Data assimilation for stochastic ocean models

Benjamin Dufée

COLLEGE MATHS, TELECOMS
DOCTORAL INFORMATIQUE, SIGNAL
BRETAGNE SYSTEMES, ELECTRONIQUE

Université de Rennes

THÈSE DE DOCTORAT DE

L'UNIVERSITÉ DE RENNES

ÉCOLE DOCTORALE N° 601
Mathématiques, télécommunications, informatique, signal, systèmes, électronique
Spécialité : *Mathématiques et leurs interactions*

Par
Benjamin DUFÉE

Data assimilation for stochastic ocean models

Thèse présentée et soutenue à Rennes, le 29 septembre 2023
Unité de recherche : INRIA / IRMAR
Thèse N° : à compléter

Rapporteurs avant soutenance :

Éric BLAYO	Professeur, Université Grenoble Alpes
Pierre DEL MORAL	Directeur de recherche, INRIA Bordeaux

Composition du Jury :

Président/e :		
Examinateurs :	Valérie MONBET	Professeur, Université de Rennes
	Éric BLAYO	Professeur, Université Grenoble Alpes
	Pierre DEL MORAL	Directeur de recherche, INRIA Bordeaux
	Serge GRATTON	Professeur, INP Toulouse
	Dan CRISAN	Professor, Imperial College London
Dir. de thèse :	Étienne MÉMIN	Directeur de recherche, INRIA Rennes

Remerciements

Je vais tout d'abord remercier mon directeur de thèse, Étienne Mémin, pour tout le temps qu'il a pu me dédier, pour ses conseils, ses orientations et sa bonne humeur. Je remercie également mes deux autres collaborateurs principaux : Dan Crisan et Gilles Tissot.

Merci aux rapporteurs et au jury de ma thèse pour le temps consacré à l'étude du manuscrit et à leurs remarques.

Je remercie mes collègues des équipes Odyssey/Fluminance et S(AI/E)erpico pour m'avoir accompagné et/ou subi pendant ces 3 ans, et tout particulièrement mon co-bureau Francesco pour les bons moments passés ensemble. Merci aussi à mes collègues d'enseignement, Anne, Marie-Pierre, Ronan et Grégoire, pour la bonne ambiance de travail.

Je vais maintenant remercier les amis et camarades qui m'ont accompagné pendant mon parcours rennais qui a démarré en 2016 avec mon M1 et qui aboutit à cette soutenance d'aujourd'hui : Agathe, Alice, Antoine, Élisa, Enguerrand, François, Grégoire (encore lui !), Gwendal, Laure, Pierre, Sébastien, Silvère, Steven et Typhène.

Un grand merci à ma famille, et tout particulièrement à mes formidables parents, pour leur soutien inconditionnel. Il n'y a rien de plus précieux...

Si ce n'est peut-être le financement 856408-STUOD du Conseil Européen de la Recherche, qui a subventionné ma thèse et au titre duquel je me permets de remercier à nouveau l'ensemble des personnes sus-mentionnées (à l'exception des exilés fiscaux qui se reconnaîtront) pour leurs contributions diverses au financement public.

Pour finir, je salue tout particulièrement mes compagnons de route les plus chers et qui m'accompagnent au quotidien : Alice, Antoine et Steven. Merci d'être là.

Table of Contents

List of Figures

Résumé en français

L'assimilation de données est le processus d'incorporation de données d'observations partielles d'un système dans un modèle d'évolution numérique de ce même système afin d'enrichir sa dynamique et de faire de l'estimation ou de la prédiction d'états du système. L'assimilation est utilisée depuis bien longtemps et demeure aujourd'hui un domaine de recherche très actif pour la prévision météorologique et les systèmes géophysiques. Une large panoplie de méthodes a été conçue en ce sens, allant du contrôle optimal au filtrage stochastique, parmi tant d'autres méthodes dont on peut trouver une description dans le panorama récent proposé par Carrassi et al. (2018). Parmi elles, trois classes de méthodes ont été particulièrement étudiées et/ou efficaces d'un point de vue opérationnel : les méthodes dites variationnelles, les filtres de Kalman d'ensemble et les filtres particulaires.

Les méthodes variationnelles sont à rapprocher de la théorie du contrôle optimal (Le Dimet and Talagrand, 1986; Simon et al., 2011; Blayo et al., 2014). Étant donné un estimé a priori (dit de background), la méthode consiste à minimiser une fonction de coût qui prend en compte les écarts au background et aux observations. Dans certains cas, la fonction est quadratique et donc facilement minimisable. Mais la plupart du temps, ces méthodes demandent en théorie un calcul explicite du modèle linéaire tangent, ce qui peut être très coûteux et bien sûr intrinsèquement lié à l'état courant du code considéré, même si de nos jours les procédures d'auto-différentiation épargnent ou rendent implicite ce calcul. D'autres techniques ont également été appliquées aux méthodes variationnelles, telles que le préconditionnement (Gürol et al., 2014) et les méthodes de Gauss-Newton (Gratton et al., 2007).

Les méthodes d'ensemble se basent sur des ensembles de réalisations d'un système dynamique donné. Parmi elles, les filtres de Kalman d'ensemble (EnKF en anglais) calculent l'ensemble filtré (ou postérieur) comme combinaisons linéaires des membres de l'ensemble non filtré (ou antérieur). On renvoie à Evensen (1994); Asch et al. (2016); Evensen (2006); Reich and Cotter (2015) pour des descriptions précises des différentes variantes de l'EnKF. Bien que les équations de Kalman aient été originellement établies sous des hypothèses de linéarité du modèle dynamique et de Gaussianité du bruit de modèle, l'application ensembliste de ces méthodes s'est montrée efficace pour des modèles non linéaires, même

en très grande dimension (Houtekamer and Mitchell, 1998, 2001, 2005; Houtekamer et al., 2005), pour lesquels le principe de superposition induit par les EnKF est a priori sujet à caution. À noter que ces méthodes épargnent tout calcul explicite du modèle linéaire tangent. Le problème de stockage des matrices de covariance en très grande dimension a mené à des reformulations des équations de Kalman en termes de matrices d'anomalie, aboutissant à ce que l'on appelle des filtres d'ensemble "square-root" (ESRF)(Anderson, 2001; Bishop et al., 2001; Pham, 2001; Tippett et al., 2003; Whitaker and Hamill, 2002a). La principale limitation de ces méthodes est la taille de l'ensemble, qui est généralement de taille très limitée (entre 10 et 200). En effet, ces approches se basent sur une approximation de rang faible de la "vraie" matrice de covariance, de taille $D_X \times D_X$, en notant D_X la dimension de l'espace d'états, par la matrice de covariance empirique, qui est au mieux de rang $N-1$, où N est la taille de l'ensemble. Les techniques dites de localisation ont été conçues pour contourner ce problème via une réduction artificielle de la dimension de l'espace d'états par des voisinages locaux (Anderson, 2007, 2012; Hamill et al., 2001; Sakov and Bertino, 2011), conduisant de facto à une augmentation du rang de la matrice de covariance empirique dégénérée. La localisation est implémentée en multipliant par un produit de Schur soit la matrice de covariance empirique, soit la matrice de covariance d'erreur d'observation par une matrice issue d'une fonction à support compact visant à tuer les corrélations résiduelles entre des points de grille éloignés, définissant ainsi un voisinage local d'intérêt. Ces méthodes sont très efficaces d'un point de vue calculatoire. Néanmoins, les membres d'ensemble postérieurs sont par construction des combinaisons par blocs de filtrages locaux, et peuvent donc ne plus être solutions du système dynamique considéré.

Les filtres particulaires constituent une autre classe de méthodes d'ensemble, bien qu'elles soient davantage à rapprocher des méthodes de Monte-Carlo. Contrairement aux filtres de Kalman d'ensemble, les membres d'ensemble postérieurs ne sont pas calculés comme combinaisons linéaires des membres antérieurs. Ils sont des copies bien choisies des membres d'ensemble (aussi appelés particules dans ce contexte) qui sont les plus probables relativement aux observations. Ainsi les combinaisons linéaires ne sont pas réalisées sur les membres d'ensemble, c'est-à-dire les états X, mais sur les atomes de la distribution de filtrage δ_X. Les filtres particulaires ne nécessitent aucune hypothèse particulière sur la nature du modèle et du bruit, ce qui présente en soi un avantage théorique incontestable. Néanmoins, les filtres particulaires classiques comme le rééchantillonnage par importance (SIR en anglais) (Rubin, 1987) ne fonctionnent pas pour des systèmes de grande dimen-

sion. En effet, comme la plupart des méthodes de Monte-Carlo, le nombre de particules doit croître exponentiellement avec la dimension de l'espace d'états pour que la méthode soit efficace (Snyder et al., 2008). Des avancées ont été faites avec des techniques additionnelles apportées au filtre particulaire classique (Cotter et al., 2020a,b; Beskos et al., 2017; Kantas et al., 2014) pour les systèmes de grande dimension.

Comme mentionné ci-dessus, ces méthodes d'ensemble se basent sur un principe de superposition, qui est complètement justifié pour les filtres particulaires (combinaisons linéaires d'atomes de distribution) mais pas pour les EnKF (combinaisons linéaires d'états). Afin de le justifier complètement, une approche consiste à utiliser la théorie des espaces de Hilbert à noyau auto-reproduisant (RKHS en anglais) liés à des caractéristiques intrinsèques du système dynamique. L'opérateur de Koopman est en ce sens un candidat d'étude idéal. La représentation spectrale de l'opérateur de Koopman permet l'extraction de fonctions propres intrinsèques au système dynamique (c'est-à-dire indépendantes de la condition initiale) et constitue en soi un outil fondamental à buts théorique (théorie ergodique (Eisner et al., 2015)) et numérique comme par exemple pour de la modélisation guidée par les données. La plupart de ces approches numériques se basent soit sur des moyennes en temps long (Mezic, 2005; Budisic, M. and Mohr, R. and Mezic, I. , 2012; Das and Giannakis, 2019; Das and D.Giannakis, 2020) avec une hypothèse d'ergodicité de la dynamique, ou sur des approximations de dimension finie de l'opérateur de Koopman comme la décomposition en modes dynamiques et ces extensions (Rowley et al., 2009; Schmid, 2010) ou sur des approximations de Galerkin et des techniques dites de "delay embedding" (Brunton et al., 2017; Giannakis et al., 2015; Giannakis, 2017). Ces méthodes permettent une projection spectrale du système issue de longues séries temporelles de données observées, mais sont par essence des approximations de dimension finie d'un opérateur de dimension infinie qui contient pourtant un spectre avec potentiellement une composante continue (Mezic, 2005), ou avec une famille infinie dénombrable de modes. Travailler dans des espaces de fonctions lisses comme les RKHS s'est montré efficace pour traiter ces problèmes (Das and Giannakis, 2019; Das and D.Giannakis, 2020; Das et al., 2021). Dans un contexte d'assimilation de données, l'idée est de remplacer de longues séries temporelles de données par un ensemble de réalisations visant à explorer, localement en temps, l'espace des phases via les RKHS. Cela permettra de générer des nouvelles trajectoires solutions par un principe de superposition pleinement justifié, légitimant l'usage de filtres de Kalman dans cette configuration.

