On the stability and the uniform propagation of chaos of a Class of Extended Ensemble Kalman-Bucy filters

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

September 18, 2016

Abstract

This article is concerned with the exponential stability and the uniform propagation of chaos properties of a class of Extended Ensemble Kalman-Bucy filters with respect to the time horizon. This class of nonlinear filters can be interpreted as the conditional expectations of nonlinear McKean-Vlasov type diffusions with respect to the observation process. We consider filtering problems with Langevin type signal processes observed by some noisy linear and Gaussian type sensors. In contrast with more conventional Langevin nonlinear drift type processes, the mean field interaction is encapsulated in the covariance matrix of the diffusion. The main results discussed in the article are quantitative estimates of the exponential stability properties of these nonlinear diffusions. These stability properties are used to derive uniform and non asymptotic estimates of the propagation of chaos properties of Extended Ensemble Kalman filters, including exponential concentration inequalities. To our knowledge these results seem to be the first results of this type for this class of nonlinear ensemble type Kalman-Bucy filters.

Keywords : Extended Kalman-Bucy filter, Ensemble Kalman filters, Monte Carlo methods, mean field particle systems, stochastic Riccati matrix equation, propagation of chaos properties, uniform estimates.

Mathematics Subject Classification : 60J60, 60J22, 35Q84, 93E11, 60M20, 60G25.

1 Introduction

From the probabilistic viewpoint, the Ensemble Kalman filter (*abbreviated EnKF*) proposed by G. Evensen in the beginning of the 1990s [15] is a mean field particle interpretation of extended Kalman type filters. More precisely, Kalman type filters (including the conventional Kalman filter and extended Kalman filters) can be interpreted as the conditional expectations of a McKean-Vlasov type nonlinear diffusion. The key idea is to approximate the Riccati equation by a sequence of sample covariance matrices associated with a series of interacting Kalman type filters.

In the linear Gaussian case these particle type filters converge to the optimal Kalman filter as the number of samples (a.k.a. particles) tends to ∞. Little is known for nonlinear and/or non Gaussian filtering problems, apart that they do not converge to the desired optimal filter. This important problem is rather well known in signal processing community. For instance, we refer the reader to [21, 23] for a more detailed discussion on these questions in discrete time settings. In this connection, we mention that these ensemble Kalman type filters differ from interacting jump type particle filters and related sequential Monte Carlo methodologies. These mean field particle methods are designed to approximate the conditional distributions of the signal given the observations. It is clearly not the scope of this article to give a comparison between these two different particle methods. For a more thorough discussion we refer the reader to the book [12] and the references therein. We also mention that the EnKF models discussed in this article slightly differ from more conventional EnKF used to approximate nonlinear filtering problems. To be more precise we design a new class of EnKF that converges to the celebrated extended Kalman filter as the number of particles goes to ∞.

These powerful Monte Carlo methodologies are used with success in a variety of scientific disciplines, and more particularly in data assimilation method for filtering high dimensional problems arising in fluid mechanics and geophysical sciences [25, 26, 27, 29, 31, 32, 34, 35, 37]. A more thorough discussion on the origins and the application domains of EnKF is provided in the series of articles [5, 13, 16, 18] and in the seminal research monograph by G. Evensen [17].

The mathematical foundations and the convergence of the EnKF have started in 2011 with the independent pioneering works of F. Le Gland, V. Monbet and V.D. Tran [23], and the one by J. Mandel, L. Cobb, J. D. Beezley [29]. These articles provide $\mathbb{L}_\delta$-mean error estimates for discrete time EnKF and show that they converge towards the Kalman filter as the number of samples tends to infinity. We also quote the recent article by D.T. B. Kelly, K.J. Law, A. M. Stuart [21] showing the consistency of Ensemble Kalman filters in continuous and discrete time settings. In the latter the authors show that the Ensemble Kalman filter is well-posed and the mean error variance does not blow up faster than exponentially. The authors also apply a judicious variance inflation technique to strengthen the contraction properties of the Ensemble Kalman filter. We refer to the pioneering article by J.L. Anderson [1, 2, 3] on adaptive covariance inflation techniques, and to the discussion given in the end of Section 2 in the present article.

In a more recent study by X. T. Tong, A. J. Majda and D. Kelly [36] the authors analyze the long-time behaviour and the ergodicity of discrete generation EnKF using Foster-Lyapunov techniques ensuring that the filter is asymptotically stable w.r.t. any erroneous initial condition. These important properties ensure that the EnKF has a single invariant measure and initialization errors of the EnKF will not dissipate w.r.t. the time parameter. Beside the importance of these properties, the only ergodicity of the particle process does not give any information on the convergence and the accuracy of the particle filters towards the optimal filter nor towards any type of extended Kalman filter, as the number of samples tends to infinity.

Besides these recent theoretical advances, the rigorous mathematical analysis of long time behaviour of these particle methods is still at its infancy. As underlined by the authors in [21], many of the algorithmic innovations associated with the filter, which are required to make a useable algorithm in practice, are derived in an *ad hoc* fashion. The divergence of ensemble Kalman filters has been observed numerically in some situations [20, 22, 28], even for stable signals. This critical phenomenon, often referred as the *catastrophic filter divergence* in data assimilation literature, is poorly understood from the mathematical perspective. Our objective is to better understand the long time behaviour of ensemble Kalman type filters from a mathematical perspective. Our stochastic methodology combines spectral analysis of random matrices with recent developments in concentration inequalities, coupling theory and contraction inequalities w.r.t. Wasserstein metrics.

These developments have been started in two recent articles [13, 14]. The first one provides uniform propagation of chaos properties of ensemble Kalman filters in the context of linear-Gaussian filtering problems. The second article is only concerned with extended Kalman-Bucy filters. It discusses the stability properties of these filters in terms of exponential concentration inequalities. These concentration inequalities allow to design confidence intervals around the true signal and extended Kalman-Bucy filters. Following these studies, we consider filtering problems with uniformly stable signal processes.

This condition on the signal is a necessary and sufficient condition to derive uniform estimates for any type of particle filters [11, 12, 13] w.r.t. the time parameter. For instance when the sensor matrix is null or for a single particle any Ensemble type type filter reduces to an independent copy of the signal. In these rather elementary cases, the stability of the signal is required to have any type of uniform estimate for any size of the systems.

We illustrate these models and our stability and observability conditions with a class of nonlinear Langevin type filtering problems, with several classes of sensor models

The first contribution of the article is to extend these results as the level of the McKean-Vlasov type nonlinear diffusion associated with the ensemble Kalman-Bucy filter. Under some natural regularity conditions we show that these nonlinear diffusions are exponentially stable, in the sense that they forget exponentially fast any erroneous initial condition. These stability properties are analyzed using coupling techniques and expressed in terms of δ-Wasserstein metrics.

The main objective of the article is to analyze the long-time behaviour of the mean field particle interpretation of these nonlinear diffusions. We present new uniform estimates w.r.t. the time horizon for the bias and the propagation of chaos properties of the mean field systems. We also quantify the fluctuations of the sample mean and covariance particle approximations.

The rest of the article is organized as follows:

Section 1.2 presents the nonlinear filtering problem discussed in the article, the Extended Kalman-Bucy filter, the associated nonlinear McKean-Vlasov diffusion and its mean field particle interpretation. The two main theorems of the article are described in Section 2. In a preliminary short section, Section 3, we show that the conditional expectations and the conditional covariance matrices of the nonlinear McKean-Vlasov diffusion coincide with the EKF. We also provide a pivotal fluctuation theorem on the time evolution of these conditional statistics. Section 4 is mainly concerned with the stability properties of the nonlinear diffusion associated with the EKF. Section 5 is dedicated to the propagation of chaos properties of the extended ensemble Kalman-Bucy filter.

1.1 Some basic notation

This section provides with some notation and terminology used in several places in the article.

Given some random variable Z with some probability measure μ and some function f on some product space $\mathbb{R}^r$, we let

$$\mu(f) = \mathbb{E}(f(Z)) = \int f(x)\ \mu(dx)$$

be the integral of f w.r.t. μ or the expectation of $f(Z)$. This notation is rather standard in probability theory. It extends to integral on Euclidian state spaces the conventional vector summation notation $\mu(f) = \sum_x \mu(x)\ f(x)$ between row vector measures $\mu = (\mu(x))_{x\in E}$ and dual column vector functions $f = (f(x))_{x\in E}$ on finite state spaces $E = \{1,\ldots,d\}$, for some parameter $d \geqslant 1$.

We let $\|.\|$ be the Euclidean norm on $\mathbb{R}^r$, for some $r \geqslant 1$. We denote by $\mathbb{S}_r$ the set of $(r\times r)$ symmetric matrices with real entries, and by $\mathbb{S}_r^+$ the subset of positive definite matrices.

We denote by $\lambda_{min}(S)$ and $\lambda_{max}(S)$ the minimal and the maximal eigenvalue of a given symmetric matrix S. We let $\rho(P) = \lambda_{max}((P+P'))/2$ be the logarithmic norm of a given square matrix P. Given $(r_1\times r_2)$ matrices P, Q we define the Frobenius inner product

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

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

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

Given some $\delta \geqslant 1$, the δ-Wasserstein distance $\mathbb{W}_\delta$ between two probability measures ν_1 and ν_2 on some normed space $(E,\|.\|)$ is defined by

$$\mathbb{W}_\delta(\nu_1,\nu_2) = \inf \mathbb{E}\left(\|Z_1 - Z_2\|^\delta\right)^{1/\delta}.$$

The infimum in the above displayed formula is taken of all pair of random variable (Z_1, Z_2) such that $\text{Law}(Z_i) = \nu_i$, with $i = 1,2$.

In the further development of the article, to avoid unnecessary repetitions we also use the letter "c" to denote some finite constant whose values may vary from line to line, but they do not depend on the time parameter.

1.2 Description of the models

Consider a time homogeneous nonlinear filtering problem of the following form

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

In the above display, (W_t, V_t) is an (r_1+r_2)-dimensional Brownian motion, X_0 is a r_1-valued random vector with mean and covariance matrix $(\mathbb{E}(X_0), P_0)$ (independent of (W_t,V_t)), the square root factors $R_1^{1/2}$ and $R_2^{1/2}$ of R_1 and R_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 $\partial A\ :\ x\in\mathbb{R}^{r_1}\mapsto A(x)\in\mathbb{R}^{(r_1\times r_1)}$.

The Extended Kalman-Bucy filter (*abbreviated EKF*) and the associated stochastic Riccati equation are defined by the evolution equations

$$\left\{\begin{array}{rcl} d\widehat{X}_t &=& A(\widehat{X}_t)\ dt + P_tB'\ R_2^{-1}\left[dY_t - B\widehat{X}_t\ dt\right] \quad \text{with} \quad \widehat{X}_0 = \mathbb{E}(X_0), \\ \partial_t P_t &=& \partial A(\widehat{X}_t)P_t + P_t\ \partial A(\widehat{X}_t)' + R - P_tSP_t \quad \text{with} \quad (R,S) := (R_1, B'R_2^{-1}B). \end{array}\right. \tag{2}$$

In the above display, B' stands for the transpose of the matrix B.

We associate with these filtering models the conditional nonlinear McKean-Vlasov type diffusion process

$$d\overline{X}_t = \mathcal{A}\left(\overline{X}_t, \mathbb{E}[\overline{X}_t \mid \mathcal{G}_t]\right)\ dt + R_1^{1/2}\ d\overline{W}_t + \mathcal{P}_{\eta_t} B' R_2^{-1}\ \left[dY_t - \left(B\overline{X}_t\ dt + R_2^{1/2}\ d\overline{V}_t\right)\right] \quad (3)$$

with the nonlinear drift function

$$\mathcal{A}(x,m) := A\,[m] + \partial A\,[m]\ (x-m).$$

In the above display $(\overline{W}_t, \overline{V}_t, \overline{X}_0)$ stands for independent copies of (W_t, V_t, X_0) (thus independent of the signal and the observation path), and $\mathcal{P}_{\eta_t}$ stands for the covariance matrix

$$\mathcal{P}_{\eta_t} = \eta_t\left[(e-\eta_t(e))(e-\eta_t(e))'\right] \quad \text{with} \quad \eta_t := \text{Law}(\overline{X}_t \mid \mathcal{G}_t) \quad \text{and} \quad e(x) := x.$$

The stochastic process defined in (3) is named the Extended Kalman-Bucy diffusion or simply the EKF-diffusion. In Section 3 (see Proposition 3.1) we will see that the $\mathcal{G}_t$-conditional expectation of the states $\overline{X}_t$ and their $\mathcal{G}_t$-conditional covariance matrices coincide with the EKF filter and the Riccati equation presented in (2).

The Ensemble Extended Kalman-Bucy filter (*abbreviated En-EKF*) coincides with the mean field particle interpretation of the nonlinear diffusion process (3).

To be more precise, let $(\overline{W}^i_t, \overline{V}^i_t, \xi^i_0)_{1\leqslant i\leqslant N}$ be N independent copies of $(\overline{W}_t, \overline{V}_t, \overline{X}_0)$. In this notation, the En-EKF is given by the McKean-Vlasov type interacting diffusion process

$$d\xi^i_t = \mathcal{A}(\xi^i_t, m_t)\ dt + R_1^{1/2} d\overline{W}^i_t + p_t B' R_2^{-1}\left[dY_t - \left(B\xi^i_t\ dt + R_2^{1/2}\ d\overline{V}^i_t\right)\right] \quad (4)$$

for any $1 \leqslant i \leqslant N$, with the sample mean and the rescaled particle covariance matrix defined by

$$m_t := \frac{1}{N}\sum_{1\leqslant i\leqslant N} \xi^i_t \quad \text{and} \quad p_t := \left(1-\frac{1}{N}\right)\ \mathcal{P}_{\eta^N_t} = \frac{1}{N-1}\sum_{1\leqslant i\leqslant N}\left(\xi^i_t - m_t\right)\left(\xi^i_t - m_t\right)' \quad (5)$$

with the empirical measures $\eta^N_t := \frac{1}{N}\sum_{1\leqslant i\leqslant N}\delta_{\xi^i_t}$. We also consider the N-particle model $\zeta_t = \left(\zeta^i_t\right)_{1\leqslant i\leqslant N}$ defined as $\xi_t = \left(\xi^i_t\right)_{1\leqslant i\leqslant N}$ by replacing the sample variance p_t by the true variance P_t (in particular we have $\xi_0 = \zeta_0$).

When $B=0$ the En-EKF reduce to N independent copies of the diffusion signal. In the same vein, for a single particle the covariance matrix is null so that the En-EKF reduces to a single independent copy of the signal. In the case $r_1 = 1$ we have

$$\mathbb{E}\left(\|m_t - X_t\|^2\right) = 2\ \text{Var}(X_t) \quad (6)$$

In these rather elementary situations, *the stability property of the signal is crucial to design some useful uniform estimates w.r.t. the time parameter.* The stability of the signal is a necessary condition to derive uniform estimates for any type of particle filters [11, 12, 13] w.r.t. the time parameter.

As mentioned in the introduction the En-EKF (4) differs from the more conventional one defined as above by replacing $\mathcal{A}(\xi^i_t, m_t)$ by the signal drift $A(\xi^i_t)$. In this context the resulting sample mean will not converge to the EKF but to the filter defined as in (2) by replacing $A(\widehat{X}_t)$ by the conditional expectations $\mathbb{E}\left(A(X_t) \mid \mathcal{G}_t\right)$. The convergence analysis of this particle model is much more involved than the one discussed in this article. The main difficulty comes from the dependency on the whole conditional distribution of the signal given the observations. We plan to analyze this class of particle filters in a future study.

1.3 Regularity conditions

1.3.1 Langevin-type signal processes

In the further development of the article we assume that the Jacobian matrix of A satisfies the following regularity conditions:

$$\left\{\begin{array}{rcl} -\lambda_{\partial A} & := & \sup_{x\in\mathbb{R}^{r_1}} \rho(\partial A(x)+\partial A(x)') < 0 \\ \\ \|\partial A(x)-\partial A(y)\| & \leqslant & \kappa_{\partial A}\,\|x-y\| \quad \text{for some } \kappa_{\partial A}<\infty. \end{array}\right. \tag{7}$$

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 Euclidean distance between x and y. A Taylor first order expansion shows that

$$(7) \Longrightarrow \langle x-y, A(x)-A(y)\rangle \leqslant -\lambda_A\,\|x-y\|^2 \quad \text{with } \lambda_A \geqslant \lambda_{\partial A}/2 > 0. \tag{8}$$

The above rather strong conditions ensure the contraction needed to ensure the stability of the EFK [14]. For linear systems $A(x) = Ax$, associated with some matrix A, the parameters $\lambda_A = \lambda_{\partial A}/2$ coincide with the logarithmic norm of A. In this situation we show in [13] (section 3.1) that the above condition cannot be relaxed to derive uniform estimates of the Ensemble Kalman-Bucy filter.

The prototype of signals satisfying these conditions are multidimensional diffusions with drift functions $(A,\partial A) = (-\partial\mathcal{V}, -\partial^2\mathcal{V})$ associated with a gradient Lipschitz strongly convex confining potential $\mathcal{V}\,:\, x\in\mathbb{R}^{r_1}\mapsto\mathcal{V}(x)\in[0,\infty[$. The logarithmic norm condition (7) is met as soon as $\partial^2\mathcal{V}\geqslant v\;Id$ with $v=2|\lambda_{\partial A}|$. Equivalently the smallest eigenvalue $\lambda_{min}(\partial^2\mathcal{V}(x))$ of the Hessian is uniformly lower bounded by v. In this case (7) is met with $\lambda_{\partial A} = v/2$.

