial A}|}$$

This implies that

$$\begin{aligned}\int_0^t e^{-2\lambda_A(t-s)}\ \tau_s^2\ ds &\leqslant 2\mathrm{tr}(P_0)^2\ e^{-2\lambda_A t}\int_0^t e^{-2\Delta_A s}\ ds + \frac{1}{\lambda_A}\ \left(\frac{\mathrm{tr}(R_1)}{\lambda_{\partial A}}\right)^2\\ &\leqslant \frac{|e^{-\lambda_A t}-e^{-\lambda_{\partial A}t}|}{|\lambda_A-\lambda_{\partial A}|}\ \mathrm{tr}(P_0)^2 + \frac{1}{\lambda_A}\ \left(\frac{\mathrm{tr}(R_1)}{\lambda_{\partial A}}\right)^2\end{aligned}$$

This implies that

$$\begin{aligned}\int_0^t e^{-2\lambda_A(t-s)}\ u_s\ ds &\leqslant \frac{\mathrm{tr}(R_1)}{2\lambda_A}+\rho(S)\left[\frac{|e^{-\lambda_A t}-e^{-\lambda_{\partial A}t}|}{|\lambda_A-\lambda_{\partial A}|}\ \mathrm{tr}(P_0)^2 + \frac{1}{\lambda_A}\ \left(\frac{\mathrm{tr}(R_1)}{\lambda_{\partial A}}\right)^2\right]\\ &= \frac{\mathrm{tr}(R_1)}{\lambda_A}\left(\frac{1}{2}+\frac{\rho(S)}{\lambda_{\partial A}}\ \frac{\mathrm{tr}(R_1)}{\lambda_{\partial A}}\right)+\frac{|e^{-\lambda_A t}-e^{-\lambda_{\partial A}t}|}{|\lambda_A-\lambda_{\partial A}|}\ \rho(S)\ \mathrm{tr}(P_0)^2\\ &:= u_t(a) := v_t(a)/4\end{aligned}$$

Applying proposition 2.3 to $\mathcal{A}_t x := -ax = -2\lambda_A x$, we find that

$$\mathbb{E}\left(\exp\left[\frac{(1-\epsilon)}{e}\ \frac{1}{8u_t(a)}\ \|\mathcal{X}_t\|^2\right]\right) \leqslant \frac{1}{2}\ \exp\left(\frac{1-\epsilon}{4e}\right)+\frac{e}{2\sqrt{2}}\ \frac{1}{\sqrt{\epsilon}} \quad \text{for any } \epsilon \in]0,1].$$

Using the fact that for any non negative real numbers x, y, λ we have

$$\frac{1}{x+e^{-\lambda}\ y} \quad = \quad \frac{1}{x}\left(1-\frac{e^{-\lambda}\ y/x}{1+e^{-\lambda}\ y/x}\right) \geqslant \frac{1}{x}\left(1-e^{-\lambda}\ y/x\right)$$

and

$$\frac{|e^{-\lambda_A t}-e^{-\lambda_{\partial A} t}|}{|\lambda_A-\lambda_{\partial A}|} \longrightarrow_{t\to\infty} 0$$

we find that

$$\frac{1}{u_t(a)} \quad \geqslant \quad (1-\epsilon)\ \frac{\lambda_A}{\lambda_{\partial A}}\ \left[\frac{\mathrm{tr}(R_1)}{\lambda_{\partial A}}\left(\frac{1}{2}+\frac{\rho(S)}{\lambda_{\partial A}}\ \frac{\mathrm{tr}(R_1)}{\lambda_{\partial A}}\right)\right]^{-1}$$

for any $t \geqslant t(\epsilon)$, for any $\epsilon \in [0,1[$ and some $t(\epsilon)$.

The end of the proof is now a direct consequence of (15) and proposition 2.3 applied to

$$\mathcal{A}_t x := -ax = -2\lambda_A x \quad \text{and} \quad u_t = v_t/4 \leqslant u = v/4 := \left[\mathrm{tr}(R_1)+\tau_s^2(P)\ \rho(S)\right]$$

with $t \in [s,\infty[$. The proof of the theorem is now completed. ■

4 Proof of theorem 1.2

We let $(\widehat{X}_t, P_t)$ be the solution of the equations (3) starting at $(\widehat{X}_0, P_0)$. We denote by $(\check{X}_t, \check{P}_t)$ the solution of these equations starting at some possibly different state $(\check{X}_0, \check{P}_0)$. Firstly we have

$$((5),\ (19) \text{ and } (20)) \implies \forall n \geqslant 1 \qquad \sup_{t\geqslant 0} \mathbb{E}\left([\|\widehat{X}_t-\check{X}_t\|^2+\|P_t-\check{P}_t\|_F^2]^n\right) < \infty$$

We couple the equations with the same observation processes. In this situation we find the evolution equation

$$\begin{aligned} d(\widehat{X}_t-\check{X}_t) \quad &= \quad \left(A(\widehat{X}_t)-A(\check{X}_t)\right)dt + P_tB'R_2^{-1}\ \left[dY_t-B\widehat{X}_tdt\right] \\ &\qquad\qquad -\check{P}_tB'R_2^{-1}\ \left[dY_t-B\check{X}_tdt\right] \\ &= \quad \left(\left[A(\widehat{X}_t)-A(\check{X}_t)\right]+\check{P}_tB'R_2^{-1}\ \left[B\ (\check{X}_t-X_t)\right]\right)dt \\ &\qquad +(P_t-\check{P}_t)B'R_2^{-1}\ \left[B(X_t-\widehat{X}_t)\right]\ dt+\check{P}_tB'R_2^{-1}\ \left[B(X_t-\widehat{X}_t)\right]dt+dM_t \end{aligned}$$

with the martingale

$$dM_t := \left[P_t-\check{P}_t\right]B'R_2^{-1/2}\ dV_t$$

This implies that

$$\begin{aligned} &d(\widehat{X}_t-\check{X}_t) \\ &= \left([A(\widehat{X}_t)-A(\check{X}_t)]+\check{P}_tS\ (\check{X}_t-\widehat{X}_t)+(P_t-\check{P}_t)S(X_t-\widehat{X}_t)\right)\ dt+dM_t \end{aligned}$$

from which we conclude that

$$\begin{aligned}
& d\|\widehat{X}_t-\check{X}_t\|^2 \\
& = 2\langle \widehat{X}_t-\check{X}_t, [A(\widehat{X}_t)-A(\check{X}_t)]-\check{P}_tS\ (\widehat{X}_t-\check{X}_t)+(P_t-\check{P}_t)S(X_t-\widehat{X}_t)\ \rangle\ dt \\
& \qquad\qquad +\mathrm{tr}\left(S\left[P_t-\check{P}_t\right]^2\right)dt+d\overline{M}_t
\end{aligned} \tag{23}$$

with the martingale

$$d\overline{M}_t=2\ \langle \widehat{X}_t-\check{X}_t, dM_t\rangle$$

The angle bracket of $\overline{M}_t$ is given by

$$\begin{aligned}
\partial_t\langle\overline{M}\rangle_t &= 4\ \mathrm{tr}\left(\left[P_t-\check{P}_t\right]S\left[P_t-\check{P}_t\right](\widehat{X}_t-\check{X}_t)(\widehat{X}_t-\check{X}_t)'\right) \\
&= 4\ \langle(\widehat{X}_t-\check{X}_t),\left[P_t-\check{P}_t\right]S\left[P_t-\check{P}_t\right](\widehat{X}_t-\check{X}_t)\rangle \\
&\leqslant 4\ \rho(S)\ \|\widehat{X}_t-\check{X}_t\|^2\ \|P_t-\check{P}_t\|_F^2\leqslant 2\ \rho(S)\ \left(\|\widehat{X}_t-\check{X}_t\|^2+\|P_t-\check{P}_t\|_F^2\right)^2
\end{aligned}$$