L'assimilation de données, et en particulier le filtrage stochastique (avec l'EnKF comme

représentant du cas linéaire gaussien), pose naturellement la question de l'aléatoire en jeu. Dans la plupart des cas, l'aléatoire considéré provient de la condition initiale de réalisations déterministes, et dans certains cas une variable aléatoire additive représentant l'erreur de modèle. Un problème classique avec ce genre d'aléas est le manque de variabilité de l'ensemble, comme il n'est pas (ou que très marginalement) généré par l'aspect chaotique de la dynamique. L'inflation de la variance a été introduite pour enrichir la variabilité de l'ensemble. L'inflation additive consiste en l'ajout de perturbations globales aléatoires aux membres d'ensemble et peut être utilisée aussi bien pour les filtres particulaires que pour les EnKF. Néanmoins, l'inflation additive constitue un ajout de variance aveugle sans structure spatiale bien précise, et qui n'est par conséquent que peu utilisée en pratique. L'inflation multiplicative est bien plus répandue et consiste en la multiplication de l'anomalie de chaque membre d'ensemble (c'est-à-dire son écart à la moyenne empirique) par un scalaire. Cette inflation est plus subtile car elle est sensée accroître la variabilité de l'ensemble dans des directions intéressantes, avec potentiellement une structure spatiale appropriée.

Une grande partie de cette thèse étudie une manière plus sophistiquée d'introduire l'aléatoire pour les modèles géophysiques. L'idée est de faire de l'aléatoire une part entière du modèle, et pas simplement par le biais d'un forçage ad hoc ou de l'inflation. Cette méthodologie, appelée incertitude de position (LU en anglais), a été initialement formalisée par Mémin (2014), et s'appuie sur une décomposition de la vitesse Lagrangienne en une composante grande-échelle, lisse en temps, et d'une composante aléatoire fortement oscillante et décorrélée en temps. Toute la dynamique est ensuite déduite de ce point de départ afin de représenter complètement l'impact des comportements à petite échelle sur le système global. De ce point de vue, le but de LU de représenter des comportements physiques cachés de la grande échelle peuvent être vus comme faisant partie du domaine de la physique stochastique pour les sciences du climat (Berner et al., 2017; Franzke et al., 2014; Gottwald et al., 2016; Palmer, 2019). Le principal intérêt de cette méthode, comparativement aux méthodes déterministes, est que la dérivation du modèle LU intègre directement le caractère aléatoire de la dynamique dans un opérateur de transport stochastique modifié, qui remplace la dérivée matérielle habituelle. Par conséquent, la méthode est complètement portée par la dynamique et dérive des lois de conservation physiques. Des précédentes études sur la quantification d'incertitude (Resseguier et al., 2017c, 2020a; Bauer et al., 2020a,b) ont montré que LU, pour les modèles considérés, fournissait un meilleur compromis biais-variance que la perturbation des conditions initiales

pour les simulations déterministes. Le cadre LU a été également utilisé pour des systèmes géophysiques grande échelle (Resseguier et al., 2017a,b,c; Bauer et al., 2020a,b; Brecht et al., 2021) et pour la simulation aux grandes échelles (LES - Large Eddies Simulation) (Kadri Harouna and Mémin, 2017; Chandramouli et al., 2020).

Après cette introduction générale, présentons maintenant un résumé succinct du plan de la thèse.

Le chapitre 1 constitue un panorama rapide des techniques d'assimilation d'intérêt dans nos différentes études. Le chapitre 2 présente la méthode stochastique LU et sa formulation pour le modèle SQG. Le modèle SQG est un modèle océanique 2D relativement simple qui recèle néanmoins des caractéristiques hautement non-linéaires (Blumen, 1978; Constantin et al., 1994, 1999, 2012; Held et al., 1995; Lapeyre and Klein, 2006). Il constitue donc un modèle jouet intéressant et non-trivial pour nos expériences. Les procédures de génération du bruit sont également détaillées dans ce chapitre. Le chapitre 3 regroupe les résultats qui ont été publiés pendant ce doctorat. La section 3.2 étudie la pertinence de la modélisation stochastique par LU pour remplacer l'inflation déterministe pour les filtres de Kalman d'ensemble dans le modèle SQG (Dufée et al., 2022). La section 3.3 décrit une méthodologie basée sur LU et les transformations de Girsanov qui permet de guider l'essaim de trajectoires vers une région d'intérêt, proche des observations. Les résultats expérimentaux montrent des améliorations significatives apportées par cette procédure (Dufée et al., 2023a). Le chapitre 4 explore le formalisme des RKHS pour l'opérateur de Koopman dans le cadre de la prédiction d'ensemble (Dufée et al., 2023b) et présente des techniques d'assimilation adaptées au formalisme théorique adopté.

Des parties de cette thèse seront largement inspirées des études publiées pendant ce doctorat (Dufée et al., 2022; Dufée et al., 2023a; Dufée et al., 2023b).

INTRODUCTION

Sequential data assimilation is the process that incorporates partial observations of a given physical system into a numerical dynamical model of this system in order to enrich its knowledge and make state estimation or prediction. It has been used for a long time and is still an active area of research for weather forecasting and geophysical systems. A wide variety of methods have been designed for this kind of purpose, from optimal control theory to stochastic filtering, among many other techniques that can be found in Carrassi et al. (2018), which proposes an up-to-date overview of data assimilation techniques. Among all these, three main classes of methods have proven effective in operational centers: variational methods, ensemble Kalman filters and particle filters.

Variational methods are part of optimal control theory (Le Dimet and Talagrand, 1986; Simon et al., 2011; Blayo et al., 2014). Given a background estimate, the method consists in minimizing a cost function that accounts for the departure from both the background and the observation. In some particular cases, the function is quadratic and thus manageable. But most of the time, these methods theoretically require the explicit computation of the tangent linear model, which can be very costly and is naturally code-related, even if nowadays auto-differentiation procedures become very efficient by sparing (or making implicit) this computation. We may also refer to the applications of preconditioning (Gürol et al., 2014) and Gauss-Newton methods (Gratton et al., 2007) to variational methods.

Ensemble methods rely on ensemble of realizations of a given dynamical system. Among them, ensemble Kalman filters (EnKF) compute the filtered (or posterior) ensemble as linear combinations of the non-filtered (or prior) ensemble members. We refer to Evensen (1994); Asch et al. (2016); Evensen (2006); Reich and Cotter (2015) for precise descriptions of the different variants of EnKF. Even if the Kalman equations were initially set in hypotheses of model linearity and noise Gaussianity, these methods were shown to be efficient with non-linear models, even in very high-dimensional systems (Houtekamer and Mitchell, 1998, 2001, 2005; Houtekamer et al., 2005), for which the superposition principle entailed by the EnKF is a priori questionable. Note that these methods completely spare the explicit computation of the tangent linear model. Computational intractability of the covariance matrices in high-dimensional problems led to reformulations of the

Kalman equations in terms of the anomaly matrices, resulting in the so called ensemble square-root filters (ESRF) (Anderson, 2001; Bishop et al., 2001; Pham, 2001; Tippett et al., 2003; Whitaker and Hamill, 2002a). The main limitation of these methods is the ensemble size, which is generally of moderate size (from tens to 100-200 at most), especially when dealing with high-dimensional systems. Indeed, these approaches rely on a low-rank approximation of the actual covariance matrix, which is of size $D_X \times D_X$, denoting D_X the size of the state space, by the empirical covariance matrix, which is at best of rank $N-1$, where N is the ensemble size. In order to address this possible sampling issue, localization techniques were designed to artificially reduce the dimension of the state space in local neighborhoods (Anderson, 2007, 2012; Hamill et al., 2001; Sakov and Bertino, 2011). This localization is implemented by a Schur product between either the empirical covariance matrix or the observation error covariance matrix by a tapering matrix with compact support that kills spurious correlations between distant grid points and defines a local window of relevance. This is computationally very efficient. Nevertheless, the posterior ensemble members are by construction a recombination of local updates, which might not be solution of the underlying system anymore.

Particle filters constitute another class of ensemble methods, but are part of the more general Monte-Carlo methods. Contrary to the Kalman filters, the posterior ensemble members are not linear combinations of the prior ones, but copies of the prior ensemble members (also called particles in this context) that are the most likely with respect to the observations. So the linear combinations are not made on the states X, but instead on the Dirac measures of the distribution δ_X. Particle filtering does not require any hypotheses on the noise or on the model, which is a theoretical great advantage. Nevertheless, standard particle filters like Sampling Importance Resampling (SIR) (Rubin, 1987) fail in high-dimensional systems, as the number of particles has to scale exponentially with the dimension so that the method works (Snyder et al., 2008). Progress was made with additionnal techniques brought to the standard particle filter (Cotter et al., 2020a,b; Beskos et al., 2017; Kantas et al., 2014) for high-dimensional problems.

As previously mentioned, these ensemble-based methodologies rely on a superposition principle, which is fully justified for the particle filter (superposition of Dirac measures) but not for EnKF (linear combinations of states). An approach in order to justify this superposition principle is to work with reproducing kernel Hilbert spaces (RKHS) that are linked with intrinsic features of the dynamical system, so the Koopman operator is a perfect candidate with this respect. The spectral representation of the Koopman operator

enables to extract intrinsic eigenfunctions of the system dynamics and constitutes in itself a powerful tool for both theoretical (ergodic theory (Eisner et al., 2015)) and numerical purposes such as data-driven modeling. Most of them rely either on longtime averages with ergodicity assumptions of the dynamics (Mezic, 2005; Budisic, M. and Mohr, R. and Mezic, I. , 2012; Das and Giannakis, 2019; Das and D.Giannakis, 2020) or on finite-dimensional approximations of the Koopman operator such as the dynamic modes representation and its extensions (Rowley et al., 2009; Schmid, 2010) or on Galerkin approximation and delay embedding (Brunton et al., 2017; Giannakis et al., 2015; Giannakis, 2017). These methods enable performing spectral projection from long time series of measured observables of the system, but are by essence finite dimensional approximations of an infinite dimensional operator that often admits a spectrum that is partly continuous or with a summable infinite number of modes (Mezic, 2005). Working in spaces of smooth functions like RKHS was shown helpful to tackle these issues (Das and Giannakis, 2019; Das and D.Giannakis, 2020; Das et al., 2021). In an ensemble data assimilation context, the idea is to substitute long time series of data by the ensemble of realizations in order to explore, locally in time, the phase space of trajectories using RKHS. This will enable to generate new trajectories through a justified superposition principle, and justifies the use of EnKF in this framework.

Data assimilation methods naturally raise the nature of the randomness at stake. In most cases, the randomness considered is related to the initial condition of deterministic realizations, and in some cases an additive random variable accounting for the model error. A classical issue with such randomness is the lack of ensemble spread, as it is not (or very marginally) generated by the chaotic dynamics. Variance inflation was introduced to try to enrich the ensemble spread. Additive inflation is based on the addition of random global perturbations to the ensemble members and can be used for both EnKF and particle filters. Still, this is a completely blind increase of variance with no particular spatial pattern, which is barely used in practice. Multiplicative inflation is much more popular as it multiplies the ensemble members' anomaly (i.e. the departure from the empirical mean) by a scaling factor. This is more subtle as this is expected to expand the spread in directions that are needed, with potentially an adequate spatial pattern.