These conditions are fairly standard in the stability theory of nonlinear diffusions, we refer the reader to the review article [30], and the references therein. Choosing $R_1 = \sigma_1^2\;Id$ and $A = -\beta\partial\mathcal{V}$, for some $\beta,\sigma_1\geqslant 0$ the signal process X_t resumes to a multidimensional Langevin-diffusion

$$dX_t = -\beta\;\partial\mathcal{V}(X_t)\;dt + \sigma_1\;dW_t \tag{9}$$

This process is reversible w.r.t. the invariant distribution. Let μ be a probability distribution on $\mathbb{R}^{r_1}$ given by

$$\mu_\beta(dx) = \frac{1}{\mathcal{Z}_\beta}\;\exp\left(-\frac{2\beta}{\sigma_1^2}\mathcal{V}(x)\right)\,dx \quad \text{with} \quad \mathcal{Z}_\beta = \int\;\exp\left(-\frac{2\beta}{\sigma_1^2}\mathcal{V}(x)\right)\,dx \in]0,\infty[.$$

In the above display dx stands for the Lebesgue measure on $\mathbb{R}^{r_1}$. The Lipschitz-continuity condition of the Hessian $\partial^2\mathcal{V}$ introduced in (7) ensures the continuity of the stochastic Riccati equation (2) w.r.t. the fluctuations around the random states $\widehat{X}_t$. We illustrate this condition with a nonlinear example given by the function

$$\mathcal{V}(x) = \frac{1}{2}\;\langle\mathcal{Q}_1 x,x\rangle + \langle q,x\rangle + \frac{1}{3}\;\langle\mathcal{Q}_2 x,x\rangle^{3/2}$$

with some symmetric positive definite matrices $(\mathcal{Q}_1,\mathcal{Q}_2)$ and some given vector $q\in\mathbb{R}^{r_1}$. In this case we have

$$\begin{array}{rcl} \partial\mathcal{V}(x) & = & q+\mathcal{Q}_1 x+\langle\mathcal{Q}_2 x,x\rangle^{1/2}\;\mathcal{Q}_2 x, \\ \partial^2\mathcal{V}(x) & = & \mathcal{Q}_1+\langle\mathcal{Q}_2 x,x\rangle^{1/2}\;\mathcal{Q}_2+\langle\mathcal{Q}_2 x,x\rangle^{-1/2}\;\mathcal{Q}_2 xx'\mathcal{Q}_2. \end{array}$$

In this situation we have

$$\|\partial^2\mathcal{V}(x)-\partial^2\mathcal{V}(y)\|\leqslant 2\ \|\mathcal{Q}_2\|^{3/2}\ \|y-x\|. \tag{10}$$

This shows that conditions (7) are met with the parameters

$$(\lambda_{\partial A},\kappa_{\partial A})=\beta\ \left(2^{-1}\lambda_{min}(\mathcal{Q}_1),\ 2\lambda_{max}^{3/2}(\mathcal{Q}_2)\right).$$

A proof of (10) is provided in [14], section 6. More generally these regularity conditions also hold if we replace in (9) the parameter σ_1 by any choice of covariance matrice R_1. Also observe that the Langevin diffusion associated with the null form $\mathcal{Q}=0$ coincides with the conventional linear-Gaussian filtering problem discussed in [13]. Stochastic gradient-flow diffusions of the form (9) arise in a variety of application domains. In mathematical finance and mean field game theory [6, 19] these Langevin models describe the interacting-collective behaviour of r_1-individuals. For instance in the Langevin model discussed in [19] the state variables $X_t=\left(X_t^i\right)_{1\leqslant i\leqslant r_1}$ represent the log-monetary reserves of r_1 banks lending and borrowing to each other. The quadratic potential function is given by

$$\langle\mathcal{Q}_1x,x\rangle=\sum_{1\leqslant i\leqslant r_1}\left(x_i-\frac{1}{r_1}\sum_{1\leqslant j\leqslant r_1}x_j\right)^2\Rightarrow\ \mathcal{Q}_1>\left(1-\frac{1}{r_1}\right)\ I_{r_1}.$$

In this context, the parameter β represents the mean-reversion rate between banks. More general interacting potential functions can be considered. Mean field type diffusion processes are also used to design low-representation of fluid flow velocity fields. These vortex-type particle filtering problems are developed in some details in the pionnering articles by E. Mémin and his co-authors [7, 8, 10, 33]. These probabilistic interpretations of the 2d-incompressible Navier-Stokes equation represent the vorticity map as a mixture of basis functions centered around each vortex.

In this connexion, we mention that our approach also applies to interacting diffusion gradient flows described by a potential function of the form

$$\mathcal{V}(x)=\sum_{1\leqslant i\leqslant r_1}\mathcal{U}_1(x_i)+\sum_{1\leqslant i\neq j\leqslant r_1}\mathcal{U}_2(x_i,x_j)$$

for some gradient Lipschitz strongly convex confining potential $\mathcal{U}_i:\mathbb{R}^i\mapsto[0,\infty[$, $i=1,2$. In this situation, we have

$$\partial^2\mathcal{U}_1\geqslant u_1\quad\text{and}\quad\partial^2\mathcal{U}_2\geqslant u_2\ I_2\Longrightarrow\ \partial^2\mathcal{V}>v\ I_{r_1}\quad\text{with}\quad v:=(u_1+(r_1-1)u_2)>0 \tag{11}$$

We further assume that

$$\begin{array}{rcl}|\partial^2\mathcal{U}_1(x_1)-\partial^2\mathcal{U}_1(y_1)| & \leqslant & \kappa_{\partial^2\mathcal{U}_1}\ |x_1-x_2|,\\ \|\partial^2\mathcal{U}_2(x_1,x_2)-\partial^2\mathcal{U}_2(y_1,y_2)\| & \leqslant & \kappa_{\partial^2\mathcal{U}_2}\ \|(x_1,x_2)-(y_1,y_2)\|.\end{array}$$

In this case, we have

$$\|\partial^2\mathcal{V}(x)-\partial^2\mathcal{V}(y)\|\leqslant\kappa_{\partial^2\mathcal{V}}\ \|x-y\|\quad\text{with}\quad\kappa_{\partial^2\mathcal{V}}:=\kappa_{\partial^2\mathcal{U}_1}+\kappa_{\partial^2\mathcal{U}_2}\ (r_1-1)\ \sqrt{2(r_1-1)}. \tag{12}$$

This shows that conditions (7) are met with

$$(\lambda_{\partial A},\kappa_{\partial A})=\beta\ \left(2^{-1}(u_1+(r_1-1)u_2),\ \kappa_{\partial^2\mathcal{U}_1}+\kappa_{\partial^2\mathcal{U}_2}\ (r_1-1)\ \sqrt{2(r_1-1)}\right).$$

The detailed proofs of (11)-(12) are provided in [14], section 6.

1.3.2 Observability conditions

To introduce our observability conditions we give a brief introduction to the class of observation processes discussed in this article. When the observation variables are the same as the ones of the signal, the signal observation has the same dimension as the signal and resumes to some equation of the form

$$dY_t = b\ X_t\ dt + \sigma_2\ dV_t \tag{13}$$

for some parameters $b \in \mathbb{R}$ and $\sigma_2 \geqslant 0$. These sensors are used in data grid-type assimilation problems when measurements can be evaluated at each cell. These fully observed models are discussed in Section 4 in [24] in the context of the Lorentz-96 filtering problems. These observation processes are also used in the article [4] for application to nonlinear and multi-scale filtering problem. In this context, the observed variables represent the slow components of the signal. When the fast components are represented by some Brownian motion with a prescribed covariance matrix, the filtering of the slow components with full observations take the form (13).

For partially observed signals we cannot expect any stability properties of the EKF and the En-EKF without introducing some structural conditions of observability and controllability on the signal-observation equation (1). To get one step further in our discussion, observe that the EKF equation (2) implies that

$$d(\widehat{X}_t - X_t) = \left[(A(\widehat{X}_t) - A(X_t)) - P_t S(\widehat{X}_t - X_t)\right]\ dt + P_t\ C' R_2^{-1/2}\ dV_t + R_1^{1/2}\ dW_t \tag{14}$$

This equation shows that the stability properties of this process depends on the nature of the real eigenvalues of the symmetric matrices $(A(x) - P_t S)_{sym}$, with $x \in \mathbb{R}^{r_1}$. In contrast with the conventional Kalman-Bucy filter, the Riccati equation (2) is a stochastic equation.

In this connection, we already mention that the sample covariance matrices p_t of the En-EKF also satisfy a stochastic Riccati type equation of the same form, up to some fluctuation martingale (see for instance (29) in Theorem 3.2 in the present article). In the same vein, we shall see in (27) that the En-EKF sample mean m_t evolution satisfies the same equation as the EKF, up to some fluctuation martingales coming from the fluctuations of the sample-covariance matrices and the ones of the sample-particles.

As a result, the stability properties of the EKF and the En-EKF are not induced by some kind of observability condition that ensure the existence of a steady state deterministic covariance matrix. The random fluctuations of the matrices $\partial A(\widehat{X}_t)$ and $\partial A(m_t)$ as well as the fluctuations of the stochastic matrices $(A(m_t) - p_t S)_{sym}$ may corrupt the stability in the EKF and the En-EKF, even if the linearized filtering problem around some chosen state is observable and controllable. For instance the empirical covariance matrices may not be invertible for small sample sizes. For a more thorough discussion on the stability properties of Kalman-Bucy filters, the EKF and Riccati equations we refer the reader to [13, 14], and the references therein.

As shown in the system above, these fluctuations enter in two different ways in the En-EKF. The first one in the drift function of the system, the other one through the diffusive part.

Therefore the fluctuations of the empirical covariances from small sample sizes corrupt the natural stabilizing effect of the observation process in the EKF filter evolution. In practice it has been observed that these fluctuations induce an underestimation of the true error covariances. As a result the En-EKF eventually ignores the information given by the observations. This lack of observation-driven component also leads to the divergence of the filter.

Last but not least, from another numerical viewpoint, the En-EKF is also know to be not robust, in the sense that arithmetic errors may accumulate even if the exact filter is stable.

All of these instability properties of the EnKF are well-known and often referred as the catastrophic filter divergence in data assimilation literature, see for instance [20, 22, 28], and the references therein. As mentioned by the authors in [22], "catastrophic filter divergence is a well-documented but mechanistically mysterious phenomenon whereby ensemble-state estimates explode to machine infinity despite the true state remaining in a bounded region". In all the situations discussed above the instability properties of Ensemble Kalman-Bucy type filters are related to some observability problem.

The stability analysis of diffusion processes is always much more documented than the ones on their possible divergence. For instance, in contrast with conventional Kalman-Bucy filters, the stability properties of the EnKF are not induced by some kind of observability or controllability condition. The only known results for discrete generation EnKF is the recent work by X. T. Tong, A. J. Majda and D. Kelly [36]. One of the main assumptions of the article is that the sensor-matrix has full rank. The authors also provide a concrete numerical example of filtering problem with sparse observations for which the EnKF experiences a catastrophic divergence. These divergence properties have been analysis in some details in the article [13] in the context of linear-Gaussian filtering problems. The full rank observation assumption avoids the EnKF to experience local or global exponential instabilities.

To quantify and control uniformly in time the propagations of these instabilities we need to introduce some strong observability condition that ensure that the system is globally and locally stable. In the further development of the article we assume that the following condition is satisfied:

$$(S) \qquad S = \rho(S)\ Id \quad \text{for some} \quad \rho(S) > 0. \tag{15}$$

The fully observed model discussed in (13) clearly satisfies condition (15) with the parameter $\rho(S) = (b/\sigma_2)^2$. Condition (15) ensures that the particle EnKF has uniformly bounded $\mathbb{L}_n$-moments for any $n \geqslant 1$. In the context of linear-Gaussian filtering problems, this condition is also essential to ensure the uniform convergence of Ensemble Kalman-Bucy filter w.r.t. the time parameter [13]. This article also provides a geometric description of the divergence regions in the set of positive covariance matrices for elementary 2-dimensional observable and controllable systems. When condition (S) is not met, we design stochastic observers driven by these matrices that diverge when the signal drift matrix is unstable (see Section 4 in [13]). (see Section 4 in [13]).

From the pure mathematical viewpoint the observability condition (S) allows to combine exponential semigroup techniques with spectral analysis and log-norm inequalities. To get some intuition and to better connect this work with [13] we give some brief comments on these spectral techniques:

For 2-dimensional linear signals $A(x) = Ax$, the existence and the uniqueness of the steady state P of the Riccati equation (2) is ensured by some appropriate observability and controllability conditions. In this context we have $\mu(A - PS) < 0$ even for unstable signal-drift matrices. This condition ensures the stability of the steady state filter.

Starting from the steady state $P_0 = P$ the EnKF filter is driven by stochastic matrices p_t that converge to P, as the size N of the ensemble tends to ∞. The stability analysis of the EnKF filter now depends on the sign of the log-norms $t \mapsto \mu(A - p_t S)$ of the stochastic matrices. The fluctuations of p_t around P are defined by the matrices

$$Q_t := \sqrt{N}\ (p_t - P) \in \mathbb{S}_{r_1} \Longleftrightarrow p_t = P + \frac{1}{\sqrt{N}}\ Q_t \in \mathbb{S}^+_{r_1} \tag{16}$$

Under condition (S) we have $\mu(A) < 0 \Rightarrow \mu(A - p_t S) = \mu((A - PS) - Q_t S) < \mu(A) < 0$ for any fluctuation matrices Q_t. When (S) is not met the local divergence domain of matrices Q_t such that $\mu(A - p_t S) = \mu((A - PS) - Q_t S) > 0$ may be very large, even when $\mu(A) < 0$. The refined analysis on the stability of these models requires to analyze in some details the random excursion of the matrices into these local divergence domains. For a more thorough discussion on these local and global divergence issues in the context of linear systems we refer the reader to section 4 in the article [13].

Last but not least, we mention that (15) is satisfied when the filtering problem is similar to the ones discussed above; that is, up to a change of basis functions. More precisely, any filtering problem (1) with $r_1 = r_2$ and s.t. $(R_2^{-1/2}B)$ is invertible can be turned into a filtering problem equipped with an identity sensor matrix; even when the original matrix $S = C'R_2^{-1}C = \mathcal{C}'\mathcal{C}$ does not satisfy (15). To check this claim we observe that

$$\mathcal{Y}_t := R_2^{-1/2} Y_t \quad \text{and} \quad \mathcal{X}_t := R_2^{-1/2} B X_t \Longrightarrow \left\{ \begin{array}{rcl} d\mathcal{X}_t &=& \mathcal{A}(\mathcal{X}_t)\ dt + \mathcal{R}_1^{1/2}\ dW_t \\ d\mathcal{Y}_t &=& \mathcal{X}_t\ dt + dV_t \end{array} \right.$$

with the drift function

$$\mathcal{A} := (R_2^{-1/2}B) \circ A \circ (R_2^{-1/2}B)^{-1} \quad \text{and the matrix} \quad \mathcal{R}_1 := R_2^{-1/2} B R_1 B' R_2^{-1/2}.$$

In this situation the filtering model $(\mathcal{X}_t, \mathcal{Y}_t)$ satisfies (15). In addition, we have

$$A = (R_2^{-1/2}B)^{-1} \circ \partial U \circ (R_2^{-1/2}B) \Rightarrow (\mathcal{A}, \partial\mathcal{A}) = (\partial U, \partial^2 U).$$

In this situation the filtering model $(\mathcal{X}_t, \mathcal{Y}_t)$ satisfies (15) and the signal process $\mathcal{X}_t$ belongs to the class of Langevin type diffusion discussed in Section 1.3.1.

2 Statement of the main results

2.1 Concentration inequalities

One of our results concerns the stability properties of the EKF-diffusion (3). It is no surprise that these properties strongly depend on logarithmic norm of the drift function A as well as on the size of covariance matrices of the signal-observation diffusion. For instance, we have the uniform moment estimate

$$\lambda_{\partial A} > 0 \Rightarrow \forall \delta \geqslant 1 \quad \sup_{t \geqslant 0} \left\{ \mathbb{E}[\|X_t\|^\delta] \vee \operatorname{tr}(P_t) \vee \mathbb{E}[\|X_t - \widehat{X}_t\|^\delta] \right\} \leqslant c. \tag{17}$$

A detailed proof of these stochastic stability properties including exponential concentration inequalities can be found in [14]. Observe that $\mathrm{tr}(P_t)$ is random so that the above inequality provides an almost sure estimate. To be more precise we use (2) to check that

$$\partial_t \mathrm{tr}\,(P_t) \leqslant -\lambda_{\partial A}\ \mathrm{tr}\,(P_t) + \mathrm{tr}(R) \Longrightarrow \mathrm{tr}(P_t) \leqslant e^{-\lambda_{\partial A} t}\ \mathrm{tr}\,(P_0) + \mathrm{tr}(R)/\lambda_{\partial A}. \tag{18}$$

The detailed proof of (18) can be found on page 32.

To get one step further in our discussion, we consider the following ratio

$$\lambda_S := \frac{\lambda_{\partial A}}{\rho(S)} \qquad \lambda_R := \frac{\lambda_{\partial A}}{\mathrm{tr}(R)} \quad \text{and} \quad \lambda_K := \frac{\lambda_{\partial A}}{\kappa_{\partial A}}.$$

Roughly speaking, the three quantities presented above measure the relative stability index of the signal drift with respect to the perturbation degree of the sensor, the one of the signal, and the modulus of continuity of the Jacobian entering into the Riccati equation. For instance, λ_S is high for sensors with large perturbations, inversely λ_R is large for signals with small perturbations. Most of our analysis relies on the behaviour of the following quantities:

$$\begin{aligned}
\lambda_{R,S} &:= (8e)^{-1}\ \lambda_R\ \sqrt{\lambda_S}\ \left[1 + \frac{2}{\lambda_R \lambda_S}\right]^{-1} \\
\widehat{\lambda}_{\partial A}/\lambda_{\partial A} &:= \left(\frac{1}{2} - \frac{2}{\lambda_K \lambda_R}\right) + \left(\frac{1}{2} - \frac{1}{\sqrt{\lambda_S}}\right)\left[1 - \frac{3}{4}\frac{1}{\sqrt{\lambda_S}}\right]
\end{aligned}$$

In the quadratic Langevin-signal filtering problem discussed in (9) and (13) with $b = 1 = \beta$ these parameters resume to

$$\lambda_S := \frac{1}{2}\ \sigma_2^2\ v, \qquad \lambda_R := \frac{1}{2r_1}\ \sigma_1^{-2}\ v \quad \text{and} \quad \lambda_K := \infty \tag{19}$$

In this situation we have

$$\widehat{\lambda}_{\partial A}/\lambda_{\partial A} := \frac{1}{2} + \frac{1}{2}\ \left(1 - 2\sqrt{2}\frac{1}{\sigma_2\ \sqrt{v}}\right)\left[1 - \frac{3\sqrt{2}}{4}\frac{1}{\sigma_2\ \sqrt{v}}\right].$$

Notice that these parameters do not depend on the dimension of the signal, nor on the diffusion parameter σ_1.

In addition, we have $\widehat{\lambda}_{\partial A}/\lambda_{\partial A} > 0$ for any choice of parameters (v, σ_2).