Recalling that $\lambda_A\geqslant\lambda_{\partial A}/2$, also observe that the drift term in (23) is bounded by

$$-\lambda_{\partial A}\ \|\widehat{X}_t-\check{X}_t\|^2+2\beta_t\ \|\widehat{X}_t-\check{X}_t\|\|P_t-\check{P}_t\|_F+\rho(S)\ \|P_t-\check{P}_t\|_F^2$$

with

$$\beta_t:=\rho(S)\|X_t-\widehat{X}_t\|$$

In much the same way we have

$$\begin{aligned}
& \partial_t(P_t-\check{P}_t) \\
& = \left(\partial A(\widehat{X}_t)P_t-\partial A(\check{X}_t)\check{P}_t\right)+\left(\partial A(\widehat{X}_t)P_t-\partial A(\check{X}_t)\check{P}_t\right)'+\check{P}_t\ S\check{P}_t-P_t\ SP_t \\
& = \left([\partial A(\widehat{X}_t)-\partial A(\check{X}_t)]P_t+\partial A(\check{X}_t)[P_t-\check{P}_t]\right)+\left([\partial A(\widehat{X}_t)-\partial A(\check{X}_t)]P_t+\partial A(\check{X}_t)[P_t-\check{P}_t]\right)' \\
& \qquad +\tfrac{1}{2}\ (\check{P}_t+P_t)S(\check{P}_t-P_t)+\tfrac{1}{2}\ (\check{P}_t-P_t)S(\check{P}_t+P_t)
\end{aligned}$$

In the last assertion we have used the matrix decomposition

$$PSP-QSQ=\frac{1}{2}\ (P+Q)S(P-Q)+\frac{1}{2}\ (P-Q)S(P+Q)$$

Recalling that

$$2^{-1}\partial_t\|P_t-\check{P}_t\|^2=2^{-1}\partial_t\langle P_t-\check{P}_t,P_t-\check{P}_t\rangle=\langle P_t-\check{P}_t,\partial_t(P_t-\check{P}_t)\rangle=\mathrm{tr}\left((P_t-\check{P}_t)\partial_t(P_t-\check{P}_t)\right)$$

we find that

$$\begin{aligned}
2^{-1}\partial_t\|P_t-\check{P}_t\|_F^2 &= 2\mathrm{tr}\left(\partial A(\check{X}_t)(P_t-\check{P}_t)^2\right)+2\mathrm{tr}\left([\partial A(\widehat{X}_t)-\partial A(\check{X}_t)]P_t(P_t-\check{P}_t)\right) \\
&\qquad -\mathrm{tr}\left((\check{P}_t+P_t)S(\check{P}_t-P_t)^2\right) \\
&\leqslant 2\mathrm{tr}\left(\partial A(\check{X}_t)(P_t-\check{P}_t)^2\right)+2\mathrm{tr}\left([\partial A(\widehat{X}_t)-\partial A(\check{X}_t)]P_t(P_t-\check{P}_t)\right)
\end{aligned}$$

This implies that

$$\partial_t \|P_t - \check{P}_t\|_F^2 \quad \leqslant \quad -2\lambda_{\partial A} \, \|P_t - \check{P}_t\|_F^2 + 2\alpha_t \|P_t - \check{P}_t\|_F \|\widehat{X}_t - \check{X}_t\|$$

with

$$\alpha_t := 2\kappa_{\partial A} \tau_t(P)$$

We set

$$\mathcal{X}_t = \begin{pmatrix} \|\widehat{X}_t - \check{X}_t\| \\ \|P_t - \check{P}_t\|_F \end{pmatrix} \in \mathcal{H} := \mathbb{R}^2 \Longrightarrow d\|\mathcal{X}_t\|^2 \leqslant \langle \mathcal{X}_t, \mathcal{A}_t \mathcal{X}_t \rangle \, dt + d\mathcal{M}_t$$

with

$$\mathcal{A}_t = \begin{pmatrix} -\lambda_{\partial A} & 2\beta_t \\ 2\alpha_t & -2\lambda_{\partial A} + \rho(S) \end{pmatrix} \quad \text{and} \quad \mathcal{M}_t = \overline{M}_t$$

Notice that

$$(\mathcal{A}_t + \mathcal{A}_t')/2 = \begin{pmatrix} -\lambda_{\partial A} & \beta_t + \alpha_t \\ \beta_t + \alpha_t & -2\lambda_{\partial A} + \rho(S) \end{pmatrix}$$

Observe that

$$\begin{aligned} \rho(\mathcal{A}_t) \quad &:= \quad \lambda_{max} \left((\mathcal{A}_t + \mathcal{A}_t')/2 \right) \\ &= \quad -\frac{1}{2} \left(3\lambda_{\partial A} - \rho(S) \right) + \sqrt{\frac{1}{4} \left(\lambda_{\partial A} - \rho(S) \right)^2 + (\beta_t + \alpha_t)^2} \leqslant -\lambda_{\partial A} + \beta_t + \alpha_t \\ &\leqslant \quad \overline{\rho}(\mathcal{A}_t) := -\left(\lambda_{\partial A} - 2\kappa_{\partial A} \tau_t(P) - \rho(S) \, \|X_t - \widehat{X}_t\| \right) \end{aligned}$$

The final step is based on the following technical lemma.

Lemma 4.1. *Assume condition (10) is satisfied for some $\alpha > 1$. In this situation, for any $\epsilon \in]0,1]$ there exists some time horizon s such that for any $t \geqslant s$ we have the almost sure estimate*

$$\delta := \frac{1}{2} \sqrt{\frac{\lambda_{\partial A}}{\rho(S)}} \Longrightarrow \ \mathbb{E} \left(\exp \left[\delta \int_s^t \left\{ \left(\overline{\rho}(\mathcal{A}_u) + (\delta - 1)\rho(S) \right) \right\} \, du \right] \mid \mathcal{F}_s \right)^{1/\delta} \tag{24}$$

$$\leqslant \mathcal{Z}_s \ \exp\left(-\left(1 - \epsilon\right) \Lambda (t - s) \right)$$

for some positive random process $\mathcal{Z}_t$ s.t. $\sup_{t \geqslant 0} \mathbb{E}\left(\mathcal{Z}_t^{\alpha\delta} \right) < \infty$.