A major part of this thesis studies a more sophisticated randomization framework for geophysical problems, that does not add stochasticity as an external forcing or inflation, but is fully part of the model and its derivation. This methodology, called Location Uncertainty (LU), was first formalized in Mémin (2014), and relies on the decomposition of the Lagrangian velocity in terms of a smooth (in time) component and a highly

oscillating random component which is time-uncorrelated. The whole dynamics is then derived from this starting point in order to fully represent the impact of the small-scale behaviors on the global system. With this respect, the aim of LU to represent some sort of hidden physics can be seen as part of the domain of stochastic physics for climate sciences (Berner et al., 2017; Franzke et al., 2014; Gottwald et al., 2016; Palmer, 2019). The main asset of this method compared to deterministic ones is that the derivation of LU implies the enforcement of the randomness in a modified stochastic transport operator that replaces the usual material derivative. Consequently, the framework is completely flow-driven and is based on actual physical conservation laws. Some previous studies on uncertainty quantification (Resseguier et al., 2017c, 2020a; Bauer et al., 2020a,b) pointed out that LU, on the models at stake, was exhibiting a better error-spread trade-off than perturbations of initial conditions for deterministic simulations. The LU framework was used to address large-scale geophysical systems (Resseguier et al., 2017a,b,c; Bauer et al., 2020a,b; Brecht et al., 2021) and for large eddy simulations (Kadri Harouna and Mémin, 2017; Chandramouli et al., 2020).

After this general introduction, let us present the outline of this thesis.

Chapter 1 will constitute a quick overview of the data assimilation techniques of interest in our different studies. Chapter 2 presents the LU stochastic methodology and the formulation of the Surface Quasi-Geostrophic (SQG) model in this framework. The SQG model is a quite simple 2D model but for which ocean highly non-linear features are at play (Blumen, 1978; Constantin et al., 1994, 1999, 2012; Held et al., 1995; Lapeyre and Klein, 2006). So it constitutes a nice and non trivial toy model for our experiments. The noise generation procedures are also detailed in this chapter. Chapter 3 recaps the findings that were published during this PhD. Section 3.2 investigates how LU is a more efficient and secure way to replace inflation for ensemble Kalman filters in the SQG system (Dufée et al., 2022). Section 3.3 describes a methodology that, based on LU and Girsanov transforms, enables to guide the ensemble members towards a desired region close to the observation. Experimental results show significant improvements brought by this additional technique (Dufée et al., 2023a). Chapter 4 investigates the RKHS formalism of the Koopman operator for ensemble forecast (Dufée et al., 2023b) and presents theoretically well-posed ensemble data-assimilation techniques adapted to this framework.

Some parts of this thesis will be highly inspired from the studies published during the PhD (Dufée et al., 2022; Dufée et al., 2023a; Dufée et al., 2023b).

Chapter 1

Filtering problem and data assimilation methods

Sequential data assimilation aims at incorporating real-world data in order to correct a realization or a set of realizations resulting from the forecast of a dynamical numerical model. This chapter will formulate data assimilation methods in the framework of stochastic filtering problems and will recap the existing methods able to solve it.
In what follows, the reference trajectory of the dynamical system will be denoted by $(X_t)_t$ with $X_t \in \mathbb{R}^{D_X}$ for all t, where D_X denotes the dimension of the state space. The observations $(Y_k)_k$ are assumed to be obtained at discrete times $(t_k)_k$ from an observation operator $\mathbf{H} : \mathbb{R}^{D_X} \to \mathbb{R}^{D_Y}$, such that, for any k,

$$Y_k = \mathbf{H}(X_{t_k}) + \eta_k, \tag{1.1}$$

where D_Y is the dimension of the observational space and η_k is the observation error.

1.1 The filtering problem

The filtering problem consists in determining, at every time t, the posterior distribution of the underlying truth of the system, denoted by π_t^a, given the observations obtained up to time t. For the sequence of observation acquisition times $(t_k)_k$, the filtering distribution is defined, for any k and any bounded measurable function φ, by

$$\pi_k^a \varphi = \mathbb{E}[\varphi(X_{t_k})|\mathcal{Y}_k], \tag{1.2}$$

where $\mathcal{Y}_k = \sigma(Y_{k:})$ is the filtration associated to the observations gathered up to time t_k.

This filtering problem can be recursively solved through a prediction-correction procedure. Let us define the prediction density, for any k and any bounded measurable function

φ, by

$$\pi_k^f \varphi = \mathbb{E}[\varphi(X_{t_k})|\mathcal{Y}_{k-1}]. \tag{1.3}$$

We also define the likelihood function $L : \mathbb{R}^{D_Y} \to \mathbb{R}$ such that, for any k and any Borelian set $B \subset \mathbb{R}^{D_Y}$,

$$\mathbb{P}(Y_k \in B|X_{t_k} = x_k) = \int_B L(y_k - \mathbf{H}(x_k))dy_k. \tag{1.4}$$

An usual choice for the likelihood is a centered Gaussian density of a prescribed standard deviation σ_k.
Then, the prediction step is defined by the transport of the posterior density of the previous step by the model, namely, for any k and any Borelian set $A \subset \mathbb{R}^{D_X}$,

$$\pi_k^f(A) = \int M_k(x_{k-1}, A)d\pi_{k-1}^a(x_{k-1}) \tag{1.5}$$

where M_k is the probability transition kernel related to the underlying dynamical model (it can be either deterministic or stochastic depending on the dynamical model at hand). Finally, Bayes' theorem yields the correction step:

$$\pi_k^a(A) = \frac{\int_A L(Y_k - \mathbf{H}(x_k))d\pi_k^f(x_k)}{\int L(Y_k - \mathbf{H}(x_k))d\pi_k^f(x_k)} \tag{1.6}$$

Considering that the posterior distribution is actually computed, an additional goal is to exhibit some optimal estimates X_k^a of the underlying truth of the system. In the classical assumption of Gaussianity of the posterior distribution, the usual estimate is the posterior mean

$$X_k^a = \int x_k d\pi_k^a(x_k). \tag{1.7}$$

In case of non Gaussianity, the maximum of the posterior distribution is another natural candidate, as it matches the posterior mean in the Gaussian case.

In practice, for ensemble methods like ensemble Kalman methods or particle filters, the posterior distribution will not be exactly computed, but estimated by an ensemble empirical distribution evaluated from a set of realizations of the dynamical system. In this case, the empirical mean is the natural estimate of the truth. Other non ensemble-based techniques, like variational methods aim at exhibiting this state estimate only, regardless of the distribution. All these methods are the topic of the rest of this chapter.

In what follows, the observation operator $\mathbf{H}$ will be assumed to be linear, and the observation error will be assumed to be Gaussian with a constant covariance error covariance matrix $\mathbf{R}$, meaning that, for all k,

$$\eta_k \sim \mathcal{N}(0, \mathbf{R})$$

Data assimilation techniques are usually split into a forecast step, essentially propagating the model in time, and an analysis step that incorporates the observation through filtering. In the following, the focus will be on the analysis (update) step, as the forecast step will not differ from one method to another. Modifications of the forecast step accounting for data will be considered in Section 3.3.

Note that the dependance on time t_k will be dropped for the description of the methods.

1.2 Variational data assimilation

Variational data assimilation basically consists, given a prior (or background) estimate of the state of the system X^f with background error covariance matrix $\mathbf{B}$ and an observation Y, to solve the most generally ill-posed problem $\mathbf{H}X = Y$. A regularization term, accounting for the departure from the prior estimate is added, so that the 3D-Var method finally consists in minimizing the following cost function

$$J(X) = \frac{1}{2}||\mathbf{H}X - Y||_{\mathbf{R}^{-1}} + \frac{1}{2}||X - X^f||_{\mathbf{B}^{-1}}. \tag{1.8}$$

Note that the cost function reads formally the same without the linear hypothesis on $\mathbf{H}$, replacing $\mathbf{H}X$ by $\mathbf{H}(X)$. In this latter case, note however that J is no longer a quadratic function.

The function J is strictly convex, so it has a unique minimum, defined as the posterior estimate X^a, that satisfies

$$\nabla J(X^a) = 0. \tag{1.9}$$

The gradient and the Hessian matrix of J read as follows:

$$\nabla J(X) = \mathbf{H}^T\mathbf{R}^{-1}(\mathbf{H}X - Y) + \mathbf{B}^{-1}(X - X^f) \tag{1.10}$$

$$\nabla^2 J = \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} + \mathbf{B}^{-1}. \tag{1.11}$$

The posterior estimate X_a can thus be expressed by

$$X^a = (\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} + \mathbf{B}^{-1})(\mathbf{H}^T\mathbf{R}^{-1}Y + \mathbf{B}^{-1}X^f) = -(\nabla^2 J)^{-1}\nabla J(0), \tag{1.12}$$

which recovers an usual result in optimization theory. For a more explicit writing of X^a , one can apply the Sherman-Morrison-Woodbury formula (Reinhardt, 2020) to the Hessian matrix and finally get

$$X^a = X^f - \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}(\mathbf{H}X^f - Y). \tag{1.13}$$

Note that the optimization problem does not necessarily need the potentially very costly inversion of the Hessian matrix. Optimal control methods that rely on an explicit computation of the tangent linear model and its adjoint (Le Dimet and Talagrand, 1986), can be used to minimize J in an efficient way. We may also refer to the applications of preconditioning (Gürol et al., 2014) and Gauss-Newton methods (Gratton et al., 2007) for variational methods.

In the case of a time series of observations $(Y_k)_{k=0,\dots,K}$ obtained over a time window $[t_0, t_K]$ and an initial estimate X_0^f, a similar procedure is the 4D-Var method, that consists in minimizing the functional

$$J(X) = \frac{1}{2}\sum_{k=1}^{K} ||\mathbf{H}X_k - Y_k||_{\mathbf{R}^{-1}} + \frac{1}{2}||X - X_0^f||_{\mathbf{B}^{-1}}, \tag{1.14}$$

where X_k is obtained from the transport of X by the dynamical model up to time t_k.

1.3 Ensemble Kalman filters

Contrary to the variational methods, ensemble methods are based on an empirical approximation of the tangent linear model from a set of realizations of the underlying dynamical model. Ensembles denoted $(X^{(n)})_{n=1,\dots,N}$ will contain N members in the state space $\mathbb{R}^{D_X}$. The empirical mean will be denoted by $\overline{X}$ and the (unbiased) ensemble covariance matrix is defined by

$$\mathbf{P} = \frac{1}{N-1}\sum_{n=1}^{N}(X^{(n)} - \overline{X})(X^{(n)} - \overline{X})^T. \tag{1.15}$$

Let us do a recap on different ensemble methods, from the classical EnKF to the localized ensemble square root filter.

1.3.1 Classical EnKF (with perturbed observations)

Starting from a prior forecast ensemble denoted $(X^{(n),f})_{n=1,\dots,N}$, the classical EnKF updates the ensemble as follows :

— The so called Kalman gain matrix is defined by

$$\mathbf{K} = \mathbf{P}^f \mathbf{H}^T (\mathbf{H}\mathbf{P}^f \mathbf{H}^T + \mathbf{R})^{-1}. \tag{1.16}$$

— The ensemble members are updated individually by

$$X^{(n),a} = X^{(n),f} - \mathbf{K}(\mathbf{H}X^{(n),f} - Y - \xi^{(n)}), \tag{1.17}$$

with $\xi^{(n)} \sim \mathcal{N}(0, \mathbf{R})$, which corresponds to random draws of the observation error.