To better connect these quantities with the stochastic stability of the EKF diffusion we discuss some exponential concentration inequalities that can be easily derived from our analysis. These concentration inequalities are of course more accurate than any type of mean square error estimate. Let $\widehat{X}_t(m,p)$ be the solution of the EKF equation (2) starting at $(\widehat{X}_0, P_0) = (m,p)$, and let $X_t(x)$ be the state of the signal starting at $X_0(x) = x$. Let $\varpi(\delta)$ be the function

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

In this notation, we have the following exponential concentration inequalities.

Theorem 2.1. *For any time horizon $t \in [0,\infty[$, and any $\delta \geqslant 0$ the probabilities of the following events*

$$\|X_t(x) - \widehat{X}_t(m,p)\|^2 \leqslant \frac{1}{2e}\,\varpi(\delta)\,\sqrt{\lambda_S}/\lambda_{R,S} + 2\,e^{-\lambda_{\partial A}t}\,\|x-m\|^2 + 8\,\varpi(\delta)\,\frac{|e^{-\lambda_A t} - e^{-\lambda_{\partial A}t}|}{|\lambda_A/\lambda_{\partial A} - 1|}\,\mathrm{tr}(p)^2/\lambda_S$$

and

$$\|\overline{X}_t(m,p) - \widehat{X}_t(m,p)\|^2 \leqslant \frac{1}{2e}\,\varpi(\delta)\,\sqrt{\lambda_S}/\lambda_{R,S} + 8\,\varpi(\delta)\,e^{-\lambda_{\partial A}t}\,\mathrm{tr}(p)^2/\lambda_S$$

are greater than $1 - e^{-\delta}$.

The proof of the first assertion is a consequence of [14, Theorem 1.1], the proof of the second one is a consequence of the $\mathbb{L}_\delta$-mean error estimate (32). These concentration inequalities show that the quantity

$$\sqrt{\lambda_S}/\lambda_{R,S} = 8e\,\lambda_S\,\frac{1}{\lambda_R\lambda_S}\,\left[1 + \frac{2}{\lambda_R\lambda_S}\right]$$

can be interpreted as the size of a confidence interval around the values of *the true signal*, as soon as the time horizon is large. It is also notable that the same quantity controls the fluctuations of the EKF diffusion around the values of the EKF. These confidence intervals are small for stable signals with small perturbations. In the quadratic Langevin-signal filtering problem discussed in (9) and (13) with $b = 1 = \beta$ the above quantity resumes to

$$\sqrt{\lambda_S}/\lambda_{R,S} = 2^4 e\,\frac{1}{v}\,r_1\,\sigma_1^2\,\left[1 + \frac{8}{v^2}\,r_1\,\left(\frac{\sigma_1}{\sigma_2}\right)^2\right].$$

For unit signal-to-noise ratio $\sigma_1 = \sigma_2$ these fluctuation parameters are small for stable signals with small perturbations. The above formula also indicate the degradation of the fluctuation parameter when the size of the system is large.

2.2 A stability theorem

We further assume that

$$(\lambda_K\lambda_R/4) \wedge \lambda_{R,S} \wedge (\lambda_S/4) > 1. \tag{20}$$

This regularity property is a purely technical condition. The condition $(\lambda_K\lambda_R/4) \wedge (\lambda_S/4)$ ensures that $0 < \widehat{\lambda}_{\partial A} \leqslant \lambda_{\partial A}$, while $\lambda_{R,S} > 1$ is used to derive $\mathbb{L}_p$-mean error estimates with some parameter $p \geqslant 1$ that depends on $\lambda_{R,S}$.

The condition (20) is clearly met as soon as λ_R and λ_S are sufficiently large. As we shall see the quantity $\widehat{\lambda}_{\partial A}$ represents the Lyapunov stability exponent of the EKF. This exponent is decomposed into two parts. The first one represents the relative contribution of the signal perturbations, the second one is related to the sensor perturbations.

In contrast with the linear-Gaussian case discussed in [13], the stochastic Riccati equation (2) depends on the states of the EKF. As shown in [14] the stability of the EKF relies on a stochastic Lyapunov exponent that depends on the random trajectories of the filter as well as on the signal-observation processes. The technical condition (20) allows to control uniformly the fluctuations of these stochastic exponents with respect to the time horizon.

A more detailed discussion on the regularity condition (20), including a series of sufficient conditions are provided in the appendix, Section 6.1. For filtering problems with an observation process of the form (13) with $\rho(S) = (b/\sigma_2)^2 = 1$ we have

$$\lambda_S = \lambda_{\partial A} \quad \Longrightarrow \quad \lambda_{R,S} := \frac{1}{8e\mathrm{tr}(R)}\ \frac{\lambda_{\partial A}^{3+1/2}}{\lambda_{\partial A}^2 + 2\mathrm{tr}(R)}.$$

In this situation (20) is met as soon as the following easy to check condition is satisfied

$$\lambda_{\partial A} > 4 \quad \text{and} \quad \mathrm{tr}(R) \leqslant \frac{\lambda_{\partial A}^2}{2}\left\{\frac{1}{2\kappa_{\partial A}} \wedge \left[\sqrt{1+\frac{1}{4e\sqrt{\lambda_{\partial A}}}}-1\right]\right\}. \tag{21}$$

A detailed proof of this assertion is provided in the end of Section 6.1. In the quadratic Langevin-signal filtering problem discussed in (9) and (13) with $b = \sigma_2$, condition (21) resumes to

$$v/8 > 1 \quad \text{and} \quad 2\sqrt{2}e\ r_1\sigma_1^2 \leqslant (v/8)\ \frac{1}{\sqrt{1+\frac{1}{2\sqrt{2}ev}}+1}.$$

These conditions are clearly much stronger than the ones discussed in [13] in the context of linear-Gaussian filtering problems. For the same type of filtering problem, exponential stability and uniform propagations of chaos for the EnKF hold as soon as $v > 0$.

Let $(\overline{X}_t, \overline{Z}_t)$ be a couple of EKF Diffusions (3) starting from two random states with mean $(\widehat{X}_0, \check{X}_0)$ and covariances matrices $(P_0, \check{P}_0)$ (and driven by the same Brownian motions $(\overline{W}_t, \overline{V}_t)$). One key feature of these nonlinear diffusions is that the $\mathcal{G}_t$-conditional expectations $(\widehat{X}_t, \check{X}_t)$ and the $\mathcal{G}_t$-conditional covariance matrices $(P_t, \check{P}_t)$ satisfy the EKF and the stochastic Ricatti equations discussed in (2).

Whenever condition (20) is satisfied we recall from [14] that 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

$$\mathbb{E}\left(\|(\widehat{X}_t, P_t) - (\check{X}_t, \check{P}_t)\|^{\delta_S} \mid \mathcal{G}_s\right)^{2/\delta_S} \leqslant \mathcal{Z}_s\ \exp\left[-(1-\epsilon)\,\widehat{\lambda}_{\partial A}(t-s)\right]\ \|(\widehat{X}_s, P_s) - (\check{X}_s, \check{P}_s)\|^2$$

with $\delta_S := 2^{-1}\sqrt{\lambda_S}$, and some random process $\mathcal{Z}_t$ satisfying the uniform moment condition

$$\sup_{t\geqslant 0} \mathbb{E}\left(\mathcal{Z}_t^{\alpha}\right) < \infty \quad \text{with} \quad \alpha = 2\lambda_{R,S}\ \delta_S. \tag{22}$$

These conditional contraction estimates can be used to quantify the stability properties of the EKF. More precisely, if we set

$$\mathbb{P}_t = \mathrm{Law}(\widehat{X}_t, P_t) \quad \text{and} \quad \check{\mathbb{P}}_t = \mathrm{Law}(\check{X}_t, \check{P}_t)$$

then the above contraction inequality combined with the uniform estimates (17) readily implies that

$$\forall t \geqslant t_0 \qquad \mathbb{W}_{\delta_S}^2(\mathbb{P}_t, \check{\mathbb{P}}_t) \leqslant c\ \exp\left[-t\ (1-\epsilon)\ \widehat{\lambda}_{\partial A}\right]$$

for any $\epsilon \in [0,1[$, with some time horizon t_0. This stability property ensures that the EKF forgets exponentially fast any erroneous initial condition. Of course these forgetting properties of the EKF do not give any information at the level of the process. One of the main objective of the article is to complement these conditional expectation stability properties at the level of the McKean-Vlasov type nonlinear EKF-diffusion (3).

Our second main result can basically be stated as follows.

Theorem 2.2. *Let $(\overline{\eta}_t, \breve{\eta}_t)$ be the probability distributions of a couple $(\overline{X}_t, \overline{Z}_t)$ of EKF Diffusions (3) starting from two possibly different random states. Assume condition (20) is met with $\delta'_S := \delta_S/4 \geqslant 2$. In this situation, for any $\epsilon \in [0,1[$ there exists some time horizon t_0 such that for any $t \geqslant t_0$ we have*

$$\mathbb{W}^2_{\delta'_S}(\overline{\eta}_t, \breve{\eta}_t) \leqslant c\ \exp\left[-t\ (1-\epsilon)\ \lambda\right] \quad \text{with} \quad \lambda \geqslant \widehat{\lambda}_{\partial A} \wedge (\lambda_{\partial A}/4). \tag{23}$$

2.3 A uniform propagation of chaos theorem

Our next objective is to analyze the long-time behaviour of the mean field type En-EKF model discussed in (4). From the practical estimation point of view, only the sample mean and the sample covariance matrices (5) are of interest since these quantities converge to the EKF and the Riccati equations, as N tends to ∞. Another important problem is to quantify the bias of the mean field particle approximation scheme. These properties are related to the propagation of chaos properties of the mean field particle model. They are expressed in terms of the collection of probability distributions

$$\mathbb{P}^N_t = \text{Law}(m_t, p_t), \qquad \mathbb{Q}^N_t = \text{Law}(\xi^1_t) \quad \text{and} \quad \mathbb{Q}_t = \text{Law}(\zeta^1_t).$$

Theorem 2.3. *Assume that (20) is met with $\delta_{R,S} := (e\lambda_{R,S}) \wedge \delta_S \geqslant 2$. In this situation, there exists some $N_0 \geqslant 1$ and some $\beta \in]0,1/2]$ such that for any $N \geqslant N_0$, we have the uniform non asymptotic estimates*

$$\text{tr}(P_0)^2 \leqslant \frac{\lambda_S}{\lambda_R}\left[\frac{1}{2} + \frac{1}{\lambda_R \lambda_S}\right] \quad \Longrightarrow \quad \sup_{t \geqslant 0} \mathbb{W}_{\delta_{R,S}}\left(\mathbb{P}^N_t, \mathbb{P}_t\right) \leqslant cN^{-\beta}. \tag{24}$$

In addition, when $\delta_{R,S} \geqslant 4$ we have the uniform propagation of chaos estimate

$$\sup_{t \geqslant 0} \mathbb{W}_2\left(\mathbb{Q}^N_t, \mathbb{Q}_t\right) \leqslant cN^{-\beta}. \tag{25}$$

Our analysis does not provide an explicit formula for the rate of convergence β. We conjecture that the optimal rate is $\beta = 1/2$ as in the linear-Gaussian case developed in [13].

For the quadratic Langevin-signal filtering model discussed in (9) and (13) with $b = 1 = \beta$, by (19) the l.h.s. condition in (24) resumes to

$$\text{tr}(P_0)^2 \leqslant \frac{\lambda_S}{\lambda_R}\left[\frac{1}{2} + \frac{1}{\lambda_R \lambda_S}\right] = r_1\ (\sigma_1\sigma_2)^2 \left[\frac{1}{2} + r_1\ \left(\frac{2}{v}\right)^2\ \left(\frac{\sigma_1}{\sigma_2}\right)^2\right].$$

We end this section with some comments on our regularity conditions.

The condition (15) is needed to control the fluctuations of the trace of the sample covariance matrices of the En-EKF, even if the trace expectation is uniformly stable. We believe that this technical observability condition can be relaxed.

Despite our efforts, our regularity conditions are stronger than the ones discussed in [13] in the context of linear-Gaussian filtering problems. The main difference here is that the signal stability is required to compensate the possible instabilities created by highly informative sensors when we initialize the filter with wrong conditions.

Next we comment the trace condition in (24). As we mentioned earlier, the stability properties of the limiting EKF-diffusion (3) are expressed in terms of a stochastic Lyapunov exponent that depends on the trajectories of the signal process. The propagation of chaos properties of the mean field particle approximation (4) depend on the long-time behaviour of these stochastic Lyapunov exponents. Our analysis is based on a refined analysis of Laplace transformations associated with quadratic type stochastic exponents. The existence of these χ-square type Laplace transforms requires some regularity on the signal process. For instance at the origin we have

$$(\mathrm{tr}(P_0)\leqslant)\ r_1\rho(P_0)\leqslant 1/(4\delta)\Longrightarrow \mathbb{E}\left(\exp\left[\delta\|X_0-\widehat{X}_0\|^2\right]\right)\leqslant e. \tag{26}$$

The proof of (26) and more refined estimates can be found in [14].

From the numerical viewpoint the trace condition in (24) is related to the initial location of the particles and the signal-observation perturbations. Signals with a large diffusion part are more likely to correct an erroneous initialization. In the same vein, the estimation problems associated with sensors corrupted by large perturbations are less sensitive to the initialization of the filter. In the reverse angle, when the signal is almost deterministic and the sensor is highly informative the particles need to be initialized close to the true value of the signal.

To better connect our work with existing literature we end our discussion with some connection with the variance inflation technique introduced by J.L. Anderson in [1, 2, 3] and further developed by D.T. B. Kelly, K.J. Law, A. M. Stuart [21] and by X. T. Tong, A. J. Majda and D. Kelly [36]. In discrete time settings this technique amounts of adding an extra positive matrix in the Riccati updating step. This strategy allows to control the fluctuations of the sample covariance matrices. In continuous time settings, this technique amounts of changing the covariance matrix $\mathcal{P}_{\eta_t}$ in the EKF diffusion (3) by $\mathcal{P}_{\eta_t}+\theta\ Id$ for some tuning parameter $\theta>0$. The resulting EKF-diffusion (3) is given by the equation

$$\begin{aligned}
d\overline{X}_t &= \left(\mathcal{A}\left(\overline{X}_t,\mathbb{E}[\overline{X}_t\mid\mathcal{G}_t]\right)-\theta\ S\ \overline{X}_t\right)\ dt+\mathcal{P}_{\eta_t}B'R_2^{-1}\ \left[dY_t-\left(B\overline{X}_t\ dt+R_2^{1/2}\ d\overline{V}_t\right)\right]\\
&\quad+\left[R_1^{1/2}\ d\overline{W}_t-\theta\ B'R_2^{-1/2}d\overline{V}_t\right]+\theta\ B'R_2^{-1}\ dY_t.
\end{aligned}$$

The stabilizing effects of the variance inflation technique are clear. The last term in the r.h.s. of the above displayed formula has no effect (by simple coupling) on the stability properties of the diffusion. The form of the drift also indicates that we increase the Lyapunov exponent by an additional factor θ (as soon as $\rho(S)>0$). In addition we increase the noise of the diffusion by a factor θ^2, in the sense that the covariance matrix of the perturbation term $R_1^{1/2}\ d\overline{W}_t-\theta\ B'R_2^{-1/2}d\overline{V}_t$ is given by $R_1+\theta^2 S$. We believe that the stability analysis of these regularized models is simplified by these additional regularity properties. This class of regularized nonlinear diffusions can probably be studied quite easily using the stochastic analysis developed in this article. We plan to develop this analysis in a forthcoming study.

3 Some preliminary results

This short section presents a couple of pivotal results. The first one ensures that the Extended Kalman-Bucy filter coincides with the $\mathcal{G}_t$-conditional expectations of the nonlinear diffusion $\overline{X}_t$. The second result shows that the stochastic processes (m_t, p_t) satisfy the same equation as $\left(\widehat{X}_t, P_t\right)$, up to some local fluctuation orthogonal martingales with angle brackets that only depend on the sample covariance matrix p_t.