The end of the proof of theorem 1.2 is a direct consequence of this lemma, so we give it first. Combining (24) with (13) we find that

$$\mathbb{E} \left(\|\mathcal{X}_t\|^\delta \mid \mathcal{F}_s \right)^{2/\delta} \leqslant \mathcal{Z}_s \ \exp\left(-\left(1 - \epsilon\right) \Lambda (t - s) \right) \, \|\mathcal{X}_s\|^2$$

This ends the proof of theorem 1.2.

Now we come to the proof of the lemma.

Proof of lemma 4.1:

For any $t \geqslant s \geqslant 0$ we have the estimate

$$\overline{\rho}(\mathcal{A}_t) \quad \leqslant \quad -\left(\lambda_{\partial A} - 2\kappa_{\partial A} \tau_s(P) \right) + \beta_t = -\Delta_{\partial A}(s) + \rho(S) \, \|X_t - \widehat{X}_t\|$$

with

$$\Delta_{\partial A}(s) := \lambda_{\partial A} - 2\kappa_{\partial A}\ \tau_s(P) \longrightarrow_{s\to\infty} \Delta_{\partial A} := \lambda_{\partial A} - 2\kappa_{\partial A}\ \mathrm{tr}(R_1)/\lambda_{\partial A} > 0$$

as soon as

$$\lambda_{\partial A} > \sqrt{2\kappa_{\partial A}\ \mathrm{tr}(R_1)}$$

For any $\epsilon \in]0,1]$, there exists some time horizon $\varsigma_\epsilon(P_0)$ such that

$$\begin{aligned} t \geqslant s \geqslant \varsigma_\epsilon(P_0) \quad &\Longrightarrow \quad (1-\epsilon) \leqslant \Delta_{\partial A}(s)/\Delta_{\partial A} \leqslant 1 \\ &\Rightarrow \quad \overline{\rho}(\mathcal{A}_t) \leqslant -(1-\epsilon)\Delta_{\partial A} + \rho(S)\ \|X_t - \widehat{X}_t\| \end{aligned}$$

On the other hand, the contraction inequality (18) implies that

$$\begin{aligned} \int_s^t \|X_r - \widehat{X}_r\|\ dr &= \int_s^t \|\phi_{r-s}(\phi_s(X_0)) - \widehat{X}_r\|\ dr \\ &\leqslant \int_s^t \|\phi_{r-s}(X_s) - \phi_{r-s}(\widehat{X}_s)\|\ dr + \int_s^t \|\phi_{r-s}(\widehat{X}_s) - \widehat{X}_r\|\ dr \\ &\leqslant \left(\int_s^t e^{-\lambda_{\partial A} r/2}\ dr\right)\ \|X_s - \widehat{X}_s\| + \int_s^t \|\phi_{r-s}(\widehat{X}_s) - \widehat{X}_r\|\ dr \\ &\leqslant 2\|X_s - \widehat{X}_s\|/\lambda_{\partial A} + \int_s^t \|\phi_{r-s}(\widehat{X}_s) - \widehat{X}_r\|\ dr \end{aligned}$$

The above inequality yields the almost sure estimate

$$\begin{aligned} &\mathbb{E}\left(\exp\left[\delta\left\{\int_s^t (\overline{\rho}(\mathcal{A}_u) + (\delta-1)\rho(S))\ du\right\}\right] \mid \mathcal{F}_s\right) \\ &\leqslant \exp\left[-\delta\left\{(1-\epsilon)\Delta_{\partial A} + (1-\delta)\rho(S)\right\}(t-s)\right] \\ &\qquad\qquad \times\ \mathcal{Z}_s\ \mathbb{E}\left(\exp\left[\delta\left\{\rho(S)\int_s^t \|\phi_{u-s}(\widehat{X}_s) - \widehat{X}_u\|\ \ du\right\}\right] \mid \mathcal{F}_s\right) \end{aligned}$$

with

$$\mathcal{Z}_s := \exp\left[2\rho(S)\|X_s - \widehat{X}_s\|/\lambda_{\partial A}\right]$$

Using the estimate $x - 1/4 \leqslant x^2$, which is valid for any x we have

$$\int_s^t ((\|\phi_{u-s}(\widehat{X}_s) - \widehat{X}_u\| - 1/4) + 1/4)\ du \leqslant (t-s)/4 + \int_s^t \|\phi_{u-s}(\widehat{X}_s) - \widehat{X}_u\|^2 du$$

we find that

$$\begin{aligned} &\mathbb{E}\left(\exp\left[\delta\left\{\int_s^t (\overline{\rho}(\mathcal{A}_u) + (\delta-1)\rho(S))\right\}\ du\right] \mid \mathcal{F}_s\right) \\ &\leqslant \exp\left[-\delta\left\{(1-\epsilon)\Delta_{\partial A} + (3/4-\delta)\rho(S)\right\}(t-s)\right] \\ &\qquad\qquad \times\ \mathcal{Z}_s\ \mathbb{E}\left(\exp\left[\delta\rho(S)\int_s^t \|\phi_{u-s}(\widehat{X}_s) - \widehat{X}_u\|^2\ du\right] \mid \mathcal{F}_s\right) \end{aligned}$$

By (10) we can also choose s sufficiently large so that

$$\delta=\frac{1}{2}\sqrt{\frac{\lambda_{\partial A}}{\rho(S)}}\Longrightarrow\delta\rho(S)4\mathrm{tr}(R_1)\ (1+\pi_{\partial A}(s))\leqslant\ \lambda_A\frac{\lambda_{\partial A}}{2}$$

In this situation, by (22) we have

$$\delta\rho(S)\leqslant\ \overbrace{\frac{\lambda_{\partial A}}{2\lambda_A}}^{\leqslant 1}\ \frac{1}{1+\pi_{\partial A}(s)}\ \frac{\lambda_A^2}{4\mathrm{tr}(R_1)}$$

$$\Rightarrow\ \mathbb{E}\left(\exp\left(\delta\rho(S)\ \int_s^t\ \|\phi_{r-s}(\widehat{X}_s)-\widehat{X}_r\|^2\ dr\right)\mid\mathcal{F}_s\right)\ \ \leqslant\ \ \exp\left[\frac{\lambda_{\partial A}}{4}\ (t-s)\right]$$