1.3.2 Ensemble Square Root Filter (ESRF)

The square root formulation was designed in Whitaker and Hamill (2002b) and provides a fully deterministic way to compute the posterior ensemble, with no need to sample the $\xi^{(n)}$s. The method proposes a different update of the ensemble members based on the anomaly matrices. As consequences of the previous Kalman equations, we have:

— The innovation term $\mathbf{H}\overline{X}^f - Y$ weighted by the Kalman gain matrix $\mathbf{K}$ accounts for how much the ensemble mean needs to be corrected:

$$\overline{X}^a = \overline{X}^f - \mathbf{K}(\mathbf{H}\overline{X}^f - Y). \tag{1.18}$$

— The ensemble covariance matrix is updated accordingly:

$$\mathbf{P}^a = \mathbf{P}^f - \mathbf{K}\mathbf{H}\mathbf{P}^f. \tag{1.19}$$

Note that, according to (1.18), the posterior empirical mean is solution of the 3D-Var problem (1.13) with respect to the prior estimate $\overline{X}^f$ and its empirical covariance matrix $\mathbf{P}^f$. Dealing with very high dimensional systems in practice, the ensemble covariance matrices raise a tractability issue. An alternative is to rewrite the previous Kalman equations

only in terms of much more tractable matrices, namely the ensemble anomaly matrices

$$\mathbf{A} = [(X^{(1)} - \overline{X}), (X^{(2)} - \overline{X}), \dots, (X^{(N)} - \overline{X})],$$

which is a $D_X \times N$ matrix, so a much more convenient one to store and manipulate. The ensemble covariance matrix is related to the anomaly matrix by

$$\mathbf{P} = \frac{1}{N-1}\mathbf{A}\mathbf{A}^T, \tag{1.20}$$

and the same relation holds for the posterior (or analysis) matrices. The idea of the square root filter is to look for the posterior ensemble anomaly in the form

$$\mathbf{A}^a = \mathbf{A}^f\mathbf{S}, \tag{1.21}$$

where $\mathbf{S}$ is a $N \times N$ matrix, such that the posterior ensemble covariance matrix satisfies the Kalman equations, especially (1.19). Injecting (1.20) in equation (1.19) of the update of $\mathbf{P}^a$, one can find, after some algebraic manipulations, that

$$\mathbf{A}^a(\mathbf{A}^a)^T = \mathbf{A}^f\left(\mathbf{I} - \frac{1}{N-1}(\mathbf{H}\mathbf{A}^f)^T(\mathbf{H}\mathbf{P}^f\mathbf{H} + \mathbf{R})^{-1}\mathbf{H}\mathbf{A}^f\right)(\mathbf{A}^f)^T.$$

So, the matrix $\mathbf{S}$ satisfying equation (1.21) is such that

$$\mathbf{S}\mathbf{S}^T = \mathbf{I} - \frac{1}{N-1}(\mathbf{H}\mathbf{A}^f)^T(\mathbf{H}\mathbf{P}^f\mathbf{H} + \mathbf{R})^{-1}\mathbf{H}\mathbf{A}^f.$$

Using the Sherman-Morrison-Woodbury formula (Anderson and Moore, 2012), we deduce that

$$\mathbf{S} = \left(\mathbf{I} + \frac{1}{N-1}(\mathbf{H}\mathbf{A}^f)^T\mathbf{R}^{-1}\mathbf{H}\mathbf{A}^f\right)^{-\frac{1}{2}}. \tag{1.22}$$

Note that this matrix has very moderate size $N \times N$ and, in practice, Singular Value Decomposition (SVD) can be used in order to compute the inverse square root. By definition of A^a, the posterior forecast is given by

$$X^a = \overline{X}^a + \mathbf{A}^a = \overline{X}^a + \mathbf{A}^f\mathbf{S}. \tag{1.23}$$

It can be shown that the mean $\overline{X}^a$ can be written as a linear combination of the members of the forecast ensemble, namely

$$\overline{X}^a = \sum_{i=1}^{N} \omega_i X^{(i),f},$$

where ω_i is the i-th coordinate of the column vector

$$\omega = \frac{1}{N}\mathbf{1} - \frac{1}{N-1}\mathbf{S}^2(\mathbf{A}^f)^T\mathbf{H}^T\mathbf{R}^{-1}(\mathbf{H}\overline{X}^f - Y). \tag{1.24}$$

Finally, combining this with equation (1.23), the members of the posterior ensemble can be computed, for $n = 1, \ldots, N$,

$$X^{(n),a} = \sum_{i=1}^{N} \omega_i X^{(i),f} + \sum_{i=1}^{N} (X^{(i),f} - \overline{X}^f)\mathbf{S}_{in}. \tag{1.25}$$

This filter is part of what is generally called Linear Ensemble Transform Kalman Filter (LETKF), because members of the posterior ensemble are linear combinations of the forecast ones. Here

$$X^{(n),a} = \sum_{i=1}^{N} c_i^{(n)} X^{(i),f},$$

with

$$c_i^{(n)} = \omega_i + \mathbf{S}_{i,n} - \frac{1}{N}\sum_{j=1}^{N} \mathbf{S}_{j,n}.$$

Algorithm 1 recaps the ESRF.

Algorithm 1 Ensemble Square Root Filter (ESRF)

Input: Initial ensemble $(X_0^{(1)}, ..., X_0^{(N)})$, sequence of observations (Y_j), observation error covariance matrix $\mathbf{R}$, observation operator $\mathbf{H}$.

for $j = 1, 2, ...$ **do**
- Propagate all the ensemble members according to the underlying dynamics to get the forecast ensemble $(X_j^{(1),f}, ..., X_j^{(N),f})$
- Get observation Y_j.
- Compute the anomaly matrix $\mathbf{A}^f$
- Compute the matrix $\mathbf{S}$ according to formula (1.22), so that $\mathbf{A}^a = \mathbf{A}^f\mathbf{S}$.
- Compute the weight vector w with formula (1.24) such that $\overline{X}^a = \sum_{i=1}^{N} \omega_i X^{(i),f}$
- Update individually the ensemble members $X^{(n),a}$ by formula (1.25).

end for

Let us mention that such a superposition principle (stability of solutions by linear combinations) is in theory only valid for linear dynamics. In the context of nonlinear dynamics, this setting has to be understood as constraining the solution to live in a particular Hilbert space of Gaussian processes spanned by the ensemble members and associated to a norm defined by the inverse ensemble covariance matrix.

Ensemble-based methods are very convenient because, as mentioned, it spares the knowledge and computation of the tangent linear operator, which is approximated by an empirical one. The counterpart of that is the low-rank approximation of the "real" covariance matrix by the empirical covariance matrices of rank at most $N-1$, which is much smaller than D_X for computational affordability. It was shown (see Houtekamer and Mitchell (1998) for example) that this can result in filter divergence, as too much confidence is put in the forecast ensemble members, possibly preventing them from being strongly corrected, even if the observation would require it.

Two possible techniques were figured out in order to counter the latter effect. The first one, called inflation, consists in artificially increase the ensemble spread to try to recover some additional diversity theoretically lost through the low-rank approximation of the covariance matrix. The second is to turn the problem around and to artificially reduce the dimension of the state space by so called localization techniques, so that, locally, the ensemble covariance matrix becomes a very decent approximation of the "real" covariance matrix.

Both aspects are detailed hereafter.

1.3.3 Inflation

The need for inflation naturally appears when dealing with deterministic models, where the only randomness considered is most commonly related to the initial conditions. After some time, the ensemble may exhibit insufficient spread, causing a progressive drift from the observations and ultimately to filter divergence.

The most common way to inflate an ensemble is called multiplicative inflation. It can be applied either on the prior, before the assimilation, or on the posterior. Given an ensemble $(X^{(1)}, ..., X^{(N)})$ and its empirical mean $\overline{X}$, multiplicative inflation consists in building an inflated ensemble $(\tilde{X}^{(1)}, ..., \tilde{X}^{(N)})$ such that, for any n,

$$\tilde{X}^{(n)} = \overline{X} + \alpha(X^{(n)} - \overline{X}),$$

where $\alpha > 1$ is called the multiplicative inflation parameter. This artificially increases the variance by inflating the anomaly.

Another method is called additive inflation and simply consists in adding random perturbations to the posterior ensemble members after the filtering step, which possibly degraded the ensemble spread:

$$\tilde{X}^{(n),a} = X^{(n),a} + \eta^{(n)}$$

with $\eta^{(n)} \sim \mathcal{N}(0, \beta^2 \mathbf{P}^a)$ and $\beta > 0$ is the additive inflation parameter. This perturbs the new initial conditions before propagating by the numerical model in order to recover some diversity.

1.3.4 Localization for the ESRF

There exist two main ways to localize the filter. The R-localization (or local analysis) aims at transforming the observation error covariance matrix $\mathbf{R}$. Another way to achieve localization is to work on the ensemble covariance matrix $\mathbf{P}$ (B-localization or covariance localization). As equations (1.22-1.25) are only based on the use of the anomaly matrices instead of the whole covariance matrices $\mathbf{P}$, R-localization is the only one compatible with the square root filter, and more generally with the LETKF equations. Sakov and Bertino (2011) give details on both localization techniques and give numerical insights showing that both methods are likely to produce similar results.

The basic idea of R-localization is to update the ensemble members locally, grid point after grid point, according to local neighborhoods of "relevance". One asset of this method is to dramatically reduce the dimension of the state space for each local update.

In the following, the respective number of points on the simulation (state space) grid and on the observation grid will be denoted d_X and d_Y.

More precisely, for each grid point $k = 1, \ldots, d_X$, let us denote p_k the projection on the k-th coordinate (or grid point). For each k, the aim is to compute the local coefficients $c_{i,k}^{(n)}$ such that

$$p_k(X^{(n),a}) = \sum_{i=1}^{N} c_{i,k}^{(n)} p_k(X^{(i),f}). \tag{1.26}$$

In order to do this, R-localization is applied, meaning that the observation error covariance matrix will be converted into a local matrix for each grid point, and then the previous formulas (1.22-1.25) will apply for this local matrix. This is done in order to kill spurious

correlations.

For each $k = 1, \dots, d_X$, a diagonal localization matrix $\mathbf{C}_k$ of size $d_Y \times d_Y$ is defined and the diagonal coefficients are given, for $l = 1, \dots, d_Y$, by

$$(\mathbf{C}_k)_{ll} = \rho\left(\frac{d(l,k)}{r_{loc}}\right), \tag{1.27}$$

where :

— $d(l,k)$ denotes the distance between the grid point k and the observation site l.

— r_{loc} is the localization radius.

— ρ is the Gaspari-Cohn function, defined for any $z \geq 0$ by

$$\rho(z) = \begin{cases} -\frac{1}{4}z^5 + \frac{1}{2}z^4 + \frac{5}{8}z^3 - \frac{5}{3}z^2 + 1 & \text{if } \ z < 1 \\ \frac{1}{12}z^5 - \frac{1}{2}z^4 + \frac{5}{8}z^3 + \frac{5}{3}z^2 - 5z + 4 - \frac{2}{3z} & \text{if } \ 1 \leq z \leq 2 \\ 0 & \text{if } \ z \geq 2 \end{cases}.$$

It basically behaves like a Gaussian, but has compact support $[0, 2]$, so that the coefficients are set to zero as soon as $d(l,k)$ is twice as big as the localization radius.

Then the inverse observation error covariance matrix is modified by the Schür product

$$\mathbf{R}_k^{-1} = \mathbf{C}_k \circ \mathbf{R}^{-1}. \tag{1.28}$$

By doing so, the grid points outside this area are no longer taken into account, as the observation error is set to infinity at these points. Finally, $\mathbf{R}$ is replaced by $\mathbf{R}_k$ in formula (1.22), which gives a localized version $\mathbf{S}_k$ of $\mathbf{S}$, and consequently a localized version of equation (1.25).