Proposition 3.1. *We have the equivalence*

$$\mathbb{E}(\overline{X}_0) = \widehat{X}_0 \quad and \quad \mathcal{P}_{\eta_0} = P_0 \Longleftrightarrow \forall t \geqslant 0 \quad \mathbb{E}(\overline{X}_t \mid \mathcal{G}_t) = \widehat{X}_t \quad and \quad \mathcal{P}_{\eta_t} = P_t.$$

Proof. Taking the $\mathcal{G}_t$-conditional expectations in (3) we find the diffusion equation

$$d\,\mathbb{E}(\overline{X}_t \mid \mathcal{G}_t) = A(\mathbb{E}(\overline{X}_t \mid \mathcal{G}_t))\; dt + \mathcal{P}_{\eta_t} B' R_2^{-1}\; \left[dY_t - B\; \mathbb{E}(\overline{X}_t \mid \mathcal{G}_t) dt\right].$$

Equivalently, if we set $\mathbb{E}(\overline{X}_t \mid \mathcal{G}_t) = \widehat{X}_t$ then we find that

$$d\,\widehat{X}_t = A(\widehat{X}_t)\; dt + \mathcal{P}_{\eta_t} B' R_2^{-1}\; \left[dY_t - B\; \widehat{X}_t\; dt\right].$$

Let us compute the evolution of $\mathcal{P}_{\eta_t}$. We set $\widetilde{X}_t = \overline{X}_t - \mathbb{E}(\overline{X}_t \mid \mathcal{G}_t) = \overline{X}_t - \widehat{X}_t$. In this notation we have

$$\begin{aligned} d\widetilde{X}_t &= \partial A(\mathbb{E}(\overline{X}_t \mid \mathcal{G}_t))\; \widetilde{X}_t\; dt \;+\; R_1^{1/2}\; d\overline{W}_t - \mathcal{P}_{\eta_t} B' R_2^{-1}\; \left[B\widetilde{X}_t dt + R_2^{1/2}\; d\overline{V}_t\right] \\ &= \left[\partial A(\mathbb{E}(\overline{X}_t \mid \mathcal{G}_t)) - \mathcal{P}_{\eta_t} S\right]\; \widetilde{X}_t\; dt \;+\; R_1^{1/2}\; d\overline{W}_t - \mathcal{P}_{\eta_t} B' R_2^{-1/2}\; d\overline{V}_t. \end{aligned}$$

This implies that

$$\begin{aligned} d(\widetilde{X}_t\widetilde{X}_t') &= \left\{\left[\partial A(\widehat{X}_t) - P_t S\right]\; \widetilde{X}_t\widetilde{X}_t'\; dt \;+ \widetilde{X}_t\widetilde{X}_t'\left[\partial A(\widehat{X}_t) - P_t S\right]' + (R + \mathcal{P}_{\eta_t} S \mathcal{P}_{\eta_t})\right\}\; dt \\ &\quad + \left[R_1^{1/2}\; d\overline{W}_t - \mathcal{P}_{\eta_t} B' R_2^{-1/2}\; d\overline{V}_t\right] \widetilde{X}_t' + \widetilde{X}_t \left[R_1^{1/2}\; d\overline{W}_t - \mathcal{P}_{\eta_t} B' R_2^{-1/2}\; d\overline{V}_t\right]'. \end{aligned}$$

Taking the $\mathcal{G}_t$-conditional expectations we conclude that

$$\begin{aligned} \partial_t \mathcal{P}_{\eta_t} &= \left[\partial A(\widehat{X}_t) - \mathcal{P}_{\eta_t} S\right]\; \mathcal{P}_{\eta_t}\; dt \;+ \mathcal{P}_{\eta_t}\left[H(\widehat{X}_t) - \mathcal{P}_{\eta_t} S\right]' + (R + \mathcal{P}_{\eta_t} S \mathcal{P}_{\eta_t}) \\ &= \partial A(\widehat{X}_t)\mathcal{P}_{\eta_t} + \mathcal{P}_{\eta_t}\partial A(\widehat{X}_t)' + R - \mathcal{P}_{\eta_t} S \mathcal{P}_{\eta_t}. \end{aligned}$$

This ends the proof of the proposition. ∎

Theorem 3.2 (Fluctuation theorem [13]). *The stochastic processes (m_t, p_t) defined in (5) satisfy the diffusion equations*

$$dm_t = A\left[m_t\right]\; dt + p_t\; B' R_2^{-1}\; (dY_t - B m_t\; dt) + \frac{1}{\sqrt{N}}\; d\overline{M}_t \tag{27}$$

with the vector-valued martingale $\overline{M}_t = \left(\overline{M}_t(k)\right)_{1\leqslant k\leqslant r_1}$ with the angle-brackets

$$\partial_t \langle \overline{M}_t(k), \overline{M}_t(k') \rangle_t = R(k,k') + (p_t S p_t)(k,k'). \tag{28}$$

We also have the matrix-valued diffusion

$$dp_t = \left(\partial A\left[m_t\right] p_t + p_t \partial A\left[m_t\right]' - p_t S p_t + R\right)\ dt + \frac{1}{\sqrt{N-1}}\ dM_t \tag{29}$$

with a symmetric matrix-valued martingale $M_t = (M_t(k,l))_{1\leqslant k,l\leqslant r_1}$ *and the angle brackets*

$$\begin{aligned}\partial_t \left\langle M(k,l), M(k',l')\right\rangle_t &= (R + p_t S p_t)(k,k')\ p_t(l,l') + (R + p_t S p_t)(l,l')\ p_t(k,k') \\ &\quad + (R + p_t S p_t)(l',k)\ p_t(k',l) + (R + p_t S p_t)(l,k')\ p_t(k,l').\end{aligned} \tag{30}$$

In addition we have the orthogonality properties

$$\left\langle M(k,l), \overline{M}(l')\right\rangle_t = \left\langle M(k,l), V(k')\right\rangle_t = \left\langle \overline{M}(l'), V(k')\right\rangle_t = 0$$

for any $1 \leqslant k, l, l' \leqslant r_1$ *and any* $1 \leqslant k' \leqslant r_2$.

Proof. We have

$$d(\xi_t^i - m_t) = \left[\partial A\left(m_t\right) - p_t B' S\right](\xi_t^i - m_t)dt + dM_t^i$$

with the martingale

$$dM_t^i := R_1^{1/2}\left(d\overline{W}_t^i - \frac{1}{N}\sum_{1\leqslant j\leqslant N} d\overline{W}_t^j\right) - p_t B' R_2^{-1/2}\left(d\overline{V}_t^i - \frac{1}{N}\sum_{1\leqslant j\leqslant N} d\overline{V}_t^j\right).$$

Notice that

$$\partial_t\langle M^i(k), M^i(k')\rangle_t = \left(1 - \frac{1}{N}\right)\ (R + p_t S p_t)(k,k')$$

and for $i \neq j$

$$\partial_t\langle M^i(k), M^j(k')\rangle_t = -\frac{1}{N}\ (R + p_t S p_t)(k,k').$$

The end of the proof follows the proof of [13, Theorem 1], thus it is skipped. This ends the proof of the theorem. ■

4 Stability properties

This section is dedicated to the long-time behaviour of the EKF-diffusion (3), mainly with the proof of Theorem 2.2. We use the stochastic differential inequality calculus developed in [13, 14]. Let $\mathcal{Y}_t$ be some nonnegative process 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 following definition

$$d\mathcal{Y}_t \leqslant \mathcal{Z}_t^+\ dt + d\mathcal{M}_t \iff \left(d\mathcal{Y}_t = \mathcal{Z}_t\ dt + d\mathcal{M}_t \quad \text{with} \quad \mathcal{Z}_t \leqslant \mathcal{Z}_t^+\right). \tag{31}$$

We recall some useful algebraic 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, and $(\alpha, \overline{\alpha})$ a couple of nonnegative parameters. In this case it is readily checked that

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

and

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

We consider a couple of diffusions $(\overline{X}_t, \overline{Z}_t)$ coupled with the same Brownian motions $(\overline{V}_t, \overline{W}_t)$ and the same observation process Y_t, and we set

$$\mathcal{F}_t := \mathcal{G}_t \vee \sigma\left((\overline{X}_s, \overline{Z}_s),\ s \leqslant t\right).$$

Next proposition provides uniform estimates of the $\mathbb{L}_\delta$-centered moments of the EKF-diffusion with respect to the time horizon.

Proposition 4.1. *Assume that $\lambda_{\partial A} > 0$. In this situation, for any $\delta \geqslant 1$ and any time horizon $s \geqslant 0$ we have the uniform almost sure estimates*

$$\begin{aligned}\mathbb{E}\left(\|\overline{X}_t - \widehat{X}_t\|^\delta \mid \mathcal{F}_s\right)^{2/\delta} &\leqslant e^{-\lambda_{\partial A}(t-s)}\ \|\overline{X}_s - \widehat{X}_s\|^2 \\ &\quad +(2\delta-1)\left[\lambda_R^{-1}(1 + 2\ (\lambda_R\lambda_S)^{-1}) + 2e^{-\lambda_{\partial A}(t+s)}\mathrm{tr}(P_0)^2\lambda_S^{-1}\right].\end{aligned} \tag{32}$$

Proof. We have

$$d(\overline{X}_t - \widehat{X}_t) = \left[\partial A(\widehat{X}_t) - P_tS\right]\ (\overline{X}_t - \widehat{X}_t)\ dt\ +\ R_1^{1/2}\ d\overline{W}_t - P_tB'R_2^{-1/2}\ d\overline{V}_t$$

This implies that

$$\begin{aligned}&d\|\overline{X}_t - \widehat{X}_t\|^2 \\ &= \left[2\langle\overline{X}_t - \widehat{X}_t, \left[\partial A(\widehat{X}_t) - P_tS\right]\ (\overline{X}_t - \widehat{X}_t)\rangle + \mathrm{tr}(R_1) + \mathrm{tr}(P_t^2S)\right]\ dt + dM_t \\ &\leqslant \left[-\lambda_{\partial A}\ \|\overline{X}_t - \widehat{X}_t\|^2 + \mathcal{U}_t\right]\ dt + dM_t\end{aligned}$$

with the process

$$\begin{aligned}\mathcal{U}_t &:= \mathrm{tr}(R) + \mathrm{tr}(P_t^2S) \leqslant \mathrm{tr}(R) + \rho(S)\mathrm{tr}(P_t)^2 \\ &\leqslant \mathrm{tr}(R) + \rho(S)\left(e^{-\lambda_{\partial A}t}\ \mathrm{tr}(P_0) + 1/\lambda_R\right)^2\end{aligned}$$

and the martingale

$$dM_t := 2\langle\overline{X}_t - \widehat{X}_t, R_1^{1/2}\ d\overline{W}_t - P_tB'R_2^{-1/2}\ d\overline{V}_t\rangle.$$

Observe that the angle bracket of this martingale satisfies the property

$$\partial_t\langle M\rangle_t \quad = \quad 4\langle\overline{X}_t - \widehat{X}_t, (R + P_tSP_t)\,(\overline{X}_t - \widehat{X}_t)\rangle \leqslant 4\|\overline{X}_t - \widehat{X}_t\|^2\ \|R + P_tSP_t\|.$$

By [14, Corollary 2.2] for any $\delta \geqslant 1$ we have

$$\mathbb{E}\left(\|\overline{X}_t - \widehat{X}_t\|^\delta \mid \mathcal{F}_s\right)^{2/\delta} \leqslant \exp\left(-\lambda_{\partial A}(t-s)\right) \|\overline{X}_s - \widehat{X}_s\|^2$$

$$+(2\delta-1)\int_s^t \exp\left(-\lambda_{\partial A}(t-u)\right) \left(\mathrm{tr}(R) + \rho(S)\mathrm{tr}(P_u)^2\right) du.$$

Observe that by (18)

$$\rho(S)\int_s^t \exp\left(-\lambda_{\partial A}(t-u)\right) \mathrm{tr}(P_u)^2 \, du$$

$$\leqslant 2\rho(S)\int_s^t \exp\left(-\lambda_{\partial A}(t-u)\right) \left[e^{-2\lambda_{\partial A}u} \mathrm{tr}(P_0)^2 + 1/\lambda_R^2\right] du$$

$$\leqslant 2(\lambda_R^2\lambda_S)^{-1} + 2\exp\left(-\lambda_{\partial A}(t+s)\right)\mathrm{tr}(P_0)^2\lambda_S^{-1}.$$

This ends the proof of the proposition. ■

Theorem 4.2. *When the initial random states $\overline{X}_0$ and $\overline{Z}_0$ have the same first and second order statistics, that is when $(\widehat{X}_0, P_0) = (\check{X}_0, \check{P}_0)$, we have the almost sure contraction estimates:*

$$\|\overline{X}_t - \overline{Z}_t\|^2 \leqslant \exp\left[-\lambda_{\partial A}t\right] \|\overline{X}_0 - \overline{Z}_0\|^2.$$

More generally, when condition (20) is met with $\lambda_S \geqslant 4^4$, for any $\epsilon \in [0,1[$ there exists some s such that for any $t \geqslant s$ and any $1 \leqslant \delta \leqslant 4^{-4}\sqrt{\lambda_S}$ we have

$$\mathbb{E}\left(\|\overline{X}_t - \overline{Z}_t\|^{2\delta} \mid \mathcal{F}_s\right)^{1/\delta} \leqslant \exp\left[-(1-\epsilon)\overline{\lambda}_{\partial A}(t-s)\right] \left[\|\overline{X}_s - \overline{Z}_s\|^2 + \overline{\mathcal{Z}}_s\right] \tag{33}$$

with some exponent $\overline{\lambda}_{\partial A} \geqslant \widehat{\lambda}_{\partial A} \wedge (\lambda_{\partial A}/2)$, and some process $\overline{\mathcal{Z}}_t$ satisfying the uniform moment condition

$$\sup_{t\geqslant 0}\mathbb{E}\left(\overline{\mathcal{Z}}_t^{\alpha/4}\right) < \infty \quad \textit{for any} \quad \alpha \leqslant \lambda_{R,S}\sqrt{\lambda_S}. \tag{34}$$

Before getting into the details of the proof of this theorem we mention that (23) is a direct consequence of (33) combined with the uniform estimates (32). Indeed, applying (33), for any $\delta \geqslant 2$ we have

$$\mathbb{E}\left(\|\overline{X}_t - \overline{Z}_t\|^\delta\right)^{1/\delta} \leqslant \exp\left[-(1-\epsilon)\overline{\lambda}_{\partial A}(t-s)/2\right] \left(\mathbb{E}\left[\|\overline{X}_s - \overline{Z}_s\|^\delta\right]^{1/\delta} + \mathbb{E}\left[\overline{\mathcal{Z}}_s^{\delta/2}\right]^{1/\delta}\right).$$

Using (32) and the fact that

$$1 \leqslant \delta/2 \leqslant 16^{-1}\sqrt{\lambda_S} \leqslant 4^{-1}\lambda_{R,S}\sqrt{\lambda_S}$$

we conclude that

$$\mathbb{W}_\delta(\eta_t, \check{\eta}_t) \leqslant c\ \exp\left[-t\ (1-\epsilon)(1-s/t)\overline{\lambda}_{\partial A}/2\right] \leqslant c\ \exp\left[-t\ (1-2\epsilon)\overline{\lambda}_{\partial A}/2\right]$$

as soon as $s/t \leqslant \epsilon$. The end of the proof of (23) is now clear.

Now we come to the proof of the theorem.

Proof of Theorem 4.2:

We have

$$d\overline{X}_t \;=\; \mathcal{A}(\overline{X}_t, \widehat{X}_t)\;dt \;+\; R_1^{1/2}\;d\overline{W}_t + P_t B' R_2^{-1}\;\left[dY_t - \left(B\overline{X}_t dt + R_2^{1/2}\;d\overline{V}_t\right)\right].$$

Using the decomposition

$$\check{P}_t S\overline{Z}_t - P_t S\overline{X}_t \;=\; -P_t S(\overline{X}_t - \overline{Z}_t) + (\check{P}_t - P_t)S\overline{Z}_t$$

we readily check that

$$d\left(\overline{X}_t - \overline{Z}_t\right)$$

$$= \left\{\left[\mathcal{A}(\overline{X}_t, \widehat{X}_t) - \mathcal{A}(\overline{Z}_t, \check{X}_t)\right] - P_t S(\overline{X}_t - \overline{Z}_t)\right\}\;dt + \left[P_t - \check{P}_t\right] S(X_t - \overline{Z}_t)\;dt + d\mathcal{M}_t$$

with the martingale

$$d\mathcal{M}_t := \left[P_t - \check{P}_t\right] B' R_2^{-1/2}\;d(V_t - \overline{V}_t)$$

$$\Rightarrow\;\; \partial_t\langle\mathcal{M}\rangle_t \;=\; \|\left[P_t - \check{P}_t\right] B' R_2^{-1/2}\|_F^2 = \mathrm{tr}\left(\left[P_t - \check{P}_t\right]^2 S\right) \leqslant \mathcal{V}_t := \rho(S)\;\|P_t - \check{P}_t\|_F^2.$$

When the initial random states $\overline{X}_0$ and $\overline{Z}_0$ are possibly different but they have the same first and second order statistics we have

$$\widehat{X}_0 = \check{X}_0 \quad\text{and}\quad P_0 = \check{P}_0 \quad\Longrightarrow\; \forall t \geqslant 0 \quad \widehat{X}_t = \check{X}_t \quad\text{and}\quad P_t = \check{P}_t.$$

In this particular situation we have

$$\mathcal{A}(\overline{X}_t, \widehat{X}_t) - \mathcal{A}(\overline{Z}_t, \check{X}_t) = \partial A(\check{X}_t)\;(\overline{X}_t - \overline{Z}_t)$$

and

$$\partial_t\left(\overline{X}_t - \overline{Z}_t\right) = \left[\partial A(\check{X}_t) - P_t S\right]\;(\overline{X}_t - \overline{Z}_t).$$

This implies that

$$\partial_t\|\overline{X}_t - \overline{Z}_t\|^2 = 2\langle(\overline{X}_t - \overline{Z}_t), \left[\partial A(\check{X}_t) - P_t S\right]\;(\overline{X}_t - \overline{Z}_t)\rangle \leqslant -\lambda_{\partial A}\;\|\overline{X}_t - \overline{Z}_t\|^2$$

This ends the proof of the first assertion.