We conclude that

$$\mathbb{E}\left(\exp\left[\delta\left\{\int_s^t(\overline{\rho}(\mathcal{A}_u)+(\delta-1)\rho(S))\right\}\ du\right]\mid\mathcal{F}_s\right)^{1/\delta}$$

$$\leqslant\exp\left[-(1-\epsilon)\ \lambda_{\partial A}\left\{1-2\ \frac{\kappa_{\partial A}}{\lambda_{\partial A}}\ \frac{\mathrm{tr}(R_1)}{\lambda_{\partial A}}-\frac{1}{4\delta}+(3/4-\delta)\ \frac{\rho(S)}{\lambda_{\partial A}}\right\}(t-s)\right]\ \mathcal{Z}_s$$

$$\leqslant\exp\left[-(1-\epsilon)\ \Lambda(t-s)\right]\ \mathcal{Z}_s$$

The last assertion comes from the formula

$$\delta=\frac{1}{2}\sqrt{\frac{\lambda_{\partial A}}{\rho(S)}}$$

$$\Longrightarrow 1+\frac{3}{4}\ \frac{\rho(S)}{\lambda_{\partial A}}-2\ \frac{\kappa_{\partial A}}{\lambda_{\partial A}}\ \frac{\mathrm{tr}(R_1)}{\lambda_{\partial A}}-\frac{1}{4\delta}-\delta\ \frac{\rho(S)}{\lambda_{\partial A}}$$

$$=1-2\ \frac{\kappa_{\partial A}}{\lambda_{\partial A}}\ \frac{\mathrm{tr}(R_1)}{\lambda_{\partial A}}+\ \sqrt{\frac{\rho(S)}{\lambda_{\partial A}}}\left[\frac{3}{4}\sqrt{\frac{\rho(S)}{\lambda_{\partial A}}}-1\right]>0$$

On the other hand we have

$$\begin{aligned}\|X_s-\widehat{X}_s\| &= \|\phi_s(X_0)-\widehat{X}_s\|\leqslant\|\phi_s(X_0)-\phi_s(\widehat{X}_0)\|+\|\phi_s(\widehat{X}_0)-\widehat{X}_s\|\\ &\leqslant e^{-\lambda_{\partial A}s/2}\ \|X_0-\mathbb{E}\left(X_0\right)\|+\|\phi_s(\widehat{X}_0)-\widehat{X}_s\|\end{aligned}$$

This shows that

$$\mathcal{Z}_s\leqslant\exp\left(\frac{\rho(S)}{\lambda_{\partial A}}\ \|\phi_s(\widehat{X}_0)-\widehat{X}_s\|\right)\ \exp\left(\frac{\rho(S)}{\lambda_{\partial A}}\ e^{-\lambda_{\partial A}s/2}\ \|X_0-\mathbb{E}\left(X_0\right)\|\right)$$

Under the assumption (10) and using (21) we have

$$\alpha\ \delta\ \frac{\rho(S)}{\lambda_{\partial A}}=\frac{\alpha}{2}\ \sqrt{\frac{\rho(S)}{\lambda_{\partial A}}}<\frac{1}{8e\sigma_{\partial A}^2}\ \frac{\lambda_{\partial A}}{\mathrm{tr}(R_1)}$$

$$\Longrightarrow\exists p>1\ :\ \sup_{s\geqslant 0}\mathbb{E}\left(\exp\left(p\alpha\delta\|\phi_s(\widehat{X}_0)-\widehat{X}_s\|(\rho(S)/\lambda_{\partial A})\right)\right)<\infty$$

We can also choose s sufficiently large so that

$$\frac{\alpha}{\alpha-1}\ \delta\rho(S)\ e^{-\lambda_{\partial A}s/2} < \lambda_{\partial A}/\chi(P_0)$$

$$\Longrightarrow \mathbb{E}\left(\exp\left(\alpha\delta\frac{p}{p-1}\ (\rho(S)/\lambda_{\partial A})\ e^{-\lambda_{\partial A}s/2}\ \|X_0-\mathbb{E}\,(X_0)\,\|\right)\right) \leqslant e$$

This ends the proof of the lemma. ■

5 Appendix

5.1 Concentration properties and Laplace estimates

This section is mainly concerned with the proof of (4).

The initial state X_0 of the signal is a Gaussian random variable with mean $\widehat{X}_0$ and some covariance matrix P_0. In this case $X_0-\widehat{X}_0 \stackrel{law}{=} P_0^{1/2}W_1$ and

$$\mathbb{E}\left(\|X_0-\widehat{X}_0\|^{2n}\right) \leqslant \rho(P_0)^n\ \mathbb{E}\left(\|W_1\|^{2n}\right)$$

Recalling that $\|W_1\|^2$ is distributed according to the chi-squared distribution with r_1 degrees of freedom we have

$$\forall\gamma < 1/(2\rho(P_0)) \qquad \mathbb{E}\left(e^{\gamma\|X_0-\widehat{X}_0\|^2}\right) \leqslant \mathbb{E}\left(e^{\gamma\rho(P_0)\|W_1\|^2}\right) = (1-2\gamma\rho(P_0))^{-r_1/2} < \infty$$

Using the fact that

$$-t-\frac{1}{2}\log(1-2t) = t^2\sum_{n\geqslant 0}\frac{2}{2+n}\ (2t)^n \leqslant \frac{t^2}{1-2t} \Rightarrow (1-2t)^{-1/2} \leqslant \exp\left(t+\frac{t^2}{1-2t}\right)$$

for any $0<t<1/2$, we check that

$$\forall 0<t<(1-r_1\epsilon)/2 \qquad t+\frac{t^2}{1-2t} = t\ \frac{1-t}{1-2t} \leqslant \frac{t}{r_1\epsilon}$$

for any $\epsilon\in]0,1/r_1[$. This yields

$$\forall 0<\gamma<(1-r_1\epsilon)/(2\rho(P_0)) \qquad \mathbb{E}\left(e^{\gamma\|X_0-\widehat{X}_0\|^2}\right) \leqslant \exp\left(\rho(P_0)\gamma/\epsilon\right)$$

Choosing $\gamma=(1-2r_1\epsilon)/(2\rho(P_0))$, with $\epsilon\in]0,1/(2r_1)[$ we find that

$$\forall\epsilon\in]0,1/(2r_1)[\qquad \mathbb{E}\left(\exp\left[\left(\frac{1}{2}-(r_1\epsilon)\right)\frac{\|X_0-\widehat{X}_0\|^2}{\rho(P_0)}\right]\right) \leqslant \exp\left(\left(\frac{1}{2}-r_1\epsilon\right)\frac{1}{\epsilon}\right)$$

We check (4) by choosing

$$\epsilon = \frac{2r_1-1}{4r_1^2} = \frac{1}{2r_1}-\frac{1}{4r_1^2} \Rightarrow \mathbb{E}\left(\exp\left[\frac{\|X_0-\widehat{X}_0\|^2}{4r_1\rho(P_0)}\right]\right) \leqslant \exp\left(\frac{r_1}{2r_1-1}\right) \leqslant e$$