In order to sum everything up, in this localized setting, the update formula is given by

$$X^{(n),a} = \sum_{k=1}^{d_X} p_k(X^{(n),a}) = \sum_{k=1}^{d_X} \sum_{i=1}^{N} c_{i,k}^{(n)} p_k(X^{(i),f}). \tag{1.29}$$

One theoretical drawback of this localization is the impossibility to guarantee the conservation of some physical balances from the prior to the posterior ensemble members, and more drastically the posterior is not solution of the global underlying dynamical system, which can become a serious issue.

Algorithm 2 Localized Ensemble Square Root Filter (LESRF)

Input: Initial ensemble $(X_0^{(1)}, ..., X_0^{(N)})$, sequence of observations (Y_j), observation error covariance matrix $\mathbf{R}$, observation operator $\mathbf{H}$, localization radius r_{loc}.

for $j = 1, 2, ...$ **do**
  Propagate all the ensemble members according to the underlying dynamics to get the forecast ensemble $(X_j^{(1),f}, ..., X_j^{(N),f})$
  Get observation Y_j.
  Compute the anomaly matrix $\mathbf{A}^f$
  for $k = 1, ..., d_X$ **do**
    Compute the diagonal localization matrix $\mathbf{C}_k$ in equation (1.27).
    Compute the localized inverse observation error covariance matrix $\mathbf{R}_k^{-1}$ with equation (1.28).
    Compute the corresponding matrix $\mathbf{S}_k$ according to formula (1.22).
    Compute the weight vector w_k with formula (1.24) such that $\overline{p_k(X)}^a = \sum_{i=1}^N \omega_i p_k(X^{(i),f})$
  end for
  Update individually and locally the ensemble members $X^{(n),a}$ by formula (1.29).
end for

1.4 Particle filters

Particle filters are part of the Monte-Carlo methods and is based, like the ensemble Kalman filters, on an ensemble representation of the distribution. It largely differs from the Kalman methodology in the sense that it no longer relies on a modification of the ensemble members through linear combinations. Instead, it affects and updates an importance weight to each ensemble member (also called particle in this framework) that represents how likely the particle is according to the observation.

Assuming that we start from a forecast ensemble $(X^{(n),f})_{n=1,...,N}$ and the empirical approximation of the prior distribution

$$\pi^f = \frac{1}{N} \sum_{n=1}^{N} \delta(X^{(n),f}),$$

where δ is the Dirac distribution, the "quality" of each prior particle is assessed through the likelihood function, and the particle is given its weight: for all n,

$$w^{(n)} = L(X^{(n),f} - Y),$$

where Y stands for the observation. Then we accordingly define the normalized weight vector

$$\overline{\mathbf{w}} = (\overline{w}^{(n)})_{n=1,...,N} \quad ; \quad \overline{w}^{(n)} = \frac{w^{(n)}}{\sum_{i=1}^{N} w^{(i)}}.$$

Without modifying the ensemble members, the posterior distribution will write

$$\pi^a = \frac{1}{N} \sum_{n=1}^{N} \overline{w}^{(n)} \delta(X^{(n),f}) \tag{1.30}$$

Another big difference with the EnKF is the relaxation of the assumptions of observation and model noise Gaussianity and model linearity from which the Kalman equations were originally derived.
The following parts detail the algorithmic translation of this distribution point of view.

1.4.1 Sampling Importance Resampling (SIR) algorithm

This is the most basic particle filter. Given the particles and the observation, the normalized weight vector $\overline{\mathbf{w}}$ is computed and a resampling procedure is applied on the ensemble, which consists in discarding particles with low weights, and to replace them by random duplicates of particles with higher weights.
Algorithm 1 gives an example of resampling procedure, called multinomial resampling. This algorithm is equivalent to draw the number of offspring for all the particles according

Algorithm 3 Multinomial Resampling

Input: Forecast ensemble $(X^{(1)}, ..., X^{(N)})$ and its normalized weight vector $\overline{\mathbf{w}}$.

Compute the cumulative normalized importance weight vector c using this **for** loop :
for $i = 1, ..., N$ **do**
  $c_i = \sum_{l=1}^{i} \overline{w}^{(l)}$
end for
for n=1,...,N **do**
  Draw one random variable $u_n \sim \mathcal{U}[0, 1]$.
  Find the resampling index $\psi(n) = \min\{i, u_n \leq c_i\}$.
  Replace particle n by particle $\psi(n)$.
end for

to a multinomial distribution with parameters $(N, \overline{\mathbf{w}})$.
The main drawback of resampling is the huge loss of diversity entailed by such a procedure.

In the extreme case where only one particle has a weight close to 1, and all the others an almost zero weight, the posterior ensemble will basically be N duplicates of one particle. This phenomenon is called ensemble degeneracy.
Degeneracy can be due to too frequent resampling procedures that are not necessary. A criterion to assess the level of degeneracy of the filter is called the Effective Sample Size (ESS), defined by

$$\mathrm{ESS}(\overline{\mathbf{w}}) = \left(\sum_{n=1}^{N} \overline{w}^{(n)2} \right)^{-1}.$$

This indicator measures the variability of the ensemble weights. If all particles have almost the same weights, then $\overline{w}^{(n)} \simeq \frac{1}{N}$ and the ESS is approximately N. On the contrary, if only a small group of particles have the largest weights (ensemble degeneracy), then the ESS will decrease dramatically. So the ESS is a good indicator to decide whether or not we should proceed to resampling. An usual choice is to perform it only if $\mathrm{ESS} < N^*$ where N^* is a fixed threshold. Algorithm 4 sums up the procedure for the SIR filter. In

Algorithm 4 Bootstrap particle filter

Input: Initial ensemble $(X_0^{(1)}, ..., X_0^{(N)})$.
  for $j = 1, 2, ...$ **do**
    Propagate all the particles according to the underlying dynamics to get the forecast ensemble $(X_j^{(1)}, ..., X_j^{(N)})$
    Get observation Y_j.
    Compute the normalized weight vector $\overline{\mathbf{w}}$
    if $\mathrm{ESS}(\overline{\mathbf{w}}) < N^*$ **then**
      Resample using Algorithm 3
    end if
  end for

practice, filter degeneracy is not due to too frequent assimilations. In high-dimensional systems, the usual choice of likelihood based on the L^2-norm over the observation grid is overdiscriminating and will, after very few steps of simulation, lead to a very small ESS (close to 1) and a highly degenerate filter. A first possible solution is to strongly constrain the truth to stay in the ensemble spread by controlling the initial condition to be very close (or even exactly) the truth, and to filter the ensemble very frequently. Additive inflation, presented in section 1.3.4, which is also called particle rejuvenation in the particle filter community, can also be used in deterministic settings to reintroduce some spread after the filtering, so that the posterior distribution is not exactly a single Dirac measure in the degenerate case.

Additional techniques were carried out by Cotter et al. (2020a,b) in order to counteract filter degeneracy and are detailed hereafter. Note that some of these methods will only apply in the framework of stochastic forecast models.

1.4.2 Jittering

Jittering aims at recovering the diversity of the posterior ensemble and is similar to the stochastic version of additive inflation. The objective of both procedures is the same, to perturb the posterior initial conditions before propagating them again according to the underlying dynamics.

In this method, particles resampled by the filter will no longer be duplicates of particles with higher weights, but will be generated according to a Metropolis-Hasting Markov Chain Monte Carlo (MCMC) method. One way to proceed is to slightly modify the noise according to which the stochastic PDE (SPDE) is driven for these particles. If the original one was driven by a standard Brownian motion B, the jittered one will be driven by

$$\rho B + \sqrt{1-\rho^2}\tilde{B}, \tag{1.31}$$

where $\tilde{B}$ is a new Brownian motion independent from B, and ρ is called the jittering parameter. It controls the amplitude of the perturbation brought by the new noise compared to the original one. In practice, ρ is chosen sufficiently close to 1 so that particles do not go to far from the duplicate, but small enough so that it indeed has an impact (typically $\rho = 0.999$ in our experiments, this way the new noise is around 20 times weaker than the original one).

Then the new realization $\tilde{X}$ will be accepted and replace the former realization X with probability

$$a(X, \tilde{X}) = \min\left(1, \frac{L(\tilde{X} - Y)}{L(X - Y)}\right). \tag{1.32}$$

Algorithm 5 recaps what happens in one jittering step.

1.4.3 Tempering

The idea of tempering is the following: given a forecast ensemble which has a bad ESS, the variance of the distribution is artificially increased using a rescaling of the log-likelihood by a scalar $\phi \in]0, 1[$, called the temperature by analogy with annealing algo-

Algorithm 5 Jittering step at time t_j using MCMC

Input: Ensemble $(X_{t_j}^{(1)}, ..., X_{t_j}^{(N)})$, the number of MCMC steps M and the jittering parameter ρ.

```
for m = 1, ..., M do
  for n = 1, ..., N do
    Generate a new proposal X̃_{t_j}^{(n)} starting from X_{t_{j-1}}^{(n)} with the jittered noise (1.31).
    Compute the weight of this new proposal w̃^{(n)} = L(X̃_{t_j}^{(n)} − Y_{t_j}).
    if w̃^{(n)} ≥ w^{(n)} then
      The acceptance probability a is equal to 1, so update X_{t_j}^{(n)} = X̃_{t_j}^{(n)}.
    else
      a = w̃^{(n)} / w^{(n)}.
      Draw a random variable u ∼ U([0, 1]).
      if u ≤ a then
        Update X_{t_j}^{(n)} = X̃_{t_j}^{(n)}
      end if
    end if
  end for
end for
```

rithms. This temperature is chosen so that the ESS is good (above a given threshold N^*) for this tempered distribution, enabling the application of resampling (jittered or not). The next temperature will be chosen in the interval $]\phi, 1[$, and this is iterated until the initial distribution is fully recovered, corresponding to temperature 1.

In other words, the Radon-Nikodym distribution of the posterior distribution with respect to the forecast, which is exactly the likelihood function $\frac{d\pi^a}{d\pi^f} \propto L(\cdot - Y)$ leads to a bad ESS ensemble, so it is decomposed as follows:

$$\frac{d\pi^a}{d\pi^f} = \frac{d\pi_1}{d\pi_0} ... \frac{d\pi_R}{d\pi_{R-1}}, \tag{1.33}$$

where $0 = \phi_0 < \phi_1 < ... < \phi_R = 1$ is the sequence of temperatures such that $\pi_0 = \pi^f$, $\pi_R = \pi^a$, and for $r = 1, ..., R$,

$$\frac{d\pi_r}{d\pi_{r-1}} = L(\cdot - Y)^{\phi_r - \phi_{r-1}}. \tag{1.34}$$

and $\mathrm{ESS}(\overline{\mathbf{w}}_r)$ denotes the ESS of the weight vector computed with this likelihood.

These tempered distributions should progressively lead to a balanced and diverse posterior ensemble.

The procedure is recapped in Algorithm 6.