More generally, we have

$$\mathcal{A}(\overline{X}_t, \widehat{X}_t) - \mathcal{A}(\overline{Z}_t, \check{X}_t)$$

$$= \partial A(\check{X}_t)\;(\overline{X}_t - \overline{Z}_t)$$

$$+\left[A(\widehat{X}_t) - A(\check{X}_t)\right] - \partial A(\check{X}_t)(\widehat{X}_t - \check{X}_t) + \left[\partial A(\widehat{X}_t) - \partial A(\check{X}_t)\right]\;(\overline{X}_t - \widehat{X}_t)$$

This yields the estimate

$$\langle \overline{X}_t - \overline{Z}_t, \Big(\mathcal{A}(\overline{X}_t, \widehat{X}_t) - \mathcal{A}(\overline{Z}_t, \check{X}_t)\Big) - P_t S(\overline{X}_t - \overline{Z}_t)\rangle$$

$$\leqslant -\frac{\lambda_{\partial A}}{2}\ \|\overline{X}_t - \overline{Z}_t\|^2 + \langle \overline{X}_t - \overline{Z}_t, \Big[\partial A(\widehat{X}_t) - \partial A(\check{X}_t)\Big]\ (\overline{X}_t - \widehat{X}_t)\rangle$$

$$+\langle \overline{X}_t - \overline{Z}_t, \Big[A(\widehat{X}_t) - A(\check{X}_t)\Big] - \partial A(\check{X}_t)(\widehat{X}_t - \check{X}_t)\rangle$$

$$\leqslant -\frac{\lambda_{\partial A}}{2}\ \|\overline{X}_t - \overline{Z}_t\|^2 + \ \|\widehat{X}_t - \check{X}_t\|\ \|\overline{X}_t - \overline{Z}_t\|\ \Big(\kappa_{\partial A}\ \|\overline{X}_t - \widehat{X}_t\| + \kappa_{\partial A} + \|\partial A\|\Big)$$

We also have

$$\langle \overline{X}_t - \overline{Z}_t, \Big[P_t - \check{P}_t\Big]\, S(X_t - \overline{Z}_t)\rangle \leqslant \|P_t - \check{P}_t\|_F\ \|\overline{X}_t - \overline{Z}_t\|\ \|S(X_t - \overline{Z}_t)\|$$

This implies that

$$d\|\overline{X}_t - \overline{Z}_t\|^2$$

$$\leqslant \Big[-\lambda_{\partial A}\ \|\overline{X}_t - \overline{Z}_t\|^2 + 2\ \|\widehat{X}_t - \check{X}_t\|\ \|\overline{X}_t - \overline{Z}_t\|\ \Big(\kappa_{\partial A}\ \|\overline{X}_t - \widehat{X}_t\| + \kappa_{\partial A} + \|\partial A\|\Big)\Big]\ dt$$

$$+\Big[2\|P_t - \check{P}_t\|_F\ \|\overline{X}_t - \overline{Z}_t\|\ \|S(X_t - \overline{Z}_t)\|\Big]\ dt + 2\sqrt{\mathcal{V}_t}\ \|\overline{X}_t - \overline{Z}_t\|\ d\overline{\mathcal{M}}_t$$

with

$$\mathcal{V}_t = \rho(S)\ \|P_t - \check{P}_t\|_F^2$$

and a rescaled continuous martingale $\overline{\mathcal{M}}_t$ such that $\partial_t\langle\overline{\mathcal{M}}\rangle_t \leqslant 1$. On the other hand, we have

$$2\ \|\overline{X}_t - \overline{Z}_t\|\ \|\widehat{X}_t - \check{X}_t\|\ \Big(\kappa_{\partial A}\ \|\overline{X}_t - \widehat{X}_t\| + \kappa_{\partial A} + \|\partial A\|\Big)$$

$$\leqslant \frac{\lambda_{\partial A}}{4}\|\overline{X}_t - \overline{Z}_t\|^2 + \frac{4}{\lambda_{\partial A}}\|\widehat{X}_t - \check{X}_t\|^2\ \Big(\kappa_{\partial A}\ \|\overline{X}_t - \widehat{X}_t\| + \kappa_{\partial A} + \|\partial A\|\Big)^2$$

and

$$2\|P_t - \check{P}_t\|_F\ \|\overline{X}_t - \overline{Z}_t\|\ \|S(X_t - \overline{Z}_t)\|$$

$$\leqslant \frac{\lambda_{\partial A}}{4}\|\overline{X}_t - \overline{Z}_t\|^2 + \frac{4}{\lambda_{\partial A}}\ \|P_t - \check{P}_t\|_F^2\ \|S(X_t - \overline{Z}_t)\|^2.$$

We conclude that

$$d\|\overline{X}_t - \overline{Z}_t\|^2 \leqslant \Big[-\frac{\lambda_{\partial A}}{2}\ \|\overline{X}_t - \overline{Z}_t\|^2 + \mathcal{U}_t\Big]\ dt + 2\sqrt{\mathcal{V}_t}\ \|\overline{X}_t - \overline{Z}_t\|\ d\overline{\mathcal{M}}_t$$

with

$$\mathcal{U}_t := \alpha_t\ \|\widehat{X}_t - \check{X}_t\|^2 + \beta_t\ \|P_t - \check{P}_t\|_F^2$$

and the parameters

$$\alpha_t := \frac{4}{\lambda_{\partial A}}\Big(\kappa_{\partial A}\ \|\overline{X}_t - \widehat{X}_t\| + \kappa_{\partial A} + \|\partial A\|\Big)^2 \quad \text{and} \quad \beta_t := \frac{4}{\lambda_{\partial A}}\ \|S(X_t - \overline{Z}_t)\|^2.$$

By (20) and (32), for any $\delta \leqslant 2^{-1}\sqrt{\lambda_S}$ and any $t \geqslant s$ we have

$$\begin{aligned}\mathbb{E}\left(\alpha_t^{\delta/4}\;\|\widehat{X}_t-\check{X}_t\|^{\delta/2}\mid\mathcal{F}_s\right)^{4/\delta} &\leqslant \mathbb{E}\left(\|\widehat{X}_t-\check{X}_t\|^{\delta}\mid\mathcal{F}_s\right)^{2/\delta}\;\mathbb{E}\left(\alpha_t^{\delta/2}\mid\mathcal{F}_s\right)^{2/\delta}\\ &\leqslant \overline{\mathcal{Z}}_s\;\exp\left(-\widehat{\lambda}_{\partial A}(1-\epsilon)(t-s)\right)\end{aligned}$$

for some process $\overline{\mathcal{Z}}_s$ satisfying the uniform moment condition (34). In the same vein we check that

$$\mathbb{E}\left(\mathcal{U}_t^{\delta/4}\mid\mathcal{F}_s\right)^{4/\delta}\vee\mathbb{E}\left(\mathcal{V}_t^{\delta/4}\mid\mathcal{F}_s\right)^{4/\delta}\leqslant\overline{\mathcal{Z}}_s\;\exp\left(-\widehat{\lambda}_{\partial A}(1-\epsilon)(t-s)\right)$$

for any $s \geqslant t_0$. By [14, Corollary 2.2] we have

$$\begin{aligned}&\mathbb{E}\left(\|\overline{X}_t-\overline{Z}_t\|^{\delta/4}\mid\mathcal{F}_s\right)^{8/\delta}\\ &\leqslant \exp\left(-\left[\frac{\lambda_{\partial A}}{2}(t-s)\right]\right)\;\|\overline{X}_s-\overline{Z}_s\|^2\\ &\qquad + n\;\overline{\mathcal{Z}}_s\int_s^t\;\exp\left(-\left[\frac{\lambda_{\partial A}}{2}(t-u)+\widehat{\lambda}_{\partial A}(1-\epsilon)(u-s)\right]\right)\,du\\ &\leqslant e^{-\frac{\lambda_{\partial A}}{2}(t-s)}\;\|\overline{X}_s-\overline{Z}_s\|^2+\frac{n\;\overline{\mathcal{Z}}_{t_0}}{|\widehat{\lambda}_{\partial A}(1-\epsilon)-\lambda_{\partial A}/2|}\;|e^{-\frac{\lambda_{\partial A}}{2}(t-s)}-e^{-\widehat{\lambda}_{\partial A}(1-\epsilon)(t-s)}|.\end{aligned}$$

The end of the proof of the theorem is now easily completed. ∎

5 Quantitative propagation of chaos estimates

5.1 Laplace exponential moment estimates

The analysis of EKF filters and their particle interpretation is mainly based on the estimation of the stochastic exponential function

$$\mathcal{E}_\Gamma(t) := \exp\left[\int_0^t \Gamma_A(s)\;ds\right]$$

with the stochastic functional

$$\Gamma_A(s) := -\left[\lambda_{\partial A}-\left(2\kappa_{\partial A}\;\mathrm{tr}(P_t)+\rho(S)\;\|X_t-\widehat{X}_t\|\;\right)\right].$$

Assume condition (20) is satisfied and set

$$\Lambda_{\partial A}\left[\epsilon,\delta\right]/\lambda_{\partial A} := 1-\frac{2}{\lambda_K\lambda_R}+\frac{1}{\lambda_S}\left(\frac{3}{4}-\delta\right)-\frac{1}{\delta}\;\frac{\epsilon\lambda_A}{2\lambda_{\partial A}}.$$

Observe that for any $\delta > 0$ we have

$$\epsilon=\frac{1}{2}\;\frac{\lambda_{\partial A}}{\lambda_A}\Longrightarrow\Lambda_{\partial A}\left[\epsilon,\sqrt{\lambda_S}/2\right]=\widehat{\lambda}_{\partial A}\geqslant\Lambda_{\partial A}\left[\epsilon,\delta\right].$$

The next technical lemma provides some key δ-exponential moments estimates. Its proof is quite technical, thus it is housed in the appendix, Section 6.2.

Lemma 5.1. • *For any $\delta > 0$ and any $0 \leqslant s \leqslant t$ we have the almost sure estimate*

$$\mathbb{E}\left((\mathcal{E}_\Gamma(t)/\mathcal{E}_\Gamma(s))^{-\delta} \mid \mathcal{F}_s\right)^{1/\delta} \leqslant \exp\left(\Lambda_\Gamma^-\ (t-s)\right) \quad \text{with} \quad \Lambda_\Gamma^- = \lambda_{\partial A}\left[1 - \frac{2}{\lambda_K \lambda_R}\right]. \tag{35}$$

- *For any $\epsilon \in [0,1]$, any $0 < \delta \leqslant e\ \epsilon\ \lambda_{R,S}$ and any initial covariance matrix P_0 such that*

$$\operatorname{tr}(P_0)^2 \leqslant \sigma^2(\epsilon,\delta) := \frac{\lambda_S}{\lambda_R}\left[\frac{1}{2} + \frac{1}{\lambda_R\lambda_S}\right](e\ \epsilon\lambda_{R,S}/\delta - 1)$$

  for any time horizon $t \geqslant 0$ we have the exponential δ-moment estimate

$$\mathbb{E}\left[\mathcal{E}_\Gamma(t)^\delta\right]^{1/\delta} \leqslant c_\delta(P_0)\ \exp\left[\Lambda_\Gamma^+(\epsilon,\delta)\ t\right] \tag{36}$$

  with the parameters

$$\begin{aligned} \Lambda_\Gamma^+(\epsilon,\delta) &:= 2\kappa_{\partial A}\sigma(\epsilon,\delta) - \Lambda_{\partial A}[\epsilon,\delta] - (\delta-1)\,\rho(S) \\ c_\delta(P_0) &:= \exp\left(1/\delta + \delta\chi(P_0)/(2\lambda_S)^2\right). \end{aligned}$$

- *For any $\epsilon \in]0,1]$ there exists some time horizon s such that for any $t \geqslant s$ and any $\delta \leqslant \sqrt{\lambda_S}/2$ we have the almost sure estimate*

$$\mathbb{E}\left(\mathcal{E}_\Gamma(t)^\delta \mid \mathcal{F}_s\right)^{1/\delta} \leqslant \mathcal{E}_\Gamma(s)\ \mathcal{Z}_s\ \exp\left(-\left\{(1-\epsilon)\widehat{\lambda}_{\partial A} + (\delta-1)\rho(S)\right\}\ (t-s)\right) \tag{37}$$

  for some positive random process $\mathcal{Z}_t$ s.t.

$$\forall \alpha \leqslant \lambda_{R,S}\ \sqrt{\lambda_S} \qquad \sup_{t\geqslant 0}\mathbb{E}\left(\mathcal{Z}_t^\alpha\right) < \infty$$

5.2 A non asymptotic convergence theorem

This section is mainly concerned with the estimation of the δ-moments of the square errors

$$\Xi_t := \|(m_t,p_t) - (\widehat{X}_t, P_t)\|^2 = \|m_t - \widehat{X}_t\|^2 + \|p_t - P_t\|_F^2$$

The analysis is based on a couple of technical lemmas.

The first one provides uniform moments estimates with respect to the time parameter.

Lemma 5.2. *There exists some $\nu > 0$ such that for any $1 \leqslant n \leqslant 1 + \nu N$ we have*

$$\sup_{t\geqslant 0}\mathbb{E}\left(\operatorname{tr}(p_t)^n\right) < \infty \qquad \sup_{t\geqslant 0}\mathbb{E}\left(\|\xi_t^1\|^n\right) < \infty \quad \text{and} \quad \sup_{t\geqslant 0}\mathbb{E}\left(\|\zeta_t^1\|^n\right) < \infty.$$

The second technical lemma provides a differential perturbation inequality in terms of the Laplace functionals discussed in Section 5.1.

Lemma 5.3. *We have the stochastic differential inequality*

$$d\Xi_t \quad \leqslant \quad \Xi_t\ \left[\Gamma_A(t) + \sqrt{2\rho(S)}\ d\Upsilon_t^{(1)}\right] + \left[\mathcal{V}_t\ dt + \sqrt{\mathcal{V}_t\ \Xi_t}\ d\Upsilon_t^{(2)}\right]$$

with a couple of orthogonal martingales s.t. $\partial_t\langle\Upsilon_\cdot^{(i)}, \Upsilon_\cdot^{(j)}\rangle_t \leqslant 1_{i=j}$ and some nonnegative process $\mathcal{V}_t$ such that

$$\sup_{t\geqslant 0}\mathbb{E}\left(\mathcal{V}_t^n\right)^{1/n} \leqslant c(n)/N \quad \text{for any } 1 \leqslant n \leqslant 1 + \nu N \text{ and some } \nu > 0.$$

The proofs of these two lemmas are rather technical thus they are provided in the appendix, Section 6.3 and Section 6.4. We are now in position to state and to prove the main result of this section.

Theorem 5.4. *Assume that* $(2^{-1}\sqrt{\lambda_S}) \wedge (e\lambda_{R,S}) \geqslant 2$. *In this situation, there exist some* $N_0 \geqslant 1$ *and some* $\alpha \in]0,1]$ *such that for any* $N_0 \leqslant N$, $1 \leqslant \delta \leqslant (4^{-1}\sqrt{\lambda_S}) \wedge (2^{-1}e\lambda_{R,S})$ *and any initial covariance matrix* P_0 *of the signal we have the uniform estimates*

$$\mathrm{tr}(P_0)^2 \leqslant \frac{1}{2}\,\frac{\lambda_S}{\lambda_R}\left[1+\frac{2}{\lambda_R\lambda_S}\right] \Longrightarrow \sup_{t\geqslant 0}\mathbb{E}[\Xi_t^{\delta}]^{1/\delta} \leqslant c/N^{\alpha}.$$

Proof. We set

$$\overline{\mathcal{E}}(t) := \mathcal{E}_{\Gamma}(t)\mathcal{E}_{\Upsilon}(t) = e^{\mathcal{L}_t}$$

with the exponential martingale

$$\mathcal{E}_{\Upsilon}(t) := \exp\left[\sqrt{2\rho(S)}\ \Upsilon_t^{(1)} - \rho(S)t\right]$$

and the stochastic process

$$\mathcal{L}_t := \int_0^t \Gamma_A(u)\ du + \sqrt{2\rho(S)}\ \Upsilon_t^{(1)} - \rho(S)t.$$

Observe that for any $\delta \geqslant 0$ we have

$$\mathcal{E}_{\Upsilon}^{-\delta}(t) \quad = \quad \exp\left[-\delta\sqrt{2\rho(S)}\ \Upsilon_t^{(1)} + \delta\rho(S)t\right] = \exp\left[\delta(1+2\delta)\rho(S)t\right]\ \mathcal{E}_{-2\delta\Upsilon}^{1/2}(t)$$

with the exponential martingale

$$\mathcal{E}_{-2\delta\Upsilon}(t) := \exp\left[-2\delta\sqrt{2\rho(S)}\ \Upsilon_t^{(1)} - 4\delta^2\rho(S)t\right].$$

In the same vein we have

$$\mathcal{E}_{\Upsilon}^{\delta}(t) \quad = \quad \exp\left[\delta\sqrt{2\rho(S)}\ \Upsilon_t^{(1)} - \delta\rho(S)t\right] = \exp\left[\delta(2\delta-1)\rho(S)t\right]\ \mathcal{E}_{2\delta\Upsilon}^{1/2}(t)$$

with the exponential martingale

$$\mathcal{E}_{2\delta\Upsilon}(t) := \exp\left[2\delta\sqrt{2\rho(S)}\ \Upsilon_t^{(1)} - 4\delta^2\rho(S)t\right].$$

This yields the estimates

$$\begin{aligned}
\mathbb{E}\left(\overline{\mathcal{E}}^{-\delta}(t)\right) &= \exp\left(\delta(1+2\delta)\rho(S)t\right)\,\mathbb{E}\left[\mathcal{E}_{\Gamma}(t)^{-\delta}\ \mathcal{E}_{-2\delta\Upsilon}^{1/2}(t)\right]\\
&\leqslant \mathbb{E}\left[\mathcal{E}_{\Gamma}(t)^{-2\delta}\right]^{1/2}\ \exp\left(\delta(1+2\delta)\ \rho(S)t\right)\\
\mathbb{E}\left(\overline{\mathcal{E}}^{\delta}(t)\right) &\leqslant \mathbb{E}\left[\mathcal{E}_{\Gamma}(t)^{2\delta}\right]^{1/2}\ \exp\left(\delta(2\delta-1)\ \rho(S)t\right).
\end{aligned}$$

Using (35) and (36) we find the estimates

$$\mathbb{E}\left(\overline{\mathcal{E}}^{-\delta}(t)\right)^{1/\delta} \leqslant \exp\left(\left[(1+2\delta)\ \rho(S) + \Lambda_{\Gamma}^{-}\right]t\right) \tag{38}$$

$$\mathbb{E}\left(\overline{\mathcal{E}}^{\delta}(t)\right)^{1/\delta} \leqslant c_{\delta}(P_0)\ \exp\left(\left[(2\delta-1)\ \rho(S) + \Lambda_{\Gamma}^{+}(\epsilon,\delta)\right]t\right). \tag{39}$$

The estimate (39) is valid for any $\epsilon \in [0,1]$ and any

$$\delta \leqslant e\ \epsilon\ \lambda_{R,S} \quad \text{and} \quad \mathrm{tr}(P_0) \leqslant \sigma(\epsilon,\delta).$$

Using the fact that

$$\begin{aligned} d\overline{\mathcal{E}}^{-1}(t) &\leqslant -e^{-\mathcal{L}_t}\ \left(\Gamma_A(t)\ dt + \sqrt{2\rho(S)}\ d\Upsilon_t^{(1)} - \rho(S)dt\right) + \frac{1}{2}\ e^{-\mathcal{L}_t}\ 2\rho(S)\ \partial_t\langle\Upsilon^{(1)}\rangle_t\ dt \\ &\leqslant -\overline{\mathcal{E}}^{-1}(t)\ \left(\Gamma_A(t)\ dt + \sqrt{2\rho(S)}\ d\Upsilon_t^{(1)}\right) \end{aligned}$$

we find the stochastic inequality

$$\begin{aligned} d(\Xi_t\ \overline{\mathcal{E}}^{-1}(t)) &\leqslant \overline{\mathcal{E}}^{-1}(t)\ d\Xi_t + \Xi_t\ d\overline{\mathcal{E}}^{-1}(t) - 2\overline{\mathcal{E}}^{-1}(t)\ \Xi_t\ \rho(S)\ dt \\ &\leqslant \overline{\mathcal{E}}^{-1}(t)\ \Xi_t\ \left[\Gamma_A(t) + \sqrt{2\rho(S)}\ d\Upsilon_t^{(1)}\right] + \overline{\mathcal{E}}^{-1}(t)\left[\mathcal{V}_t\ dt + \sqrt{\mathcal{V}_t\ \Xi_t}\ d\Upsilon_t^{(2)}\right] \\ &\quad -\overline{\mathcal{E}}^{-1}(t)\ \Xi_t\ \left[\Gamma_A(t)\ dt + \sqrt{2\rho(S)}\ d\Upsilon_t^{(1)}\right] - 2\overline{\mathcal{E}}^{-1}(t)\ \Xi_t\ \rho(S)\ dt \\ &= \overline{\mathcal{E}}^{-1}(t)\left[(\mathcal{V}_t - 2\ \Xi_t\ \rho(S))\ dt + \sqrt{\mathcal{V}_t\ \Xi_t}\ d\Upsilon_t^{(2)}\right]. \end{aligned}$$