■

More generally, for any non negative random variable Z such that

$$\mathbb{E}\left(Z^{2n}\right)^{1/(2n)} \leqslant z\,\sqrt{n} \quad \text{for some parameter } z \geqslant 0$$

and for any $n \geqslant 1$ we have

$$\mathbb{E}\left(Z^{2n}\right) \leqslant (z^2 n)^n \leqslant \frac{e}{\sqrt{2}}\ \left(\frac{e}{2}\ z^2\right)^n \mathbb{E}(V^{2n})$$

for some Gaussian and centre random variable V with unit variance. We check this claim using Stirling approximation

$$\begin{aligned} \mathbb{E}(V^{2n}) &= 2^{-n}\frac{(2n)!}{n!} \\ &\geqslant e^{-1}\ 2^{-n}\frac{\sqrt{4\pi n}\ (2n)^{2n}\ e^{-2n}}{\sqrt{2\pi n}\ n^n\ e^{-n}} = \sqrt{2}e^{-1}\ \left(\frac{2}{e}\right)^n\ \ n^n \end{aligned} \tag{25}$$

By proposition 11.6.6 in [7], the probability of the following event

$$(Z/z)^2 \leqslant \frac{e^2}{\sqrt{2}}\ \left[\frac{1}{2} + \left(\delta + \sqrt{\delta}\right)\right] \tag{26}$$

is greater than $1 - e^{-\delta}$, for any $\delta \geqslant 0$

The above estimate also implies that

$$\mathbb{E}\left(\exp\left(tZ^2\right)\right) \leqslant \frac{e}{\sqrt{2}}\ \mathbb{E}\left(\exp\left(t(z)V^2\right)\right) = \frac{e}{\sqrt{2}}\ \frac{1}{\sqrt{1-2t(z)}} \quad \text{with} \quad t(z) = \frac{e}{2}\ z^2 t < 1/2$$

from which we check that

$$\begin{aligned} t(z) \leqslant (1-\epsilon)/2 \Rightarrow\ \mathbb{E}\left(\exp\left(tZ^2\right)\right) &\leqslant \tfrac{e}{\sqrt{2}}\ \exp\left(t(z)\left[1+\tfrac{t(z)}{1-2t(z)}\right]\right) \\ &\leqslant \tfrac{e}{\sqrt{2}}\ \exp\left(t(z)\left[1+\tfrac{1}{2}\left(\tfrac{1}{\epsilon}-1\right)\right]\right) = \tfrac{e}{\sqrt{2}}\ \exp\left(\tfrac{t(z)}{2}\left[1+\tfrac{1}{\epsilon}\right]\right) \end{aligned}$$

In summary we have

$$t \leqslant (1-\epsilon)/(z^2 e) \Rightarrow \mathbb{E}\left(\exp\left(tZ^2\right)\right) \leqslant \frac{e}{\sqrt{2}}\ \exp\left(\frac{e}{4}\ z^2 t\left[1+\frac{1}{\epsilon}\right]\right)$$

Choosing $t = (1-\epsilon)/(z^2 e)$ we conclude that

$$\forall \epsilon \in]0,1] \qquad \mathbb{E}\left(\exp\left[\frac{1-\epsilon}{e}\ \left(\frac{Z}{z}\right)^2\right]\right) \leqslant \frac{e}{\sqrt{2}}\ \exp\left(\frac{1}{4}\ \frac{1-\epsilon^2}{\epsilon}\right)$$

5.2 Proof of theorem 2.1

When $\mathcal{U}_t = \mathcal{V}_t = 0$ we have

$$d\|\mathcal{X}_t\|^2 \leqslant \rho(\mathcal{A}_t)\ \|\mathcal{X}_t\|^2\ dt + \sqrt{\mathcal{W}_t}\ \|\mathcal{X}_t\|^2\ d\mathcal{N}_t$$

with the martingale

$$d\mathcal{N}_t := \frac{1}{\sqrt{\mathcal{W}_t}\ \|\mathcal{X}_t\|^2}\ d\mathcal{M}_t \implies \partial_t\langle\mathcal{N}\rangle_t \leqslant 1$$

This implies that

$$\begin{aligned} d\log\|\mathcal{X}_t\|^2 &\leqslant \|\mathcal{X}_t\|^{-2}\ \left(\rho(\mathcal{A}_t)\ \|\mathcal{X}_t\|^2\ dt + \sqrt{\mathcal{W}_t}\ \|\mathcal{X}_t\|^2\ d\mathcal{N}_t\right) - \frac{1}{2}\ \mathcal{W}_t\ dt \\ &= \left(\rho(\mathcal{A}_t) - \frac{1}{2}\ \mathcal{W}_t\right)\ dt + \sqrt{\mathcal{W}_t}\ d\mathcal{N}_t \end{aligned}$$

from which we prove that

$$\exp\left(-\int_0^t \rho(\mathcal{A}_s)\ ds\right)\ \|\mathcal{X}_t\|^2 \leqslant \|\mathcal{X}_0\|^2\ \mathcal{E}_t$$

with the exponential martingale

$$\mathcal{E}_t := \exp\left(\int_0^t \sqrt{\mathcal{W}_s}\ d\mathcal{N}_s - \frac{1}{2}\ \int_0^t \mathcal{W}_s\ ds\right)$$

Next we provide a proof of the second assertion based on the above formula. For any $n \geqslant 0$, we observe that

$$\exp\left(-n\int_0^t \rho(\mathcal{A}_s)\ ds\right)\ \|\mathcal{X}_t\|^{2n} \leqslant \exp\left(\frac{n(n-1)}{2}\ \int_0^t \mathcal{W}_s\ ds\right)\ \|\mathcal{X}_0\|^{2n}\ \mathcal{E}_t(n)$$

with the collection of exponential martingales

$$\mathcal{E}_t(n) := \exp\left(n\int_0^t \sqrt{\mathcal{W}_s}\ d\mathcal{N}_s - \frac{n^2}{2}\ \int_0^t \mathcal{W}_s\ ds\right)$$

This implies that

$$\mathbb{E}\left(\exp\left(-n\ \int_0^t \left(\rho(\mathcal{A}_s) + \frac{(n-1)}{2}\ \mathcal{W}_s\right)\ ds\right)\ \|\mathcal{X}_t\|^{2n} \mid \mathcal{F}_0\right) \leqslant \|\mathcal{X}_0\|^{2n}$$

Arguing as above we use the decomposition

$$\begin{aligned} \mathbb{E}\left(\|\mathcal{X}_t\|^n \mid \mathcal{F}_0\right) &= \mathbb{E}\left(\exp\left(\frac{n}{2}\ \int_0^t \left(\rho(\mathcal{A}_s) + \frac{(n-1)}{2}\ \mathcal{W}_s\right)\ ds\right)\right. \\ &\qquad \left.\exp\left(-\frac{n}{2}\ \int_0^t \left(\rho(\mathcal{A}_s) + \frac{(n-1)}{2}\ \mathcal{W}_s\right)\ ds\right)\ \|\mathcal{X}_t\|^n \mid \mathcal{F}_0\right) \end{aligned}$$

to check that

$$\mathbb{E}\left(\|\mathcal{X}_t\|^n \mid \mathcal{F}_0\right) \leqslant \mathbb{E}\left(\exp\left(n\ \int_0^t \left(\rho(\mathcal{A}_s) + \frac{(n-1)}{2}\ \mathcal{W}_s\right)\ ds\right)\ \mid \mathcal{F}_0\right)^{1/2}\ \|\mathcal{X}_0\|^n$$