Algorithm 6 Tempered particle filter

Input: Initial ensemble $(X_0^{(1)}, ..., X_0^{(N)})$.
  for $j = 1, 2, ...$ **do**
    Propagate all the particles according to the underlying dynamics to get the forecast ensemble $(X_{t_j}^{(1)}, ..., X_{t_j}^{(N)})$
    Get observation Y_{t_j}.
    Compute the normalized weight vector $\overline{\mathbf{w}}$
    Set $\phi_0 = 0, r = 1$
    while $\mathrm{ESS}(\overline{\mathbf{w}}_{r-1}) < N^*$ **do**
      Set $p = 2$ and $\phi_r = \frac{1}{p}$.
      while $\mathrm{ESS}(\overline{\mathbf{w}}_r) < N^*$ **do**
        $p = p + 1$
      end while
      Resample using Algorithm 3 with $\overline{\mathbf{w}}_r$ as the weight vector.
      Apply jittering using Algorithm 5 with the tempered likelihood (10).
      r=1
    end while
    Resample and jitter with $r = R$ and $\phi_R = 1$ to get the posterior ensemble.
  end for

1.5 Convergence results

This section recaps the existing convergence results for ensemble methods when the ensemble size N goes to $+\infty$.

1.5.1 Ensemble Kalman filters

The convergence properties are quite straightforward to obtain at time 0. As the initial ensemble is generated from i.i.d. samples of the initial background distribution $p_0 \sim (X_b, \mathbf{B})$, the strong law of large numbers provides, when $N \to +\infty$, the convergence of the empirical mean and covariance matrix respectively towards X_b and $\mathbf{B}$. The idea is to propagate this result in time with continuity arguments for both the forecast and the analysis step. We refer to Le Gland et al. (2011) or Mandel et al. (2011) for proofs for the standard EnKF with perturbed observations. Kwiatkowski and Mandel (2015) formulated the results of convergence for the ensemble square root filter:

Theorem 1. *Assume that the state space is a finite dimensional or separable Hilbert space, the initial state X_0 has a distribution such that all moments exist, $\mathbb{E}(|X_0|^p) < \infty$ for $1 \leq p < \infty$, the initial ensemble $(X_0^{(n)})_{n=1,\dots,N}$ is an iid sample from this distribution, and the model is linear. Then, for any iteration k, the ensemble mean $\overline{X}_k^{N,a}$ and covariance matrix $\mathbf{P}_k^{N,a}$ resulting from the ensemble square root filter respectively converge to the mean $\overline{X}_k^a$ and covariance matrix $\mathbf{P}$ resulting from the (non ensemble) Kalman filter, in L^p for $1 \leq p < \infty$, as $N \to \infty$, with the convergence rate $\frac{1}{\sqrt{N}}$. Namely,*

$$||\overline{X}_k^{N,a} - \overline{X}_k^a||_p \leq \frac{C_{p,k}}{\sqrt{N}}, \tag{1.35}$$

$$||\mathbf{P}_k^{N,a} - \mathbf{P}_k^a||_p \leq \frac{C_{p,k}}{\sqrt{N}}, \tag{1.36}$$

for all $1 \leq p < \infty$, and all $N \in \mathbb{N}$, where $C_{p,k}$ is a constant that does not depend on the ensemble size or the dimension of the state space.

This theorem quantifies the convergence of the ensemble square root filter distribution towards the Kalman filter distribution which is, in the linear and Gaussian case, solution of the filtering problem. Note that, in a non linear or non Gaussian case, this theorem ensures the convergence towards a distribution which is not necessarily the filtering distribution.

We may also refer to Bishop and Del Moral (2023) for recent progress on theoretical results for the ensemble Kalman-Bucy filter in the linear Gaussian case that, among many other results, spares the time dependance in the previous fluctuation estimate.

1.5.2 Particle filters

As particle filters deal with probability distributions, the convergence results are stated for the topology associated with the weak convergence of distributions. The law of large numbers and a central limit theorem were respectively proven in Del Moral (1997) and Del Moral and Guionnet (1999). We also refer to Crisan and Doucet (2002) for an early overview of convergence results for particle filters. This first theorem states a law of large numbers for the SIR algorithm.

Theorem 2. *Under the assumptions that the transition kernel driving the underlying process X is Feller and that the likelihood function L is bounded, continuous and strictly positive, then, for any iteration k, the posterior distribution $\pi_k^{N,a}$ resulting from the SIR*

algorithm (equation (1.30)) converges almost surely towards the filtering distribution π_k^a *(equation (1.2)). Namely, for any test function* $f : \mathbb{R}^d \to \mathbb{R}$,

$$\lim_{N\to+\infty} \int f(x) \mathrm{d}\pi_k^{N,a}(x) = \int f(x) \mathrm{d}\pi_k^a(x)$$

The assumptions of the theorem are essentially here to ensure the continuity of the time transition between $\pi_k^{N,a}$ and $\pi_{k+1}^{N,a}$ and hence to propagate in time the convergence property, as it was needed for the ensemble square root filter convergence result. This law of large numbers for the SIR can be accompanied with a convergence result for the estimate MSE:

Theorem 3. *Under the same hypothesis, for any test function* $f : \mathbb{R}^d \to \mathbb{R}$ *and any* $N \in \mathbb{N}$,

$$\mathbb{E}\left[\left(\int f \mathrm{d}\pi_k^{N,a} - \int f \mathrm{d}\pi_k^a\right)^2\right] \leq C_k \frac{||f||^2}{N}.$$

These results are much stronger than its EnKF counterparts. They ensure the convergence towards the filtering distribution, even out of the classical linear Gaussian setting. With this respect, particle filter has an undeniable theoretical advantage in a data assimilation perspective. We also refer to Beskos et al. (2014) for the theoretical justification of tempering and jittering procedures.

Nevertheless, Bishop and Del Moral (2023) point out that, in a linear Gaussian setting, the particle filter is unable to keep track of unstable reference signals, with the example of McKean-Vlasov diffusions, whereas the EnKF, which is in this case a mean-field approximation of the diffusion process, is able to do it. This illustrates that the EnKF, with its ability to explore the neighborhood of the state space spanned by the ensemble members, can compensate an unstable forecast step, whereas the SIR algorithm is more dependent on the forecast reliability.

1.6 Synthesis and choice of the PhD focus

After this quick overview, let us point out that most of the contents of this thesis will be focused on the ensemble square root filter and its localized version. The choice to elaborate on this family of methods is motivated by the early experiments of the PhD, and we refer to the conclusive chapter for more precise explanations. As a short summary of this conclusion, despite an undeniably more reliable theory, the particle filter and its

variations with tempering and jittering did not prove effective enough (in terms of MSE) compared to the ensemble square root filter, and even less compared to its localized version. In order to make particle filter effective (for a short period of time) in our case, we needed a very careful control of the ensemble spread around the observations, a quite linear dynamical behavior, and a much higher computational cost, coming either from a bigger ensemble or jittering and tempering procedures. As a consequence, the contents of the thesis will be mainly focused on the EnKF, even if some bridges with particle filters and variational methods will be attempted in the end of Chapter 4.

Chapter 2

Stochastic modeling under Location Uncertainty (LU)

The purpose of this chapter is to present the principles of modeling under Location Uncertainty and how it modifies the formulation of transport equations. The SQG equations are stated in this framework and noise generation procedures are suggested.

2.1 General principles of LU

Location Uncertainty models the impact of the small scales on the global flow dynamics. It is a stochastic framework that relies on the following decomposition of the Lagrangian velocity of a fluid particle positioned at X_t in a spatial domain $\Omega \subset \mathbb{R}^d$:

$$\mathrm{d}X_t = v(X_t, t)\mathrm{d}t + \sigma(X_t, t)\mathrm{d}B_t, \tag{2.1}$$

where v is a smooth-in-time component (referred to as the large-scale component in the following) and $\sigma \mathrm{d}B_t$ is a highly oscillating random component, built from a (cylindrical) Wiener process B_t (ie a well-defined Brownian motion taking values in a functional space) (Prato and Zabczyk, 1992). This latter component is uncorrelated in time. Let us point out that the above relation is rigorously understood in its integral form and corresponds thus only to a practical shortcut notation. The correlation operator σ is defined through a bounded matrix kernel $\hat{\sigma}$, for any function $f \in (L^2(\Omega))^d$, by

$$\sigma(x, t)f = \int_\Omega \hat{\sigma}(x, y, t)f(y)\mathrm{d}y. \tag{2.2}$$

From this correlation operator, the covariance matrix kernel $\hat{q}$ reads

$$\hat{q}(x, y, t) = \int_\Omega \hat{\sigma}(x, x', t)\hat{\sigma}(y, x', t)^T \mathrm{d}x', \tag{2.3}$$

and the associated covariance operator Q is given by

$$Q(x,t)f = \int_\Omega \hat{q}(x,y,t)f(y)\mathrm{d}y. \tag{2.4}$$

The random velocity is Gaussian and distributed as

$$\sigma \mathrm{d}B_t \sim \mathcal{N}(0, Qdt). \tag{2.5}$$

Moreover, at each time t, the covariance operator $Q(\cdot,t)$ is self-adjoint, non-negative definite and compact. Thus it admits an orthonormal eigenfunction basis $\{\phi_n(\cdot,t)\}_{n\in\mathbb{N}}$ with non-negative eigenvalues $(\lambda_n(t))_{n\in\mathbb{N}}$. This entails an alternative convenient spectral definition of the noise as

$$\sigma(x,t)\mathrm{d}B_t = \sum_{n\in\mathbb{N}} \sqrt{\lambda_n(t)}\phi_n(x,t)d\beta_t^n, \tag{2.6}$$

where the β^n are i.i.d standard Brownian motions. From (2.6), the noise variance tensor a is then defined by

$$a(x,t) = \sum_{n\in\mathbb{N}} \lambda_n(t)\phi_n(x,t)\phi_n(x,t)^T. \tag{2.7}$$

It can be noticed that the variance tensor has the physical dimension of a viscosity (ie m^2/s). The properties and structure of the noise will of course depend on the procedure used to generate the orthonormal basis functions. The one used in our experiments will be presented later. In the deterministic case, a transported tracer Θ has zero material derivative :$D_t\Theta = \partial_t\Theta + v\cdot\nabla\Theta = 0$. In the LU framework, the material derivative is replaced by the stochastic transport operator

$$\mathbb{D}_t\Theta = \mathrm{d}_t\Theta + (v^*\mathrm{d}t + \sigma\mathrm{d}B_t)\cdot\nabla\Theta - \frac{1}{2}\nabla\cdot(a\nabla\Theta)\mathrm{d}t, \tag{2.8}$$

where

— $\mathrm{d}_t\Theta = \Theta(x,t+\mathrm{d}t) - \Theta(x,t)$ is the forward time increment of the tracer.

— The effective advection velocity is defined by

$$v^* = v - \frac{1}{2}\nabla\cdot a. \tag{2.9}$$

— The term $\sigma\mathrm{d}B_t\cdot\nabla\Theta$ is a non-Gaussian multiplicative noise corresponding to the

tracer's transport by the small-scale flow.

— The last term is a diffusion term, as the variance tensor a is definite positive.