For any $\delta \geqslant 2$, this implies that

$$\begin{aligned} d(\Xi_t\ \overline{\mathcal{E}}^{-1}(t))^{\delta} &\leqslant \delta\ \Xi_t^{\delta-1}\ \overline{\mathcal{E}}^{-\delta}(t)\ \left[(\mathcal{V}_t - 2\ \Xi_t\ \rho(S))\ dt + \sqrt{\mathcal{V}_t\ \Xi_t}\ d\Upsilon_t^{(2)}\right] \\ &\quad +\delta\ \Xi_t^{\delta-1}\ \overline{\mathcal{E}}(t)^{-\delta}\ \frac{(\delta-1)}{2}\mathcal{V}_t\ dt \\ &= \delta\ \Xi_t^{\delta-1}\ \overline{\mathcal{E}}^{-\delta}(t)\ \left[\left(\frac{\delta+1}{2}\ \mathcal{V}_t - 2\ \Xi_t\ \rho(S)\right)\ dt + \sqrt{\mathcal{V}_t\ \Xi_t}\ d\Upsilon_t^{(2)}\right]. \end{aligned}$$

Taking the expectation we obtain

$$\begin{aligned} \partial_t\mathbb{E}\left[(\Xi_t\ \overline{\mathcal{E}}^{-1}(t))^{\delta}\right] &\leqslant \frac{\delta(\delta+1)}{2}\ \mathbb{E}\left[\ \left(\Xi_t\ \overline{\mathcal{E}}^{-1}(t)\right)^{\delta-1}\ \overline{\mathcal{E}}^{-1}(t)\ \mathcal{V}_t\right] \\ &\quad - 2\delta\ \rho(S)\ \mathbb{E}\left[\left(\Xi_t\ \overline{\mathcal{E}}^{-1}(t)\right)^{\delta}\right]. \end{aligned}$$

On the other hand using Lemma 5.3 and the Laplace estimate (38) we have

$$\begin{aligned} &\mathbb{E}\left(\left(\Xi_t\ \overline{\mathcal{E}}^{-1}(t)\right)^{\delta-1}\ \overline{\mathcal{E}}^{-1}(t)\ \mathcal{V}_t\right) \\ &\leqslant \mathbb{E}\left(\left(\Xi_t\ \overline{\mathcal{E}}^{-1}(t)\right)^{\delta}\right)^{1-1/\delta}\ \mathbb{E}\left(\overline{\mathcal{E}}^{-2\delta}(t)\right)^{1/(2\delta)}\ \mathbb{E}\left(\mathcal{V}_t^{2\delta}\right)^{1/(2\delta)} \\ &\leqslant \frac{c}{N}\ \exp\left(\left[(1+4\delta)\ \rho(S) + \Lambda_\Gamma^-\right]t\right)\ \mathbb{E}\left(\left(\Xi_t\ \overline{\mathcal{E}}^{-1}(t)\right)^{\delta}\right)^{1-1/\delta}. \end{aligned}$$

This yields

$$
\begin{aligned}
&\partial_t \mathbb{E}\left(\left(\Xi_t \overline{\mathcal{E}}^{-1}(t)\right)^{\delta}\right)^{1/\delta} \\
&\leqslant \frac{1}{\delta}\ \mathbb{E}\left(\left(\Xi_t \overline{\mathcal{E}}^{-1}(t)\right)^{\delta}\right)^{\frac{1}{\delta}-1}\ \partial_t \mathbb{E}\left(\left(\Xi_t \overline{\mathcal{E}}^{-1}(t)\right)^{\delta}\right) \\
&\leqslant -2\rho(S)\mathbb{E}\left(\left(\Xi_t \overline{\mathcal{E}}^{-1}(t)\right)^{\delta}\right)^{1/\delta} + \frac{(\delta+1)}{2}\frac{c}{N}\ \exp\left(\left[(1+4\delta)\rho(S)+\Lambda_{\Gamma}^{-}\right]t\right)
\end{aligned}
$$

from which we conclude that

$$
\begin{aligned}
\mathbb{E}\left(\left(\Xi_t\ \overline{\mathcal{E}}^{-1}(t)\right)^{\delta}\right)^{1/\delta} &\leqslant \exp\left\{-2\rho(S)t\right\}\ \mathbb{E}\left(\Xi_0^{\delta}\right)^{1/\delta} + \frac{c}{N}\ \exp\left\{\left((1+4\delta)\rho(S)+\Lambda_{\Gamma}^{-}\right)t\right\} \\
&\leqslant \frac{c}{N}\ \exp\left\{\left((1+4\delta)\rho(S)+\Lambda_{\Gamma}^{-}\right)t\right\}.
\end{aligned}
$$

By Cauchy-Schwarz inequality we also have

$$
\mathbb{E}\left(\Xi_t^{\delta/2}\right)^{2/\delta} = \mathbb{E}\left(\overline{\mathcal{E}}(t)^{\delta/2}\ \left(\Xi_t\ \overline{\mathcal{E}}^{-1}(t)\right)^{\delta/2}\right)^{2/\delta} \leqslant \mathbb{E}\left(\left(\Xi_t\ \overline{\mathcal{E}}^{-1}(t)\right)^{\delta}\right)^{1/\delta} \mathbb{E}\left(\overline{\mathcal{E}}^{\delta}(t)\right)^{1/\delta}.
$$

Using (39) we conclude that for any $\epsilon \in [0,1]$ and any $\delta \leqslant e\ \epsilon\ \lambda_{R,S}$ and $\mathrm{tr}(P_0) \leqslant \sigma(\epsilon,\delta)$,

$$
\mathbb{E}\left(\Xi_t^{\delta/2}\right)^{2/\delta} \leqslant c_{\delta}(P_0)\ \frac{c}{N}\ \exp\left\{\left(6\delta\rho(S)+\Lambda_{\Gamma}^{-}+\Lambda_{\Gamma}^{+}(\epsilon,\delta)\right)t\right\}. \tag{40}
$$

On the other hand, by [14, Theorem 2.1] for any $\delta \geqslant 1$ we also have

$$
\begin{aligned}
\mathbb{E}\left(\Xi_t^{\delta/2} \mid \mathcal{F}_s\right)^{2/\delta} &\leqslant \mathbb{E}\left[\exp\left(\delta\ \int_s^t \left\{\Gamma_A(u)\ +(\delta-1)\rho(S)\right\}\ du\right) \mid \mathcal{F}_s\right]^{1/\delta} \\
&\quad \times \left\{\Xi_s + \frac{1}{N}\ \frac{\delta+1}{2}\ \int_s^t\ \mathbb{E}\left[\overline{\mathcal{V}}_u^{\delta} \mid \mathcal{F}_s\right]^{1/\delta}\ du\right\}
\end{aligned}
\tag{41}
$$

with the rescaled process

$$
\overline{\mathcal{V}}_t := \exp\left(\int_s^t \left[-\Gamma_A(u)+2(1-\delta)\rho(S)\right]du\right)\ \mathcal{V}_t
$$

of the process $\mathcal{V}_t$ defined in Lemma 5.3.

On the other hand using (37) for any $\epsilon \in]0,1]$ there exists some time horizon $s = s(\epsilon)$ such that for any $t \geqslant s$ and any $\delta \leqslant \frac{1}{2}\ \sqrt{\lambda_S}$ we have the almost sure estimate

$$
\mathbb{E}\left(\Xi_t^{\delta/2} \mid \mathcal{F}_s\right)^{2/\delta} \leqslant \mathcal{Z}_s\ \exp\left(-(1-\epsilon)\,\widehat{\lambda}_{\partial A}(t-s)\right) \times \left\{\Xi_s + \frac{\delta+1}{2}\ \int_s^t\ \mathbb{E}\left[\overline{\mathcal{V}}_u^{\delta} \mid \mathcal{F}_s\right]^{1/\delta}\ du\right\}
$$

with some process $\mathcal{Z}_s$ such that

$$
\sup_{t\geqslant 0} \mathbb{E}\left(\mathcal{Z}_t^{\alpha}\right) < \infty \quad \text{for any} \quad \alpha \leqslant \frac{2}{e}\ \sqrt{\lambda_S}\ \left(\leqslant \frac{1}{2}\ \lambda_{R,S}\ \sqrt{\lambda_S}\right).
$$

Combining Cauchy-Schwarz inequality with (35) and Lemma 5.3 we readily check that

$$
\begin{aligned}
&\mathbb{E}\left[\overline{\mathcal{V}}_u^{\delta} \mid \mathcal{F}_s\right]^{1/\delta} \\
&= \mathbb{E}\left[\mathcal{V}_u^{\delta}\ \exp\left(\delta\int_s^u \left[-\Gamma_A(v)+2(1-\delta)\rho(S)\right]dv\right) \mid \mathcal{F}_s\right]^{1/\delta} \\
&\leqslant \mathbb{E}\left[\mathcal{V}_u^{2\delta}\right]^{1/(2\delta)}\ \exp\left(2(1-\delta)\rho(S)(u-s)\right)\ \mathbb{E}\left[\left(\mathcal{E}_\Gamma(u)/\mathcal{E}_\Gamma(s)\right)^{-2\delta} \mid \mathcal{F}_s\right]^{1/(2\delta)} \\
&\leqslant \frac{c}{N}\ \exp\left[\left(2(1-\delta)\rho(S)+\Lambda_\Gamma^-\right)(u-s)\right].
\end{aligned}
$$

This yields the estimate

$$
\begin{aligned}
\mathbb{E}\left(\Xi_t^{\delta/2} \mid \mathcal{F}_s\right) \leqslant \mathcal{Z}_s^{\delta/2}\ \exp\left(-\frac{\delta}{2}\ (1-\epsilon)\ \widehat{\lambda}_{\partial A}\ (t-s)\right) \\
\times\left\{\Xi_s+\frac{c}{N}\ \exp\left[\left(2(1-\delta)\rho(S)+\Lambda_\Gamma^-\right)(t-s)\right]\right\}^{\delta/2}.
\end{aligned}
$$

This implies that for any $1\leqslant\delta/2\leqslant\frac{1}{4}\ \sqrt{\lambda_S}$ we have

$$
\begin{aligned}
\mathbb{E}\left(\Xi_t^{\delta/2} \mid \mathcal{F}_s\right) \leqslant\ c\ \mathcal{Z}_s^{\delta/2}\ \exp\left(-\frac{\delta}{2}\ (1-\epsilon)\ \widehat{\lambda}_{\partial A}\ (t-s)\right) \\
\times\left\{\Xi_s^{\delta/2}+\frac{1}{N^{\delta/2}}\ \exp\left[\frac{\delta}{2}\ \left(2(1-\delta)\rho(S)+\Lambda_\Gamma^-\right)(t-s)\right]\right\}.
\end{aligned}
$$

Taking the expectation and choosing $\epsilon\leqslant 1/2$, there exists some time horizon t_0 such that for any $s\geqslant 0$ and any $\tau\geqslant s+t_0$

$$
\begin{aligned}
&\mathbb{E}\left(\Xi_\tau^{\delta/2}\right)^{2/\delta} \\
&\leqslant c\ \exp\left(-\widehat{\lambda}_{\partial A}\ (\tau-(s+t_0))/2\right)\left\{1+\frac{1}{N}\ \exp\left[\left(2(1-\delta)\rho(S)+\Lambda_\Gamma^-\right)(\tau-(s+t_0)\right]\right\}
\end{aligned}
$$

for any $2\leqslant\delta\leqslant 1+\nu N$ for some $\nu>0$, and for some finite constant $c(\delta)<\infty$. This implies that for any time horizon $t\geqslant 0$ and any

$$2\leqslant\delta\leqslant 2^{-1}\sqrt{\lambda_S}\wedge(1+\nu N)$$

we have

$$
\mathbb{E}\left(\Xi_{s+t_0+t}^{\delta/2}\right)^{2/\delta}\leqslant c\exp\left(-\frac{\widehat{\lambda}_{\partial A}}{2}\ t\right)\left\{1+\frac{1}{N}\ \exp\left[\left(2(1-\delta)\rho(S)+\Lambda_\Gamma^-\right)t\right]\right\}.
$$

This yields the uniform estimates

$$
\sup_{u\in[t+t_0,\infty[}\mathbb{E}\left(\Xi_u^{\delta/2}\right)^{2/\delta}=\sup_{s\geqslant 0}\mathbb{E}\left(\Xi_{s+t_0+t}^{\delta/2}\right)^{2/\delta}\leqslant c\ \left\{\exp\left[-\frac{\widehat{\lambda}_{\partial A}}{2}\ t\right]+\frac{1}{N}\ \exp\left[\lambda_\Gamma t\right]\right\}
$$

with the parameters

$$\lambda_\Gamma \quad := \quad \Lambda_\Gamma^- - 2\rho(S) = \frac{\lambda_{\partial A}}{2}\left[\left(1-\frac{4}{\lambda_K\lambda_R}\right)+\left(1-\frac{4}{\lambda_S}\right)\right] > 0.$$

On the other hand, by (40) for any time horizon $t \geqslant 0$ and any $\delta \leqslant e\ \lambda_{R,S}$ and any P_0 s.t. $\mathrm{tr}(P_0) \leqslant \sigma(1, e\lambda_{R,S}/2)$ we have the uniform estimates

$$\sup_{s\in[0,t_0+t]} \mathbb{E}\left(\Xi_s^{\delta/2}\right)^{2/\delta} \quad \leqslant \quad c\ \frac{1}{N}\ \exp\left[\lambda'_\Gamma t\right]$$

with

$$\lambda'_\Gamma \quad := \quad 5e\lambda_{\partial A}\ \lambda_{R,S}/\lambda_S + \Lambda_\Gamma^- + \Lambda_\Gamma^+(1, e\lambda_{R,S}/2).$$

We conclude that for any time horizon $t \geqslant 0$

$$\sup_{s\geqslant 0}\mathbb{E}\left(\Xi_s^{\delta/2}\right)^{2/\delta} \leqslant c\ \left\{\ \exp\left[-\frac{\widehat{\lambda}_{\partial A}}{2}\ t\right] + \frac{1}{N}\ \exp\left[(\lambda_\Gamma \vee \lambda'_\Gamma)t\right]\right\}.$$

Choosing $t = t(N)$ such that

$$t = t(N) := \log N/\left\{\widehat{\lambda}_{\partial A}/2 + (\lambda_\Gamma \vee \lambda'_\Gamma)\right\},$$

we conclude that

$$\sup_{s\geqslant 0}\mathbb{E}\left(\Xi_s^{\delta/2}\right)^{2/\delta} \leqslant c\ N^{-\alpha} \quad \text{with} \quad \alpha = \frac{\widehat{\lambda}_{\partial A}}{\widehat{\lambda}_{\partial A} + 2\left(\lambda_\Gamma \vee \lambda'_\Gamma\right)} \in]0,1].$$

This ends the proof of the theorem. ∎

Corollary 5.5. *Assume that* $(4^{-1}\sqrt{\lambda_S}) \wedge (2^{-1}e\lambda_{R,S}) \geqslant 2$. *In this situation, there exists some* $N_0 \geqslant 1$ *and some* $\alpha \in]0,1]$ *such that for any* $N_0 \leqslant N$ *and any initial covariance matrix* P_0 *of the signal*

$$\mathrm{tr}(P_0)^2 \leqslant \frac{1}{2}\ \frac{\lambda_S}{\lambda_R}\left[1+\frac{2}{\lambda_R\lambda_S}\right] \Longrightarrow \sup_{t\geqslant 0}\mathbb{E}\left(\|\xi_t^1 - \zeta_t^1\|^2\right) \leqslant c(P_0)/N^\alpha$$

for some finite constant $c(P_0) < \infty$ *whose value depends on* P_0.