This ends the proof of the first assertion

More generally, we have

$$d\|\mathcal{X}_t\|^2 \leqslant \left[\rho(\mathcal{A}_t)\ \|\mathcal{X}_t\|^2 + \mathcal{U}_t\right]\ dt + d\mathcal{M}_t$$

This yields

$$\begin{aligned}
& d\|\mathcal{X}_t\|^{2n} \\
& \leqslant n\ \|\mathcal{X}_t\|^{2(n-1)}\ d\|\mathcal{X}_t\|^2 + \tfrac{n(n-1)}{2}\ \|\mathcal{X}_t\|^{2(n-2)}\ \left[\mathcal{W}_t\ \|\mathcal{X}_t\|^4 + \mathcal{V}_t\ \|\mathcal{X}_t\|^2\right]\ dt \\
& \leqslant -\Lambda_n(\mathcal{A}_t,\mathcal{W}_t)\ \|\mathcal{X}_t\|^{2n}dt + n\ \left[\tfrac{(n-1)}{2}\ \mathcal{V}_t\ + \mathcal{U}_t\ \right]\ \|\mathcal{X}_t\|^{2(n-1)}\ dt + n\ \|\mathcal{X}_t\|^{2(n-1)}\ d\mathcal{M}_t
\end{aligned}$$

with

$$-\Lambda_n(\mathcal{A}_t,\mathcal{W}_t) := \ n\ \rho(\mathcal{A}_t)\ + \frac{n(n-1)}{2}\ \mathcal{W}_t$$

Observe that

$$\begin{aligned}
\Lambda_n(\mathcal{A}_t,\mathcal{W}_t) - \Lambda_{n-1}(\mathcal{A}_t,\mathcal{W}_t) &= -n\ \rho(\mathcal{A}_t)\ - n(n-1)\ \mathcal{W}_t/2 \\
&\qquad +(n-1)\ \rho(\mathcal{A}_t)\ + (n-1)(n-2)\ \mathcal{W}_t/2 \\
&= -\rho(\mathcal{A}_t) - (n-1)\mathcal{W}_t
\end{aligned}$$

This shows that

$$\overline{\mathcal{U}}_t/\mathcal{U}_t = \exp\left(\int_0^t \left[\Lambda_n(\mathcal{A}_s,\mathcal{W}_s) - \Lambda_{n-1}(\mathcal{A}_s,\mathcal{W}_s)\right] ds\right) = \overline{\mathcal{V}}_t/\mathcal{V}_t$$

We set

$$\mathcal{Y}^n_t := \exp\left(\int_0^t \Lambda_n(\mathcal{A}_s,\mathcal{W}_s)ds\right)\ \|\mathcal{X}_t\|^{2n}$$

Notice that

$$\begin{aligned}
& \exp\left(\int_0^t \Lambda_n(\mathcal{A}_s,\mathcal{W}_s)ds\right)\ \mathcal{U}_t\ \|\mathcal{X}_t\|^{2(n-1)} \\
& = \|\mathcal{X}_t\|^{2(n-1)}\ \exp\left(\int_0^t \Lambda_{n-1}(\mathcal{A}_s,\mathcal{W}_s)ds\right) \\
& \qquad \times \exp\left(\int_0^t \left[\Lambda_n(\mathcal{A}_s,\mathcal{W}_s) - \Lambda_{n-1}(\mathcal{A}_s,\mathcal{W}_s)\right] ds\right)\ \mathcal{U}_t = \mathcal{Y}^{n-1}_t\overline{\mathcal{U}}_t
\end{aligned}$$

This shows that

$$d\mathcal{Y}^n_t \leqslant n\ \mathcal{Y}^{n-1}_t\ \left[\ \overline{\mathcal{U}}_t + \frac{(n-1)}{2}\ \overline{\mathcal{V}}_t\ \right]\ dt + d\overline{\mathcal{M}}_t$$

with the martingale

$$d\overline{\mathcal{M}}_t := n\ \exp\left(\int_0^t \Lambda_n(\mathcal{A}_s,\mathcal{W}_s)ds\right)\ \|\mathcal{X}_t\|^{2(n-1)}\ d\mathcal{M}_t$$

This implies that

$$\partial_t \mathbb{E}\left(\mathcal{Y}_t^n \mid \mathcal{F}_0\right) \leqslant n\ \mathbb{E}\left(\mathcal{Y}_t^{n-1}\ \left[\ \overline{\mathcal{U}}_t + \frac{(n-1)}{2}\ \overline{\mathcal{V}}_t\ \right] \mid \mathcal{F}_0\right)$$

Using Hölder inequality we have

$$\mathbb{E}\left(\ \overline{\mathcal{U}}_t\ \mathcal{Y}_t^{n-1} \mid \mathcal{F}_0\right) \leqslant \mathbb{E}\left(\overline{\mathcal{U}}_t^n \mid \mathcal{F}_0\right)^{1/n} \mathbb{E}\left(\mathcal{Y}_t^n \mid \mathcal{F}_0\right)^{1-1/n}$$

This yields the estimate

$$\partial_t \mathbb{E}\left(\mathcal{Y}_t^n \mid \mathcal{F}_0\right) \leqslant n\ \mathbb{E}\left(\mathcal{Y}_t^n \mid \mathcal{F}_0\right)^{1-1/n}\ \left[\ \mathbb{E}\left(\overline{\mathcal{U}}_t^n \mid \mathcal{F}_0\right)^{1/n} + \frac{(n-1)}{2}\ \mathbb{E}\left(\overline{\mathcal{V}}_t^n \mid \mathcal{F}_0\right)^{1/n}\ \right]$$

and therefore

$$\begin{aligned}\partial_t \mathbb{E}\left(\mathcal{Y}_t^n \mid \mathcal{F}_0\right)^{1/n} &= \frac{1}{n}\ \mathbb{E}\left(\mathcal{Y}_t^n \mid \mathcal{F}_0\right)^{-(1-1/n)}\ \partial_t \mathbb{E}\left(\mathcal{Y}_t^n \mid \mathcal{F}_0\right) \\ &\leqslant \mathbb{E}\left(\overline{\mathcal{U}}_t^n \mid \mathcal{F}_0\right)^{1/n} + \frac{(n-1)}{2}\ \mathbb{E}\left(\overline{\mathcal{V}}_t^n \mid \mathcal{F}_0\right)^{1/n}\end{aligned}$$

We conclude that

$$\mathbb{E}\left(\mathcal{Y}_t^n \mid \mathcal{F}_0\right)^{1/n} \leqslant \mathbb{E}\left(\mathcal{Y}_0^n \mid \mathcal{F}_0\right)^{1/n} + \int_0^t\ \left[\mathbb{E}\left(\overline{\mathcal{U}}_s^n \mid \mathcal{F}_0\right)^{1/n} + \frac{(n-1)}{2}\ \mathbb{E}\left(\overline{\mathcal{V}}_s^n \mid \mathcal{F}_0\right)^{1/n}\right]\ ds$$