Note that the expression of the transport operator is given here for a divergence-free noise. In the case of a compressible random field, the modified advection incorporates an additional term related to the noise divergence (Resseguier et al., 2017a). One key property of LU (for an incompressible random term) is that, under the same ideal boundary conditions as in the deterministic case, it conserves the energy of the transported random tracer Θ:

$$\mathrm{d}\int_{\Omega}\Theta^2(x)\mathrm{d}x = 0. \tag{2.10}$$

Most importantly, this energy conservation property holds pathwise (i.e for any realization of the Brownian noise) (Bauer et al., 2020a; Resseguier et al., 2017a). A few general remarks on the stochastic transport operator can be done at this point. Compared to the usual material derivative, it is embedded with new additional terms. The viscosity associated with the mixing of the advected quantity by the noise is now described by the variance tensor a. It is no longer a scalar variable but a (positive definite) diffusion matrix. It is thus richer than the classical eddy viscosity models (Boussinesq, 1877) derived from an analogy with the molecular friction (the so called Boussinesq's assumption). The multiplicative noise corresponds to a backscattering of energy that is exactly compensated by the loss of energy brought by the stochastic diffusion (meaning the second order differential term associated to the stochastic transport operator in equation (2.8), which is indeed diffusive, in the sense that it dissipates energy). This balance is the reason why we have the energy conservation property and can be seen as a redistribution of the tracer (in the same way as a deterministic advection equation). Finally, the modified advection corresponds to the statistical effect induced by the small-scale inhomogeneity on the large scale component structuration (Bauer et al., 2020a). With this term, the particles are statistically encouraged to migrate from regions with high variance (i.e. of high kinetic turbulent energy) to regions with low variance. This modification of advection, the backscattering carried by the advection noise and the balanced diffusion matrix are the three distinctive ingredients of the modeling under location uncertainty. It is important to point out that this balance only holds globally. Locally in space, these three terms may play their own role without any balance. This scheme has been used for the modeling of large scale flow dynamics for numerous flow configurations. From a data assimilation point of view, it has been used in an optimal control setting for 3D flow reconstruction

and for joint parameter estimation in a 1D shallow water model (Chandramouli et al., 2020; Yang and Mémin, 2019), where it was shown to provide an interesting trade-off with the so called weak constraint variational assimilation.

2.2 The LU-SQG system

In this work we apply the LU framework to the 2D Surface Quasi-Geostrophic (SQG) model. This model constitutes an idealized dynamics for surface oceanic currents. Yet, it involves many realistic features of real world phenomena such as front, strong multiscale eddies, driven by a 3D like turbulent energy cascade, see Constantin et al. (1999); Lapeyre and Klein (2006) for details.

The SQG model relies on a (deterministic or stochastic) transport equation of the buoyancy field b (in $m.s^{-2}$) (or density fluctuation $b = -\frac{g\rho'}{\rho_0}$, with g the gravitational constant and ρ' the density fluctuation around a base constant density ρ_0):

$$D_t b = 0, \tag{2.11}$$

coupled with a kinematic equation

$$b = N(-\Delta)^{\frac{1}{2}}\psi \tag{2.12}$$

and the incompressibility constraints

$$v = \nabla^{\perp}\psi; \quad \nabla \cdot \sigma dB_t = 0, \tag{2.13}$$

linking the buoyancy field to the velocity field v, where ψ (in $m^2.s^{-1}$) is the stream function, N_{strat} (in s^{-1}) is the stratification, and $\nabla f = (\partial_x f, \partial_y f)^T$, $\nabla^{\perp} f = (-\partial_y f, \partial_x f)$ respectively stand for the classical and orthogonal gradients, while Δ denotes the Laplace operator. As shown in Resseguier et al. (2017b) for noise of moderate amplitude, the stochastic dynamics simply consist in replacing the material derivative D_t by the stochastic transport operator $\mathbb{D}_t$ given by (2.8). We show in Figure 2.1 several realizations of the stochastic dynamics obtained with a so called SVD noise (see Section 2.3.2) in a four-vortex configuration with double periodic boundary conditions (see Figure 3.1) and the numerical setup described in Section 3.1. Note that for noise of higher amplitude, the elliptic relation (2.12) is modified and involves an additional diffusion operator associated

to the noise Resseguier et al. (2017c). This case will not be considered here.

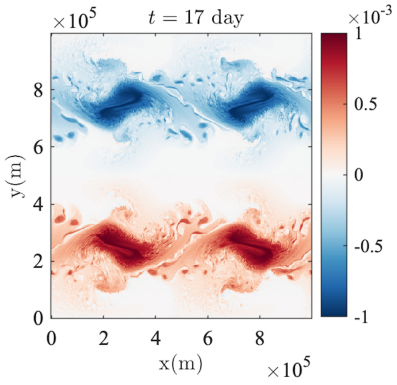
It can be noticed that the large-scale component of the different runs remain quite close after 17 days. They mainly differ by their small-scale features and vortices (subfigures (a,b,c,d)). We also added the states of these realizations at day 72 as an example to point out that the system is chaotic and may lead to significant large scale differences (subfigures (e,f,g,h)), at least on the time range studied in this work. This system is unforced and involves in its numerical implementation a small hyperviscosity term. It is hence decaying in the long run. However, as shown in Figure 2.1, the system remains turbulent at the end of the time period on which we will focus in this study. So called mesoscale eddies as well as submesoscale eddies, filaments and fronts can be observed for the different realizations at day 72 of Figure 2.1.

2.3 Noise generation procedures

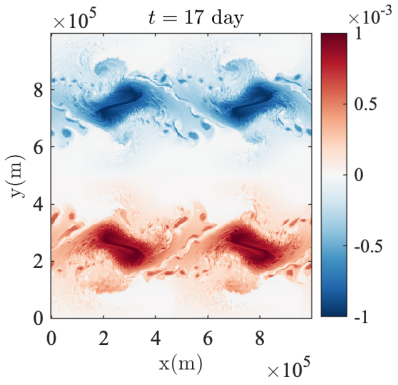
The LU scheme obviously depends on the noise parameterization chosen. For instance, for a homogeneous noise associated to Fourier basis functions, the variance tensor is homogeneous and constant (even diagonal for a divergence free flow). Hence, there is no modified advection. Note however that homogeneous noise is in theory restricted to periodic domains or to dynamics defined over the whole space. For a stationary noise, the variance field is constant in time and thus not related to the evolving large-scale components. The ability to build a flow-dependent noise enables us to improve probabilistic forecasting skills (Brecht et al., 2021). For the SQG dynamics, several noise parameterizations have been compared and assessed through different statistical proper scores (Resseguier et al., 2020a). One of the main findings of this work is that a time-varying inhomogeneous parameterization, termed as SVD (Singular Value Decomposition) noise, provides the best quantification results. Still, this method is very costly compared to stationary noises based on purely data-driven methods. The rest of this section is dedicated to both types of methods, which will be used in the experiments presented afterwards.

2.3.1 Data-driven stationary noise (POD noise)

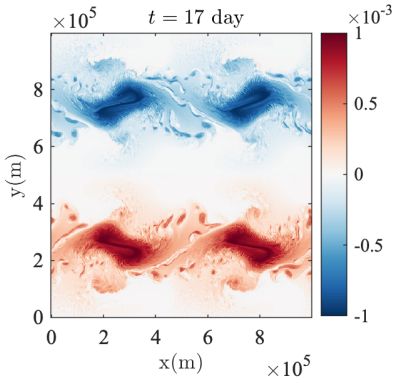
This method relies on a data-driven method called proper orthogonal decomposition (POD) to estimate the empirical orthogonal functions in the spectral representation of equation (2.6). By a slight abuse of notation in the following, this noise will be referred



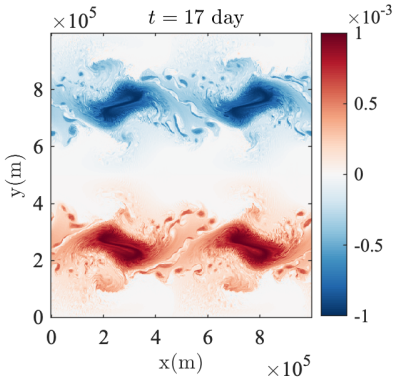
(a) Realization 1



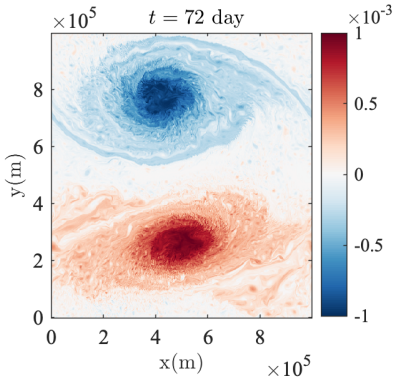
(b) Realization 2



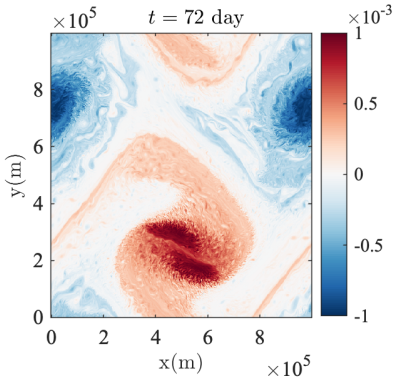
(c) Realization 3



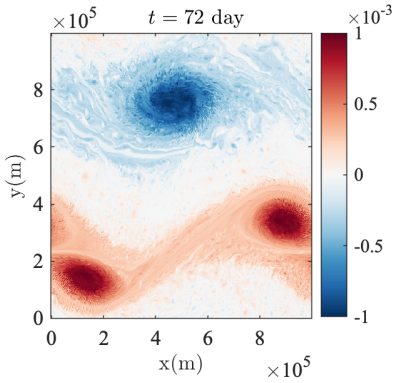
(d) Realization 4



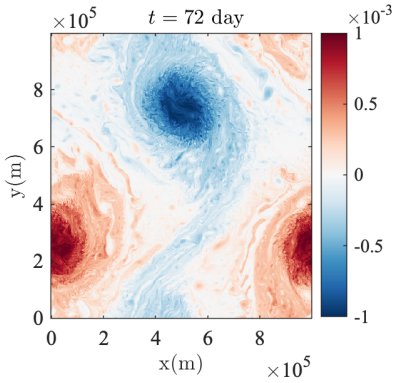
(e) Realization 1



(f) Realization 2



(g) Realization 3



(h) Realization 4

Figure 2.1 – Four different realizations for the stochastic dynamics at day 17 and day 72.

to as POD noise. We give some details in what follows.

The data considered here is a time-series of velocity snapshots $\{v(x,t_n), n=1,...,N_t\}$ coming from pre-processed high-dimensional data. From these, the temporal mean $\overline{v}$ is computed and then removed to get the fluctuation snapshots around this mean $v' = v - \overline{v}$. The temporal covariance tensor associated to v' is then defined by $C = (c_{i,j})_{i,j=1,...,N_t}$ with

$$c_{i,j} = \frac{1}{|\Omega|}\int_\Omega v'(x,t_i)v'(x,t_j)dx := \frac{1}{|\Omega|}\langle v'(\cdot,t_i), v'(\cdot,t_j)\rangle_\Omega \tag{2.14}$$

The covariance operator being compact, self-adjoint and definite positive, the eigen problem

$$CB = \Lambda B$$

can be solved in order to estimate the set of real positive decaying eigenvalues $\Lambda = (\lambda_n)_{n=1,...,N_t}$ and the corresponding set of temporal orthogonal eigenvectors $B = (b_{i,j})_{i,j=1,...,N_t} = (b_i(t_j))_{i,j=1,...,N_t}$. These temporal modes are normalized such that

$$\overline{b_i(\cdot)b_j(\cdot)} = \lambda_i \delta_{i,j}. \tag{2.15}$$

Then the spatial modes $(\phi_n)_{n=1,...,N_t}$ are defined by

$$\phi_n(x) = \overline{b_n(\cdot)v'(x,\cdot)} \tag{2.16}$$

and are orthonormal with respect to the spatial inner product defined in (2.14).
Then, each snapshot can be recovered through the following spectral decomposition :

$$v(x,t_i) = \overline{v}(x) + \sum_{n=1}^{N_t} b_n(t_i)\phi_n(x). \tag{2.17}$$

We make the additional assumption that, for Δt small enough, the small-scale random velocity $\dfrac{\sigma \mathrm{d}B_t}{\Delta t}$ only lives in the space spanned by only a subset of small modes $(\phi_k)_{M\leq k\leq M+K}$, meaning that

$$\frac{1}{\Delta t}\sigma(x)\mathrm{d}B_t \simeq \sum_{k=M}^{M+K}\sqrt{\lambda_k}\phi_k(x)\xi_k, \tag{2.18}$$

where $(\xi_k)_k$ are i.i.d. standard Gaussian variables.