Proof. Using (4) we have

$$\begin{aligned} d(\xi_t^1 - \zeta_t^1) \quad &= \quad \left[(\partial A(m_t) - p_t S)\ \xi_t^1 + p_t S X_t + A(m_t) - \partial A(m_t)\ m_t\right]\ dt \\ &\quad - \left[(\partial A(\widehat{X}_t) - P_t S)\ \zeta_t^1 + P_t S X_t + A(\widehat{X}_t) - \partial A(\widehat{X}_t)\ \widehat{X}_t\right]\ dt + d\mathcal{M}_t \end{aligned}$$

with the martingale

$$d\mathcal{M}_t := (p_t - P_t)B' R_2^{-1/2} d(V_t - \overline{V}_t^1).$$

This yields

$$
\begin{aligned}
& d(\xi^1_t-\zeta^1_t)\\
&= \Big[(\partial A(m_t)-p_tS)\ (\xi^1_t-\zeta^1_t)+(p_t-P_t)S(X_t-\zeta^1_t)+(\partial A(m_t)-\partial A(\widehat{X}_t))\ \zeta^1_t\Big]\ dt\\
&+\Big[\Big(A(m_t)-A(\widehat{X}_t)\Big)+\Big(\partial A(\widehat{X}_t)-\partial A(m_t)\Big)\ m_t+\partial A(\widehat{X}_t)\ (\widehat{X}_t-m_t)\ \Big]\ dt+d\mathcal{M}_t
\end{aligned}
$$

with

$$
\sum_{1\leqslant k\leqslant r_1}\partial_t\langle \mathcal{M}(k),\mathcal{M}(k)\rangle_t\leqslant 2\rho(S)\ \|\ p_t-P_t\|_F^2.
$$

This implies that

$$
\begin{aligned}
& d\|\xi^1_t-\zeta^1_t\|^2\\
&\leqslant 2\langle \xi^1_t-\zeta^1_t,d(\xi^1_t-\zeta^1_t)\rangle+2\rho(S)\ \|\ p_t-P_t\|_F^2\ dt\\
&\leqslant \big\{-\lambda_{\partial A}\|\xi^1_t-\zeta^1_t\|^2+2\rho(S)\ \|\ p_t-P_t\|_F^2\ \ +2\|\xi^1_t-\zeta^1_t\|\\
&\times\Big[\|p_t-P_t\|_F\ \|S(X_t-\zeta^1_t)\|+\big(\kappa_{\partial A}\ \ \big(\|\zeta^1_t\|+\|m_t\|\big)+2\|\partial A\|\big)\ \ \|m_t-\widehat{X}_t\|\Big]\Big\}\ dt+d\overline{\mathcal{M}}_t
\end{aligned}
$$

with the martingale

$$
d\overline{\mathcal{M}}_t=2\langle \xi^1_t-\zeta^1_t,d\mathcal{M}_t\rangle.
$$

Notice that

$$
\begin{aligned}
& 2\rho(S)\ \|\ p_t-P_t\|_F^2\\
&+2\|\xi^1_t-\zeta^1_t\|\Big[\|p_t-P_t\|_F\ \|S(X_t-\zeta^1_t)\|+\big(\kappa_{\partial A}\ \ \big(\|\zeta^1_t\|+\|m_t\|\big)+2\|\partial A\|\big)\ \ \|m_t-\widehat{X}_t\|\Big]\\
&\leqslant \frac{\lambda_{\partial A}}{2}\ \|\xi^1_t-\zeta^1_t\|^2\times\epsilon_t
\end{aligned}
$$

with the process

$$
\begin{aligned}
& \epsilon_t:=2\rho(S)\ \|\ p_t-P_t\|_F^2\\
&+4\Big[\|p_t-P_t\|_F^2\ \|S(X_t-\zeta^1_t)\|^2+\big(\kappa_{\partial A}\ \ \big(\|\zeta^1_t\|+\|m_t\|\big)+2\|\partial A\|\big)^2\ \ \|m_t-\widehat{X}_t\|^2\Big]/\lambda_{\partial A}.
\end{aligned}
$$

By Theorem 5.4 we have

$$
\sup_{t\geqslant 0}\mathbb{E}(\epsilon_t)\leqslant c(P_0)/N^\alpha
$$

as soon as $(4^{-1}\sqrt{\lambda_S})\wedge(2^{-1}e\lambda_{R,S})\geqslant 2$ and initial covariance matrix P_0 of the signal is chosen so that

$$
\mathrm{tr}(P_0)^2\leqslant\frac{1}{2}\ \frac{\lambda_S}{\lambda_R}\left[1+\frac{2}{\lambda_R\lambda_S}\right].
$$

This implies that

$$\partial_t \mathbb{E}\left(\|\xi_t^1 - \zeta_t^1\|^2\right) \leqslant -\frac{\lambda_{\partial A}}{2}\ \mathbb{E}\left(\|\xi_t^1 - \zeta_t^1\|^2\right) + c(P_0)/N^{\alpha}.$$

The end of the proof of the corollary is now a direct consequence of Gronwall lemma. ■

6 Appendix

6.1 Regularity conditions

Notice that for any $\alpha, x \geqslant 0$ we have

$$\frac{x}{1+1/x} > 2\alpha \quad\Longleftrightarrow\quad x > \alpha\left(1+\sqrt{1+2/\alpha}\right)$$

and by (20)

$$\lambda_{R,S} > (8e)^{-1}\ \lambda_R\ \sqrt{\lambda_S}\ \left[1+\frac{1}{\lambda_R\sqrt{\lambda_S}}\right]^{-1}.$$

This shows that

$$(8e)^{-1}\ \frac{\lambda_R\ \sqrt{\lambda_S}}{\left[1+\frac{1}{\lambda_R\sqrt{\lambda_S}}\right]} > \alpha \quad\Longleftrightarrow\quad \lambda_R\ \sqrt{\lambda_S} > 4\ \alpha\ e\ \left(1+\sqrt{1+1/(2\alpha e)}\right) \Longrightarrow \lambda_{R,S} > \alpha.$$

Also observe that

$$\lambda_S > 4 \quad \text{and} \quad \lambda_R\ > 2\ \alpha\ e\ \left(1+\sqrt{1+1/(2\alpha e)}\right) \Longrightarrow \lambda_{R,S} > \alpha.$$

This yields the sufficient condition

$$(\lambda_K\lambda_R)\wedge\lambda_S > 4 \quad \text{and} \quad \lambda_R\ \sqrt{\lambda_S} > 4e\ \left(1+\sqrt{1+1/(2e)}\right) \Longrightarrow (20).$$

Also observe that for any $\alpha \geqslant 1$ we have

$$(\lambda_K/\alpha)\wedge(\lambda_S/4) > 1 \quad \text{and} \quad \lambda_R\ > 2\ \alpha\ e\ \left(1+\sqrt{1+1/(2\alpha e)}\right)$$

$$\Longrightarrow (\lambda_K\lambda_R/4)\wedge(\lambda_{R,S}/\alpha)\wedge(\lambda_S/4) > 1.$$

We end this section with the proof of (21). Whenever $\rho(S) = 1$ condition (20) takes the form

$$\lambda_{\partial A} > 4 \qquad \lambda_{\partial A}^2 > 4\ \kappa_{\partial A}\ \mathrm{tr}(R) \quad \text{and} \quad \lambda_{\partial A}^{3+1/2} > 4^2 e\left[\mathrm{tr}(R)^2 + \frac{1}{2}\ \mathrm{tr}(R)\ \lambda_{\partial A}^2\right]$$

The r.h.s. inequality can be restated as

$$\left(\frac{\lambda_{\partial A}^2}{2}\right)^2\left(1+\frac{1}{4e}\ \frac{1}{\sqrt{\lambda_{\partial A}}}\right) > \left(\mathrm{tr}(R)+\frac{\lambda_{\partial A}^2}{2}\right)^2$$

which is equivalent to

$$\operatorname{tr}(R) < \left(\frac{\lambda_{\partial A}^2}{2}\right)\left[\left(1+\frac{1}{4e}\,\frac{1}{\sqrt{\lambda_{\partial A}}}\right)^{1/2} - 1\right]$$

This ends the proof of the sufficient condition (21). ■

6.2 Proof of Lemma 5.1

We have

$$-\Gamma_A(t) \;=\; \lambda_{\partial A} - \left(2\kappa_{\partial A}\,\operatorname{tr}(P_t) + \rho(S)\,\|X_t - \widehat{X}_t\|\,\right) \leqslant \lambda_{\partial A}\left[1 - 2/(\lambda_K\lambda_R)\right]$$

The end of the proof of (35) is now clear. Observe that

$$\begin{aligned}
\mathcal{E}_\Gamma(t)^\delta &= \exp\left[\delta\int_0^t \left[\left(2\kappa_{\partial A}\,\operatorname{tr}(P_s) + \rho(S)\,\|X_s - \widehat{X}_s\|\,\right) - \lambda_{\partial A}\right]\,ds\right]\\
&\leqslant \exp\left[\delta\lambda_{\partial A}\left[\frac{2}{\lambda_K}\,\left(\operatorname{tr}(P_0) + \frac{1}{\lambda_R}\right) - 1\right]\,t\right]\,\exp\left[\delta\,\rho(S)\,\int_0^t \|X_s - \widehat{X}_s\|\,ds\right].
\end{aligned}$$

We let $\phi_t(x) = X_t$ be the stochastic flow of signal starting at $X_0 = x$. We recall the contraction inequality

$$\|\phi_t(x) - \phi_t(y)\| \leqslant \exp\left(-\lambda_{\partial A}t/2\right)\,\|x - y\|. \tag{42}$$

A proof of (42) can be found in [14, Section 3.1]. This inequality implies that

$$\begin{aligned}
\int_0^t \|X_r - \widehat{X}_r\|\,dr &= \int_0^t \|\phi_r(X_0) - \widehat{X}_r\|\,dr\\
&\leqslant \int_0^t \|\phi_r(X_0) - \phi_r(\widehat{X}_0)\|\,dr + \int_0^t \|\phi_r(\widehat{X}_0) - \widehat{X}_r\|\,dr\\
&\leqslant \left(\int_0^t e^{-\lambda_{\partial A}r/2}\,dr\right)\,\|X_0 - \widehat{X}_0\| + \int_0^t \|\phi_r(\widehat{X}_0) - \widehat{X}_r\|\,dr\\
&\leqslant 2\|X_0 - \widehat{X}_0\|/\lambda_{\partial A} + \int_0^t \|\phi_r(\widehat{X}_0) - \widehat{X}_r\|\,dr.
\end{aligned}$$

This implies that

$$\exp\left[\delta\,\rho(S)\,\int_0^t \|X_s - \widehat{X}_s\|\,ds\right] \leqslant \exp\left[2\delta\|X_0 - \widehat{X}_0\|/\lambda_S\right]\,\exp\left[\delta\rho(S)\,\int_0^t \|\phi_s(\widehat{X}_0) - \widehat{X}_s\|\,ds\right].$$

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

$$\int_0^t \left(\left(\|\phi_u(\widehat{X}_0) - \widehat{X}_u\| - 1/4\right) + 1/4\right)\,du \leqslant t/4 + \int_0^t \|\phi_r(\widehat{X}_0) - \widehat{X}_r\|^2 dr.$$

We find that

$$\exp\left[\delta\rho(S)\int_0^t \|X_s-\widehat{X}_s\|\;ds\right] \leqslant \exp\left[2\delta\|X_0-\widehat{X}_0\|/\lambda_S\right]\;\exp\left(\delta t\rho(S)/4\right) \times\exp\left[\delta\rho(S)\int_0^t\;\|\phi_s(\widehat{X}_0)-\widehat{X}_s\|^2\;ds\right].$$

This yields

$$\mathbb{E}\left[\exp\left[\delta\;\rho(S)\int_0^t \|X_s-\widehat{X}_s\|\;ds\right]\mid X_0\right] \leqslant \exp\left(t\delta\rho(S)/4\right)\exp\left[2\delta\|X_0-\widehat{X}_0\|/\lambda_S\right] \times\mathbb{E}\left[\exp\left[\delta\rho(S)\int_0^t\;\|\phi_s(\widehat{X}_0)-\widehat{X}_s\|^2\;ds\right]\right).$$

We also have the series of inequalities

$$\frac{1}{\rho(S)}\;\frac{1}{1+\pi_{\partial A}(0)}\;\frac{\lambda_A^2}{4\mathrm{tr}(R)}$$

$$\geqslant\frac{1}{\rho(S)}\;\frac{\lambda_{\partial A}^2}{4}\;\frac{1}{1+\pi_{\partial A}(0)}\;\frac{1}{4\mathrm{tr}(R)}\geqslant\frac{\lambda_S\lambda_R}{2\times 4^2}\;\frac{1}{1/2+\mathrm{tr}(P_0)^2(\rho(S)/\mathrm{tr}(R))+\rho(S)\mathrm{tr}(R)/\lambda_{\partial A}^2}$$

$$=\frac{1}{4^2}\;\lambda_S\;\lambda_R\;\left(2\;\frac{\lambda_R}{\lambda_S}\;\mathrm{tr}(P_0)^2+\left[1+\frac{2}{\lambda_R\lambda_S}\right]\right)^{-1}$$

$$\geqslant\frac{e}{8e}\;\sqrt{\lambda_S}\;\lambda_R\;\left[1+\frac{2}{\lambda_R\lambda_S}\right]^{-1}\;\left(1+2\;\frac{\lambda_R}{\lambda_S}\;\mathrm{tr}(P_0)^2\left[1+\frac{2}{\lambda_R\lambda_S}\right]^{-1}\right)^{-1}$$

$$=e\;\lambda_{R,S}\;\left(1+2\;\frac{\lambda_R}{\lambda_S}\;\mathrm{tr}(P_0)^2\left[1+\frac{2}{\lambda_R\lambda_S}\right]^{-1}\right)^{-1}.$$

This shows that

$$\delta\rho(S)\leqslant\frac{\epsilon}{1+\pi_{\partial A}(0)}\;\frac{\lambda_A^2}{4\mathrm{tr}(R)}$$

for some $\epsilon\in[0,1]$ as soon as

$$\mathrm{tr}(P_0)^2\leqslant\frac{1}{2}\;\frac{\lambda_S}{\lambda_R}\left[1+\frac{2}{\lambda_R\lambda_S}\right]\left(\frac{e}{\delta}\;\epsilon\;\lambda_{R,S}-1\right)\quad\text{for any }\delta\leqslant e\;\epsilon\;\lambda_{R,S}.$$

The end of the proof of (36) is a direct consequence of [14, Theorem 3.2].

The last assertion resumes to [14, Lemma 4.1]. This ends the proof of the lemma. ∎

6.3 Proof of Lemma 5.2

Using (29) we have

$$\begin{aligned} d\mathrm{tr}(p_t) &= \left(\mathrm{tr}((\partial A\,[m_t]+\partial A\,[m_t]')p_t)-\mathrm{tr}(Sp_t^2)+\mathrm{tr}(R)\right)\;dt+\frac{1}{\sqrt{N-1}}\;d\mathcal{M}_t \\ &\leqslant \left[-\lambda_{\partial A}\;\mathrm{tr}(p_t)-r_1^{-1}\rho(S)\;\mathrm{tr}(p_t)^2+\mathrm{tr}(R)\right]\;dt+\frac{1}{\sqrt{N-1}}\;d\mathcal{M}_t \end{aligned}$$

with a martingale $\mathcal{M}_t$ with an angle bracket

$$\partial_t \langle \mathcal{M} \rangle_t = 4\mathrm{tr}((R + p_t S p_t) p_t) \leqslant 4\mathrm{tr}(p_t) \ \left(\rho(R) + \rho(S) \ \mathrm{tr}(p_t)^2\right).$$

Using [13, Lemma 4.1] we have

$$1 \leqslant n \leqslant 1 + \frac{(N-1)}{2r_1} \frac{\rho(S)}{\lambda_{max}(S)} \Longrightarrow \sup_{t \geqslant 0} \mathbb{E}\left(\mathrm{tr}(p_t)^n\right) < \infty.$$

By (27) we have

$$dm_t = \left[A\left[m_t\right] - p_t S m_t + p_t S X_t\right] dt + p_t \ B' R_2^{-1/2} \ dV_t + \frac{1}{\sqrt{N}} \ d\overline{M}_t.$$

Since $\overline{M}_t$ is independent of V_t we have

$$d\|m_t\|^2 = \left(2 \left\langle m_t, \left[A\left[m_t\right] - p_t S m_t + p_t S X_t\right]\right\rangle + \mathrm{tr}(R + p_t S p_t)\right) \ dt + d\widetilde{M}_t$$

with the martingale

$$d\widetilde{M}_t = 2 \left\langle m_t, p_t \ B' R_2^{-1/2} \ dV_t \right\rangle + 2 \ \frac{1}{\sqrt{N}} \left\langle m_t, \ d\overline{M}_t \right\rangle$$

and the angle bracket

$$\partial_t \langle \widetilde{M} \rangle_t \ = \ 4 \left\langle m_t, (R + p_t S p_t) m_t \right\rangle / N + 4 \left\langle m_t, (p_t S p_t) m_t \right\rangle \leqslant \mathcal{V}_t \ \|m_t\|^2$$

with

$$\mathcal{V}_t \ := 4 \left[\mathrm{tr}(R + p_t S p_t)/N + \mathrm{tr}(p_t S p_t)\right].$$

Observe that

$$\begin{aligned}
\left\langle m_t, A\left[m_t\right]\right\rangle &= \left\langle m_t - 0, A\left[m_t\right] - A(0)\right\rangle + \left\langle m_t, A\left[0\right]\right\rangle \\
&\leqslant -\lambda_A \ \|m_t\|^2 + \|A(0)\| \ \|m_t\| \leqslant -(\lambda_A/2) \ \|m_t\|^2 + \|A(0)\|^2/(2\lambda_A).
\end{aligned}$$

This yields the estimate

$$d\|m_t\|^2 \ \leqslant \ \left(-\lambda_A \ \|m_t\|^2 + \|A(0)\|^2/\lambda_A + 2\|m_t\| \ \|p_t S\| \ \|X_t\| + \mathrm{tr}(R + p_t S p_t)\right) \ dt + d\widetilde{M}_t$$

from which we find that

$$d\|m_t\|^2 \leqslant \left(-\frac{\lambda_A}{2} \ \|m_t\|^2 + \mathcal{U}_t\right) \ dt + \sqrt{\mathcal{V}_t} \ d\mathcal{N}_t$$

with $\partial_t \langle \mathcal{N} \rangle_t \leqslant 1$ and

$$\mathcal{U}_t \ := \ \|A(0)\|^2/\lambda_A + \|p_t S\|^2 \ \|X_t\|^2/\lambda_A + \mathrm{tr}(R + p_t S p_t).$$

Arguing as in the proof of Theorem 4.2 we conclude that

$$\forall 1 \leqslant 3n \leqslant 1 + (N-1)/(2r_1) \qquad \sup_{t \geqslant 0} \mathbb{E}\left(\|m_t\|^{2n}\right) < \infty.$$

Using (4) we have

$$d\xi_t^1 = \left[(\partial A\left[m_t\right] - p_t S)\ \xi_t^1 + p_t S X_t + A\left[m_t\right] - \partial A\left[m_t\right]\ m_t\right]\ dt + d\mathcal{M}_t$$

with the martingale

$$d\mathcal{M}_t := R_1^{1/2} d\overline{W}_t^1 + p_t B' R_2^{-1/2} d(V_t - \overline{V}_t^1).$$

This implies that

$$\begin{aligned} d\|\xi_t^1\|^2 &= \Big[2\langle \xi_t^1, \left[(\partial A\left[m_t\right] - p_t S)\ \xi_t^1 + p_t S X_t + A\left[m_t\right] - \partial A\left[m_t\right]\ m_t\right]\rangle \\ &\qquad\qquad + \mathrm{tr}(R) + 2\mathrm{tr}(p_t S p_t)\Big]\ dt + d\overline{\mathcal{M}}_t \\ &\leqslant \left[-(\lambda_{\partial A}/2)\ \|\xi_t^1\|^2 + \mathcal{U}_t\right]\ dt + d\overline{\mathcal{M}}_t \end{aligned}$$

with

$$d\overline{\mathcal{M}}_t := 2\langle \xi_t^1, d\mathcal{M}_t\rangle \Longrightarrow \partial_t \langle \overline{\mathcal{M}}\rangle_t \leqslant \mathcal{V}_t\ \|\xi_t^1\|^2$$

and

$$\begin{aligned} \mathcal{U}_t &:= 2\|p_t S X_t + A\left[m_t\right] - \partial A\left[m_t\right]\ m_t\|^2/\lambda_{\partial A} + \mathrm{tr}(R) + 2\mathrm{tr}(S p_t^2), \\ \mathcal{V}_t &= 4\left(\mathrm{tr}(R) + 2\mathrm{tr}(p_t S p_t)\right). \end{aligned}$$