Using the decomposition

$$\mathbb{E}\left(\|\mathcal{X}_t\|^n \mid \mathcal{F}_0\right) = \mathbb{E}\left(\exp\left(-\frac{1}{2}\int_0^t \Lambda_n(\mathcal{A}_s, \mathcal{W}_s)ds\right)\ \ \exp\left(\frac{1}{2}\int_0^t \Lambda_n(\mathcal{A}_s, \mathcal{W}_s)ds\right)\ \|\mathcal{X}_t\|^n \mid \mathcal{F}_0\right)$$

and Cauchy-Schwartz inequality we check that

$$\mathbb{E}\left(\|\mathcal{X}_t\|^n \mid \mathcal{F}_0\right)^{2/n} \leqslant \mathbb{E}\left(\exp\left(-\int_0^t \Lambda_n(\mathcal{A}_s, \mathcal{W}_s)ds\right) \mid \mathcal{F}_0\right)^{1/n}\ \mathbb{E}\left(\mathcal{Y}_t^n \mid \mathcal{F}_0\right)^{1/n}$$

This implies that

$$\begin{aligned}\mathbb{E}\left(\|\mathcal{X}_t\|^n \mid \mathcal{F}_0\right)^{2/n} \leqslant\ & \mathbb{E}\left(\exp\left(-\int_0^t \Lambda_n(\mathcal{A}_s, \mathcal{W}_s)ds\right)\right)^{1/n} \\ & \times \left[\|\mathcal{X}_0\|^2 + \int_0^t\ \left[\mathbb{E}\left(\overline{\mathcal{U}}_s^n \mid \mathcal{F}_0\right)^{1/n} + \frac{(n-1)}{2}\ \mathbb{E}\left(\overline{\mathcal{V}}_s^n \mid \mathcal{F}_0\right)^{1/n}\right]\ ds\right]\end{aligned}$$

This ends the proof of the theorem. ■

5.3 Proof of proposition 2.3

By (15), for any $m \geqslant 1$ we have the uniform estimate

$$\mathbb{E}\left(\|\mathcal{X}_t\|^{2m}\right)^{1/m} \leqslant (u_t(a) + m v_t(a)) \Rightarrow 2\ \mathbb{E}\left(\|\mathcal{X}_t\|^{2m}\right) \leqslant\ (2u_t(a))^m + (2v_t(a))^m\ m^m$$

Using Stirling approximation (25) we have

$$(2\gamma\ v_t(a))^m m^m \leqslant \frac{e}{\sqrt{2}}\ (e\gamma v_t(a))^m\ 2^{-m}\frac{(2m)!}{m!} \Rightarrow \sum_{m\geqslant 0}(2v_t(a))^m\ m^m \leqslant \frac{e}{\sqrt{2}}\ \frac{1}{\sqrt{1-2ev_t(a)\gamma}}$$

for any $\gamma < 1/(2ev_t(a))$. This yields

$$2\ \mathbb{E}\left(e^{\gamma\|\mathcal{X}_t\|^2}\right) \leqslant e^{2\gamma u_t(a)} + \frac{e}{\sqrt{2}}\ \frac{1}{\sqrt{1-2e\gamma v_t(a)}}$$

Choosing $\gamma = (1-\epsilon)/(2ev_t(a))$, with $\epsilon \in]0,1]$ we find that

$$\mathbb{E}\left(\exp\left[\frac{(1-\epsilon)}{e}\ \frac{1}{2v_t(a)}\ \|\mathcal{X}_t\|^2\right]\right) \leqslant \frac{1}{2}\ e^{\frac{1-\epsilon}{e}\frac{u_t(a)}{v_t(a)}} + \frac{e}{2\sqrt{2}}\ \frac{1}{\sqrt{\epsilon}}$$

This ends the proof of (16). Now we come to the proof of (17).
We have

$$d\|\mathcal{X}_t\|^2 \leqslant \left[-a\ \|\mathcal{X}_t\|^2 + \mathcal{U}_t\right]\ dt + d\mathcal{M}_t$$

This implies that

$$\int_0^t\ \|\mathcal{X}_s\|^2\ ds \leqslant \int_0^t\ e^{-as}\ \left[\int_0^s\ e^{au}\ \mathcal{U}_u\ du\right]\ ds + \int_0^t\ e^{-as}\ \left[\int_0^s\ e^{au}\ d\mathcal{M}_u\right]\ ds$$

On the other hand, by an integration by part we have

$$a\int_0^t\ e^{-as}\left(\int_0^s\ e^{au}\ \mathcal{U}_u\ du\right)ds = \int_0^t\left(1-e^{-a(t-s)}\right)\ \mathcal{U}_s\ ds$$

and

$$a\int_0^t\ e^{-as}\left(\int_0^s\ e^{au}\ d\mathcal{M}_u\right)ds = \int_0^t\left(1-e^{-a(t-s)}\right)\ \ d\mathcal{M}_s$$

This implies that

$$a\ \int_0^t\ \|\mathcal{X}_s\|^2\ ds \leqslant \int_0^t \mathcal{U}_s\ ds + \overline{\mathcal{M}}_t^{(t)}$$

with the terminal state $\overline{\mathcal{M}}_t^{(t)}$ of the collection of martingales $\overline{\mathcal{M}}_u^{(t)}$ on $[0,t]$ defined by

$$\forall 0 \leqslant u \leqslant t \qquad \overline{\mathcal{M}}_u^{(t)} := \int_0^u\left(1-e^{-a(t-s)}\right)\ \ d\mathcal{M}_s$$

$$\Longrightarrow \quad \partial_u\langle\overline{\mathcal{M}}^{(t)}\rangle_u \leqslant v\ \|\mathcal{X}_u\|^2$$

Therefore for any $\gamma \geqslant 0$ we have

$$\gamma\ \left[\left(a-\frac{\gamma}{2}\ v\right)\ \int_0^t\ \|\mathcal{X}_s\|^2\ ds-\int_0^t\mathcal{U}_s\ ds\right]\leqslant\gamma\overline{\mathcal{M}}_t^{(t)}-\frac{\gamma^2}{2}\langle\overline{\mathcal{M}}^{(t)}\rangle_t$$

This implies that

$$\mathbb{E}\left(\exp\left[\gamma\ \left[\left(a-\frac{\gamma}{2}\ v\right)\ \int_0^t\ \|\mathcal{X}_s\|^2\ ds-\int_0^t\mathcal{U}_s\ ds\right]\right]\mid\mathcal{F}_0\right)\leqslant 1$$

Using the decomposition

$$\exp\left[\frac{\gamma}{2}\ \left(a-\frac{\gamma}{2}\ v\right)\ \int_0^t\ \|\mathcal{X}_s\|^2\ ds\right]$$

$$=\exp\left[\frac{\gamma}{2}\ \left[\left(a-\frac{\gamma}{2}\ v\right)\ \int_0^t\ \|\mathcal{X}_s\|^2\ ds-\int_0^t\mathcal{U}_s\ ds\right]\right]\times\exp\left[\frac{\gamma}{2}\ \int_0^t\mathcal{U}_s\ ds\right]$$