The end of the proof follows the same arguments as above, so it is skipped. This completes the proof of the lemma. ■

6.4 Proof of Lemma 5.3

By (27) and (29) we have

$$d(p_t - P_t) = \Pi_t\ dt + d\mathcal{M}_t \quad \text{and} \quad d(m_t - \widehat{X}_t) = \overline{\Pi}_t\ dt + d\overline{\mathcal{M}}_t$$

with the drift terms

$$\begin{aligned} \Pi_t &= \left(\partial A(m_t) p_t - \partial A(\widehat{X}_t) P_t\right) + \left(\partial A(m_t) p_t - \partial A(\widehat{X}_t) P_t\right)' \\ &\qquad + (P_t - p_t) S P_t + ((P_t - p_t) S P_t)' - (P_t - p_t) S (P_t - p_t) \\ \overline{\Pi}_t &= (A(m_t) - A(\widehat{X}_t)) - p_t S (m_t - \widehat{X}_t) + (p_t - P_t) S (X_t - \widehat{X}_t) \end{aligned}$$

and the martingales

$$d\mathcal{M}_t := \frac{1}{\sqrt{N-1}}\ dM_t \qquad d\overline{\mathcal{M}}_t := (p_t - P_t)\ B' R_2^{-1/2}\ dV_t + \frac{1}{\sqrt{N}}\ d\overline{M}_t.$$

Using the decomposition

$$\partial A(m_t) p_t - \partial A(\widehat{X}_t) P_t = \partial A(m_t)(p_t - P_t) + (\partial A(m_t) - \partial A(\widehat{X}_t)) P_t,$$

we check that

$$\Pi_t = \left[\partial A(m_t) - \tfrac{1}{2}(p_t + P_t)S\right](p_t - P_t) + (p_t - P_t)\left[\partial A(m_t) - \tfrac{1}{2}(p_t + P_t)S\right]'$$
$$+(\partial A(m_t) - \partial A(\widehat{X}_t))P_t + P_t(\partial A(m_t) - \partial A(\widehat{X}_t))'.$$

This implies that

$$\langle p_t - P_t, \Pi_t\rangle \leqslant -\lambda_{\partial A} \|p_t - P_t\|_F^2 + 2\kappa_{\partial A} \operatorname{tr}(P_t) \|p_t - P_t\|_F \|m_t - \widehat{X}_t\|$$

from which we prove that

$$\begin{aligned} d\|p_t - P_t\|_F^2 &= 2\langle p_t - P_t, d(p_t - P_t)\rangle \\ &\quad + \frac{2}{N-1}\left[\operatorname{tr}((R + p_t S p_t)p_t) + \operatorname{tr}(R + p_t S p_t)\operatorname{tr}(p_t)\right] dt \\ &\leqslant \Big\{-2\lambda_{\partial A} \|p_t - P_t\|_F^2 + 4\kappa_{\partial A} \operatorname{tr}(P_t) \|p_t - P_t\|_F \|m_t - \widehat{X}_t\| \\ &\quad + \frac{2}{N-1}\left[\operatorname{tr}((R + p_t S p_t)p_t) + \operatorname{tr}(R + p_t S p_t)\operatorname{tr}(p_t)\right]\Big\} dt + d\mathcal{N}_t \end{aligned}$$

with the martingale

$$d\mathcal{N}_t = \frac{2}{\sqrt{N-1}} \langle p_t - P_t, dM_t\rangle = \operatorname{tr}((p_t - P_t)dM_t).$$

After some computations we find that

$$\partial_t \langle \mathcal{N}\rangle_t \leqslant \frac{4}{N-1} \|p_t - P_t\|_F^2 \operatorname{tr}(p_t(R + p_t S p_t)).$$

In much the same vein we have

$$\begin{aligned} \langle m_t - \widehat{X}_t, \overline{\Pi}_t\rangle &= \langle m_t - \widehat{X}_t, (A(m_t) - A(\widehat{X}_t)) - p_t S(m_t - \widehat{X}_t) + (p_t - P_t)S(X_t - \widehat{X}_t)\rangle \\ &\leqslant -\lambda_A \|m_t - \widehat{X}_t\|^2 + \rho(S) \|X_t - \widehat{X}_t\| \|p_t - P_t\|_F \|m_t - \widehat{X}_t\|. \end{aligned}$$

This implies that

$$\begin{aligned} d\|m_t - \widehat{X}_t\|^2 &= 2\langle (m_t - \widehat{X}_t), d(m_t - \widehat{X}_t)\rangle \\ &\quad + \left(\operatorname{tr}(S(p_t - P_t)^2) + \frac{1}{N} \operatorname{tr}(R + p_t S p_t)\right) dt \\ &\leqslant \Big\{-2\lambda_A \|m_t - \widehat{X}_t\|^2 + 2\rho(S) \|X_t - \widehat{X}_t\| \|p_t - P_t\|_F \|m_t - \widehat{X}_t\| \\ &\quad + \rho(S)\|p_t - P_t\|_F^2 + \frac{1}{N} \operatorname{tr}(R + p_t S p_t)\Big\} dt + d\overline{\mathcal{N}}_t \end{aligned}$$

with the martingale

$$d\overline{\mathcal{N}}_t = 2\langle (m_t - \widehat{X}_t), d\overline{\mathcal{M}}_t\rangle = 2\langle (m_t - \widehat{X}_t), (p_t - P_t) B' R_2^{-1/2} dV_t\rangle + \frac{2}{\sqrt{N}} \langle (m_t - \widehat{X}_t), d\overline{M}_t\rangle.$$

In addition we have

$$\begin{aligned}\partial_t \langle \overline{\mathcal{N}} \rangle_t &\leqslant 4\rho(S)\ \|m_t - \widehat{X}_t\|^2 \|p_t - P_t\|_F^2 + \frac{4}{N}\ \langle (m_t - \widehat{X}_t), (R + p_t S p_t)(m_t - \widehat{X}_t) \rangle \\ &\leqslant 2\rho(S)\ \left(\|m_t - \widehat{X}_t\|^2 + \|p_t - P_t\|_F^2 \right)^2 + \frac{4}{N}\ \|m_t - \widehat{X}_t\|^2\ \mathrm{tr}(R + p_t S p_t).\end{aligned}$$

Combining the above estimates we find that

$$\begin{aligned}d\,\Xi_t \leqslant &\left\{ -2\lambda_A\ \|m_t - \widehat{X}_t\|^2 + 2\ \|p_t - P_t\|_F\ \|m_t - \widehat{X}_t\|\ \left(2\kappa_{\partial A}\ \mathrm{tr}(P_t) + \rho(S)\ \|X_t - \widehat{X}_t\| \right) \right\}\ dt \\ &\qquad - (2\lambda_{\partial A} - \rho(S))\, \|p_t - P_t\|_F^2\ dt \\ &+ \frac{1}{N} \left\{ \mathrm{tr}(R + p_t S p_t) + \frac{2N}{N-1}\ \left[\mathrm{tr}((R + p_t S p_t) p_t) + \mathrm{tr}(R + p_t S p_t) \mathrm{tr}(p_t) \right]\ \right\}\ dt + d\mathcal{N}_t + d\overline{\mathcal{N}}_t.\end{aligned}$$

Recalling that

$$2\lambda_A \geqslant \lambda_{\partial A} > 0 \quad \text{and} \quad 2\lambda_{\partial A} - \rho(S) \geqslant \lambda_{\partial A},$$

this yields the estimate

$$\begin{aligned}d\,\Xi_t &\leqslant \left\{ -\lambda_{\partial A}\ \Xi_t + 2\ \|p_t - P_t\|_F\ \|m_t - \widehat{X}_t\| \left(2\kappa_{\partial A}\ \mathrm{tr}(P_t) + \rho(S)\ \|X_t - \widehat{X}_t\|\ \right) \right\}\ dt \\ &\quad + \frac{1}{N} \left(\frac{4N}{N-1}\ \mathrm{tr}(R + p_t S p_t)\ \mathrm{tr}(p_t) + \ \mathrm{tr}(R + p_t S p_t) \right)\ dt + d\overline{\mathcal{N}}_t + d\mathcal{N}_t.\end{aligned}$$

On the other hand using the inequality $2ab \leqslant a^2 + b^2$ we prove that

$$\begin{aligned}d\,\Xi_t &\leqslant -\left\{ \lambda_{\partial A} - \left(2\kappa_{\partial A}\ \mathrm{tr}(P_t) + \rho(S)\ \|X_t - \widehat{X}_t\|\ \right) \right\}\ \Xi_t\ dt \\ &\quad + \frac{1}{N}\ \left(1 + \frac{4}{1 - 1/N}\ \mathrm{tr}(p_t) \right)\ \left[\mathrm{tr}(R) + \mathrm{tr}(S) \mathrm{tr}(p_t)^2 \right]\ dt + d\overline{\mathcal{N}}_t + d\mathcal{N}_t\end{aligned}$$

from which we conclude that

$$d\,\Xi_t \leqslant \left(\Gamma_A(t)\ \Xi_t\ + \frac{1}{N}\ \mathcal{U}_t \right) dt + d\,\Upsilon_t \quad \text{with} \quad \mathcal{U}_t := (1 + 8\mathrm{tr}(p_t))\ \left[\mathrm{tr}(R) + \rho(S)\ \mathrm{tr}(p_t)^2 \right]$$

and the martingale $\Upsilon_t := \Upsilon_t^{(1)} + \Upsilon_t^{(2)}$ given by

$$\begin{aligned}d\,\Upsilon_t^{(1)} &:= 2\ \langle (m_t - \widehat{X}_t), (p_t - P_t)\ B' R_2^{-1/2}\ dV_t \rangle \\ d\,\Upsilon_t^{(2)} &:= \frac{2}{\sqrt{N}}\ \langle (m_t - \widehat{X}_t), d\overline{M}_t \rangle + \frac{2}{\sqrt{N-1}}\ \langle p_t - P_t, dM_t \rangle.\end{aligned}$$

Observe that

$$\begin{aligned}\langle \Upsilon^{(1)}, \Upsilon^{(2)} \rangle_t &= 0 \\ \partial_t \langle \Upsilon^{(1)} \rangle_t &\leqslant 2\rho(S)\ \left(\|m_t - \widehat{X}_t\|^2 + \|p_t - P_t\|_F^2 \right)^2 \leqslant 2\rho(S)\ \Xi_t^2 \\ \partial_t \langle \Upsilon^{(2)} \rangle_t &\leqslant \frac{4}{N} \left[\ \|m_t - \widehat{X}_t\|^2\ + 2\ \|p_t - P_t\|_F^2\ \mathrm{tr}(p_t) \right]\ \mathrm{tr}(R + p_t S p_t) \leqslant \frac{4}{N}\ \mathcal{U}_t\ \Xi_t.\end{aligned}$$

This ends the proof of the lemma. ■

Acknowledgements

We would like to thank an anonymous referee for the solving of a technical issue in the earlier proof of Theorem 3.2. Her/His detailed comments also greatly improved the presentation of the article.

References

[1] J.L. Anderson. An ensemble adjustment Kalman filter for data assimilation. Mon. Weather Rev., vol. 129, no. 12, pp. 2884–2903 (2001).

[2] J.L. Anderson. An adaptive covariance inflation error correction algorithm for ensemble filters. Tellus A vol. 59, no. 2, , pp. 210–224 (2007).

[3] J L Anderson. Spatially and temporally varing adaptive covariance inflation for ensemble filters. Tellus A, vol. 61, pp. 72–83 (2009).

[4] T. Berry, J. Harlim. Linear theory for filtering nonlinear multiscale systems with model error. Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 470, no 2167 (2014).

[5] G. Burgers, P. Van Leeuwen, G. Evensen. On the analysis scheme in the ensemble Kalman filter. Monthly Weather Review 126, pp. 1719–1724 (1998).

[6] R. Carmona, J.-P. Fouque, and L.-H. Sun. Mean Field Games and Systemic Risk, Communications in Mathematical Sciences. Vol. 13, no. 4, pp. 911–933 (2015).

[7] A Cuzol, P Hellier, É Memin. A low dimensional fluid motion estimator. International Journal of Computer Vision 75, no. 3, pp. 329–349 (2007).

[8] A Cuzol, É Memin. Vortex and source particles for fluid motion estimation. 5th Int. Conf. on Scale-Space and PDE methods in Computer Vision, Scale-Space'05, Hofgeismar, Germany, Apr. (2005).

[9] A Cuzol, É Memin. A stochastic filter for fluid motion tracking. 10th IEEE International Conference on Computer Vision, ICCV'05, Beijing, China, Oct. (2005).

[10] A Cuzol, É Memin. A stochastic filtering technique for fluid flows velocity fields tracking. IEEE PAMI (Pattern Analysis and Machine Intelligence), 31(7):1278-1293 (2009).

[11] P. Del Moral. Feynman-Kac formula. Genealogical and interacting particle approximations. Springer New York. Series: Probability and Applications (2004).

[12] P. Del Moral. Mean field simulation for Monte Carlo integration. Chapman & Hall/CRC Monographs on Statistics & Applied Probability [626p.] (2013)

[13] P. Del Moral, J. Tugaut. On the stability and the uniform propagation of chaos properties of Ensemble Kalman-Bucy filters. Arxiv:1605.09329 (2016).

[14] P. Del Moral, A. Kurtzmann, J. Tugaut. On the stability and the exponential concentration of extended Kalman-Bucy filters, ArXiv:1606.08251 (2016).

[15] G. Evensen. Sequential data assimilation with a non-linear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J Geophys Res 99(C5): vol.10 pp. 143–162 (1994)

[16] G. Evensen. The Ensemble Kalman Filter: theoretical formulation and practical implementation. Ocean Dynamics vol. 53, pp. 343–367 (2003).

[17] G. Evensen. Data assimilation : The ensemble Kalman filter, Springer, Berlin (2007).

[18] G. Evensen, J. Hove, H.C. Meisingset, E. Reiso, K.S. Seim. Using the EnKF for assisted history matching of a North Sea Reservoir Model SPE 106184 (2007).

[19] J.-P. Fouque and L.-H. Sun. Systemic Risk Illustrated, Handbook on Systemic Risk, Eds J.-P. Fouque and J. Langsam. Cambridge University Press (2013).

[20] G. Gottwald and A. J. Majda. A mechanism for catastrophic filter divergence in data assimilation for sparse observation networks. Nonlin. Processes Geophys, vol. 20, pp. 705–712 (2013).

[21] D.T. B. Kelly, K.J. Law, A. M. Stuart. Well-posedness and accuracy of the Ensemble Kalman filter in discrete and continuous time. Nonlinearity, vol. 27, pp. 2579–2603 (2014).

[22] D. Kelly, A. J. Majda, and X. T. Tong. Concrete ensemble kalman filters with rigorous catastrophic filter divergence. Proc. Natl. Acad. Sci. (2015).

[23] F. Le Gland, V. Monbet, V.D. Tran. Large sample asymptotics for the ensemble Kalman filter The Oxford Handbook of Nonlinear Filtering, chapter 22, pp. 598–631 (2011).

[24] J. Harlim, B. Hunt. Local Ensemble Transform Kalman Filter: An Efficient Scheme for Assimilating Atmospheric Data. *Preprint* (2005).

[25] E. Kalnay. Atmospheric modeling, data assimilation, and predictability. Cambridge university press (2003).

[26] K. A. Lisaeter, J. Rosanova, G. Evensen. Assimilation of ice concentration in a coupled ice-ocean model, using the Ensemble Kalman Filter. Ocean Dynamics, vol. 53, pp. 368–388 (2003).

[27] A. J. Majda and J. Harlim. Filtering complex turbulent systems. Cambridge University Press, Cambridge UK (2012).

[28] A. J. Majda and J. Harlim. Catastrophic filter divergence in filtering nonlinear dissipative systems. Comm. Math. Sci., vol. 8, pp. 27–43 (2008).

[29] J. Mandel, L. Cobb, J. D. Beezley. On the convergence of the ensemble Kalman filter Applications of Mathematics, vol. 56, no. 6, pp. 533–541 (2011).

[30] P. A. Markowich, C. Villani. On the trend to equilibrium for the Fokker-Planck equation: An interplay between physics and functional analysis. VI Workshop on Partial Differential Equations, Part II, Mat. Contemp., Rio de Janeiro, vol. 19, pp. 1–29 (2000).

[31] G. Naevdal, L. M. Johnsen, S. I. Aanonsen, E.H. Vefring. Reservoir monitoring and continuous model updating using ensemble Kalman filter. SPE journal, vol. 10, no. 1, pp. 66–74 (2005).

[32] E. Ott, B. R. Hunt, I. Szunyogh, A. V. Zimin, E. J. Kostelich, M. Corazza, E. Kalnay, D. Patil, and J. A. Yorke. A local ensemble Kalman filter for atmospheric data assimilation. Tellus A, vol. 56, pp. 415–428 (2004).

[33] N. Papadakis, É Mémin, A. Cuzol, N. Gengembre. Data assimilation with the weighted ensemble Kalman filter. Tellus A 62, no. 5, pp. 673–697 (2010).

[34] A. Seiler, G. Evensen, J.-A. Skjervheim, J. Hove, and J. G. Vab. Using the EnKF for history matching and uncertainty quantification of complex reservoir models. in Computational Methods for Large Scale Inverse Problems and Quantification of Uncertainty, Edited by L. Biegler, G. Biros, O. Ghattas, M. Heinkenschloss, D. Keyes, B. Mallick, L. Tenorio, B. van Bloemen Waanders, K Willcox John Wiley & Sons, Ltd. (2010).

[35] J-A. Skjervheim, G. Evensen, S.I. Aanonsen, B.O. Ruud and T.A. Johansen. Incorporating 4D seismic data in reservoir simulation models using ensemble Kalman filter. SPE 95789 (2005).

[36] X. T. Tong, A. J. Majda and D. Kelly. Nonlinear stability and ergodicity of ensemble based Kalman filters. Nonlinearity, vol. 29, no. 2, pp. 657–691 (2016).

[37] X.-H. Wen, W.H. Chen. Real-Time Reservoir Model Updating Using Ensemble Kalman Filter SPE-92991-MS, SPE Reservoir Simulation Symposium, The Woodlands, Texas (2005).

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