Replacing $\gamma/2$ by γ, by Cauchy-Schwartz inequality we find that

$$\mathbb{E}\left(\exp\left[v\ \gamma\ \left(\frac{a}{v}-\gamma\right)\ \int_0^t\ \|\mathcal{X}_s\|^2\ ds\right]\right)\leqslant\mathbb{E}\left(\exp\left[2\gamma\int_0^t\mathcal{U}_s\ ds\right]\right)^{1/2}$$

for any $\gamma\leqslant a/v$. Observe that

$$\gamma\ \left(\frac{a}{v}-\gamma\right):=\frac{\alpha}{v}\leqslant c^2\quad\text{with}\quad c=\frac{a}{v}$$

We also have

$$\gamma\ \left(\frac{a}{v}-\gamma\right)=\frac{\alpha}{v}\Longleftrightarrow\gamma\in\left\{c/2-\sqrt{(c/2)^2-\alpha/v},c/2+\sqrt{(c/2)^2-\alpha/v}\right\}$$

Choosing the smallest value we prove that

$$\begin{aligned}\mathbb{E}\left(\exp\left[v\ \alpha/v\ \int_0^t\ \|\mathcal{X}_s\|^2\ ds\right]\right)^2\ &\leqslant\ \mathbb{E}\left(\exp\left[\left((a/v)-\sqrt{(a/v)^2-4\alpha/v}\right)\int_0^t\mathcal{U}_s\ ds\right]\right)\\ &=\ \mathbb{E}\left(\exp\left[\frac{4\alpha/v}{(a/v)+\sqrt{(a/v)^2-4\alpha/v}}\ \int_0^t\mathcal{U}_s\ ds\right]\right)\end{aligned}$$

for any $\beta=\alpha/v\leqslant a^2/(2v)^2$, or equivalently

$$\mathbb{E}\left(\exp\left[v\ \beta\ \int_0^t\ \|\mathcal{X}_s\|^2\ ds\right]\right)\ \leqslant\ \mathbb{E}\left(\exp\left[\frac{4\beta}{(a/v)+\sqrt{(a/v)^2-4\alpha/v}}\ \int_0^t\mathcal{U}_s\ ds\right]\right)^{1/2}$$

This ends the proof of the proposition. ■

References

[1] J.S. Baras, A. Bensoussan, and M.R. James. Dynamic observers as asymptotic limits of recursive filters: special cases. SIAM J. Appl. Math., 48:1147–1158, 1988.

[2] M. Boutayeb, H. Rafaralahy, and M. Darouach. Convergence analysis of the extended Kalman filter used as an observer for nonlinear deterministic discrete-time systems. IEEE transactions on automatic control, 42, 1997.

[3] S. Bonnabel, J.J. Slotine. A contraction theory-based analysis of the stability of the Extended Kalman Filter Arxiv. 1211.6624 dec. (2012).

[4] R. S. Bucy, P. D. Joseph. Filtering for Stochastic Processes with Applications to Guidance, John Wiley & Sons (1968); 2nd Edition, AMS Chelsea Publ. (2005).

[5] R. S. Bucy, R. E Kalman. New Results in Linear Filtering and Prediction Theory, Transactions of the ASME - Journal of Basic Engineering, vol. 83, pp. 95–107 (1961).

[6] P. Del Moral, J. Tugaut. On the stability and the uniform propagation of chaos properties of Ensemble Kalman-Bucy filters, Arxiv (2016).

[7] P. Del Moral. Mean field simulation for Monte Carlo integration. Chapman & Hall/CRC Monographs on Statistics & Applied Probability (2013).

[8] V. Karvonen. Stability of linear and non-linear Kalman filters Master thesis, University of Helsinki (2014).

[9] R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems. Journal of Basic Engineering vol. 82, no. 35 (1960).

[10] A. J. Krener. The convergence of the extended Kalman filter. /em In Rantzer, A. and Byrnes, C. I., editors. Directions in Mathematical Systems Theory and Optimization. vol. 286, Lecture Notes in Control and Information Sciences, chapter 12, pp. 173–182. Springer-Verlag Berlin Heidelberg (2003).

[11] W. Lohmiller and J.J.E. Slotine. On contraction analysis for nonlinear systems. Automatica, vol. 34, no. 6, pp. 683–696 (1998).

[12] W. Lohmiller, J.J.E. Slotine. Contraction Analysis of Nonlinear Distributed Systems. International Journal of Control, vol. 78, no. 9, (2005).

[13] D. Luenberger. An introduction to observers. IEEE Trans. on Automatic Control, vol. 16, no. 6, pp. 596–602 (1971).

[14] K. Reif, S. Gunther ; E. Yaz, R. Unbehauen. Stochastic stability of the discrete-time extended Kalman filter. IEEE Transactions on Automatic Control, vol. 44, no. 4, pp. 714–728 (1999).

[15] K. Reif, S. Gunther ; E. Yaz, R. Unbehauen. Stochastic stability of the continuous-time extended Kalman filter. IEE Proceedings - Control Theory and Applications, vol. 147, no. 1, pp. 45-52 (2000).

[16] L. Ozbek, E. Koksal Babacan , M. Efe. Stochastic stability of the discrete-time constrained extended Kalman filter Turk J Elec Eng & Comp Sci, Vol.18, No.2 (2010).

[17] M. B. Rhudy and Y. Gu. Online Stochastic Convergence Analysis of the Kalman Filter. Hindawi Publishing Corporation International Journal of Stochastic Analysis Volume 2013, Article ID 240295 (2013).

[18] D. Simon. Optimal state estimation: Kalman, H-infinity, and nonlinear approaches. John Wiley & Sons (2006).

[19] F. Sonnemann Reif, K. and R. Unbehauen. An EKF-based nonlinear observer with a prescribed degree of stability. Automatica, vol. 34, pp. 1119–1123, (1998).

[20] Y. K. Song, J.W. Grizzle. The extended Kalman filter as a local asymptotic observer. Estimation and Control, vol. 5, pp. 59–78 (1995).

Pierre DEL MORAL
INRIA Bordeaux Research Center (France) & UNSW School of Mathematics and Statistics (Australia)
p.del-moral@unsw.edu.au

Aline KURTZMANN
Université de Lorraine, Institut Elie Cartan de Lorraine
CNRS, Institut Elie Cartan de Lorraine, UMR 7502, Vandoeuvre-lès-Nancy, F-54506, France.
aline.kurtzmann@univ-lorraine.fr

Julian TUGAUT
Univ Lyon, Université Jean Monnet, CNRS UMR 5208, Institut Camille Jordan
Maison de l'Université, 10 rue Tréfilerie, CS 82301, 42023 Saint-Etienne Cedex 2, France
tugaut@math.cnrs.fr