The noise modes are divergence-free and stationary by construction, so the global structure of the noise will not vary in time. The POD noise is very easy and cheap to compute. Still, in case of chaotic geophysical models like the SQG system, it may not offer

enough spread from a data assimilation perspective.

2.3.2 Flow-driven non stationary noise (SVD noise)

An alternative is to build a time-evolving noise that embraces the small-scale variability at any time. The method relies on the creation of pseudo-observations at each point of the simulation grid, and then on the diagonalization of the associated empirical matrix to extract a proper basis to decompose the noise. Here the domain Ω is a two-dimensional regular grid of size $d_x \times d_y$. The pseudo observations are constructed from the running velocity fluctuations around a sample mean, more precisely around a velocity field composed of local spatial means computed at each grid point. At each time t and each grid point $x_{i,j}$, a spatial window $W_{i,j}$ of size $d_w \times d_w$ (with d_w odd), much smaller than the whole simulation grid, is built around the point, together with the model boundary conditions (periodicity, replication,...) if the current point is on the border. Then a pseudo-observation is given by a draw of the velocity in the following set :

$$I(x_{i,j}, t) = \{v(x_{k,l}, t) \,|k, l \in W_{i,j}\} \,. \tag{2.19}$$

Proceeding to n_o draws within the window, and iterating over all the grid points, a global pseudo-observation matrix V is built :

$$V = \begin{pmatrix} v_1^1 & \cdots & v_1^{n_o} \\ \vdots & \cdots & \vdots \\ v_{d_x d_y}^1 & \cdots & v_{d_x d_y}^{n_o} \end{pmatrix}, \tag{2.20}$$

whose size is $(2 \times d_x \times d_y) \times n_o$ (the 2 comes from the two components of the velocity v). Then the mean over the n_o pseudo-observations $\langle V \rangle$ is removed

$$V' = V - \langle V \rangle \tag{2.21}$$

and Singular Value Decomposition is applied to the fluctuation matrix V' to diagonalize the corresponding second order empirical moment. This way, the matrix Φ of the left eigenvectors on which we can decompose the noise as in (2.6) is obtained.
Let us denote ℓ the simulation grid scale and $L = d_w \ell$ the spatial scale of the sliding window used to compute the noise. The procedure described above provides a noise $\sigma_L \mathrm{d}B_t$ at scale L, which is artificial and it must be downscaled to the true simulation scale ℓ.

Kadri Harouna and Mémin (2017) proposed a rescaling of the variance tensors based on 3D turbulent cascade assumption :

$$a_\ell = \left(\frac{\ell}{L}\right)^{\frac{4}{3}} a_L,$$

which relies on an estimation of the velocity fluctuations at the simulation scale ℓ. Then the effective noise for the simulation grid is

$$\sigma \mathrm{d}B_t = d_w^{-\frac{2}{3}} \sigma_L \mathrm{d}B_t. \tag{2.22}$$

Let us note that the SQG model has the particularity to be associated to 3D like spectrum while it is a 2D flow. Resseguier et al. (2020b) also show that the Uncertainty Quantification (UQ) results for the SVD noise forecast reliability are good for all the metrics presented, in particular they are much better than Perturbations of the Initial Conditions (PIC), often used in ensemble data assimilation and that tends to produce underdispersive ensembles.. It also proves better than the POD noise, whose stationarity is highly detrimental for non-stationary unforced flows we are focusing here. The SVD noise has the advantage to be purely flow-driven; it requires no data and remains simple to implement. As the basis depends both on time and on the ensemble member at hand, some adaptations should probably be devised for realistic models in order to reduce the noise computational cost. One can envisage, for instance, a combination with wavelet basis or flow-based criterion to decide across time when the noise basis should be updated.

2.4 Synthesis

This chapter presented the principles of the Location Uncertainty formalism, based on a semi-martingale decomposition of the Lagrangian velocity. This framework entails a modification of the usual material derivative towards a stochastic transport operator, preserving the physical conservation laws. The LU version of the Surface Quasi-Geostrophic model was presented and two noise generation procedures were exposed, one describing a data-driven stationary noise, and a second one a flow-driven non-stationary noise.

Chapter 3

Application of data assimilation to stochastic models

This chapter gathers our main findings in the application of ensemble data assimilation methods to the stochastic SQG system. LU is first compared to deterministic inflation procedures. Then, a noise calibration procedure is presented and relies on Girsanov transforms to achieve a guiding role for the ensemble towards a desired region of interest, close to the observations.

3.1 Generation of the observations

This very short subsection describes the numerical and observation setup that will be common to both studies of Chapter 3.

The stochastic and deterministic simulations are run on a simulation grid, G_s, of size 64×64, meaning that each realization X is a 64×64 matrix, or equivalently a vector of size 4096 (because here only the buoyancy field is observed and simulated, the velocity being obtained through the inversion of a fractional Laplacian). The actual physical size of the domain being 1000km $\times$ 1000km, two neighbor grid points are distant of about 15km. An observation Y on a coarser observation grid G_o, of size 16×16, is assumed to be available every day (i.e. every 600 time steps of the dynamics). It is generated as follows:

— A trajectory $(Z_t)_t$ is run from the deterministic model (PDE) at a very fine resolution grid G_f, of size 512×512.

— Then, a convolution-decimation procedure (Lindeberg, 1993) is applied, which is the composition of a Gaussian filter $\mathbf{G}_\sigma$ and a decimation operator $\mathbf{D}$ subsampling one pixel out of two. The Gaussian filtering writes

$$\mathbf{G}_\sigma(Z_t)(x) = (g_\sigma * Z_t)(x),$$

where g_σ is a two-dimensional Gaussian function. For any observable f defined on a grid $(x_{11}, ..., x_{2n,2n})$, we define $\mathbf{D}f$ on a decimated subgrid $(x'_{11}, ..., x'_{n,n})$ by

$$\mathbf{D}f(x'_{ij}) = f(x_{2i,2j}).$$

This convolution-decimation operator $\mathbf{D} \circ \mathbf{G}_\sigma$ is then performed three times in order to fit to the targeted simulation grid.

— Finally, a projection operator $\mathbf{P}$ is applied from G_s to G_o, the latter being a subset of the first, and the observation is defined, for all t, by

$$Y_t = \mathbf{P} \circ (\mathbf{D} \circ \mathbf{G}_\sigma)^3(Z_t) + \eta_t,$$

where

$$\eta_t \sim \mathcal{N}(\mathbf{0}, \mathbf{R}) \quad \text{and} \quad \mathbf{R} = diag(r_1^2, ..., r_{d_Y}^2) \tag{3.1}$$

is the diagonal observation covariance matrix. It will often be considered that $\mathbf{R} = r^2\mathbf{I}_{D_Y}$, where D_Y is the number of points on the observation grid. As a consequence, the operator $\mathbf{P}$ plays the role of the observation operator $\mathbf{H}$ in the Kalman equations.

In practice, in our experiments, r is set to 10^{-5} (1% of the maximum amplitude of the initial buoyancy field) and G_σ is a Gaussian filter applied on 5×5 patches with $\sigma = 2$.

Numerical setup : The simulations are performed with a pseudo-spectral code in space (Resseguier et al., 2020b). The time-scheme is a fourth-order Runge-Kutta scheme for the deterministic PDE, and an Euler-Maruyama scheme for the SPDEs (Kloeden and Platen, 1999). Regardless of the resolution and stochasticity, we use a standard hyperviscosity model to dissipate the energy at the resolution cut-off. The resulting implemented dynamics is :

$$\mathbb{D}_t b = \alpha \Delta^4 b \mathrm{d}t, \tag{3.2}$$

with a hyperviscosity coefficient $\alpha = (5 \times 10^{29} m^8.s^{-1}) D_X^{-8}$, where D_X is the grid resolution (here 512 for the fine-resolution PDE used to generate the observations, and 64 for the ensemble members). The boundary conditions are double-periodic. As mentioned before, the equations are mostly handled in the Fourier space, where the following SQG relation between velocity and buoyancy Fourier transforms can be used:

$$\hat{v} = ik^\perp \frac{\hat{b}}{N_{strat}||k||}, \tag{3.3}$$

with k is the horizontal wave-vector, $k^{\perp}$ the orthogonal horizontal wave-vector and N_{strat} is the stratification.

3.2 Comparison between LU and deterministic variance inflation for EnKF

The aim of this first study was to assess the benefits brought by stochastic dynamics in an up-to-date version of the ensemble Kalman filter with localization and prior multiplicative inflation (details in Section 1.3). In particular, we wish to observe whether or not the stochastic dynamics brings by itself an efficient and practical inflation procedure for ensemble Kalman filtering.

The test case considered in this study is the following: an ensemble of $N = 100$ particles is started from the very same initial condition at day 0, which consists in two cold vortices to the north and two warm vortices to the south (Figure 3.1):

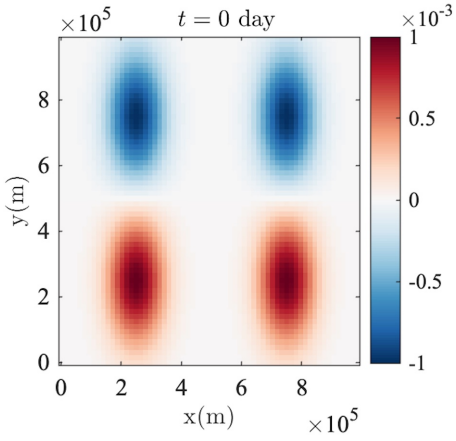


Figure 3.1 – Initial condition of the buoyancy field for all ensemble members.

The initial field is mathematically defined on each grid point $p = \begin{pmatrix} x \\ y \end{pmatrix}$ by the following formula:

$$b_0(p) = F\left(p - p_1\right) - F(p - p_2) - F(p - p_3) - F(p - p_4), \tag{3.4}$$

where the vortices initial cores are

$$p_1 = \begin{pmatrix} 250\text{km} \\ 250\text{km} \end{pmatrix} \quad ; \quad p_2 = \begin{pmatrix} 750\text{km} \\ 250\text{km} \end{pmatrix} \quad ; \quad p_3 = \begin{pmatrix} 250\text{km} \\ 750\text{km} \end{pmatrix} \quad ; \quad p_4 = \begin{pmatrix} 750\text{km} \\ 750\text{km} \end{pmatrix}$$

and the function F is defined by

$$F(r) = B_0 \exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right) \tag{3.5}$$

with $B_0 = 10^{-3} m.s^{-2}$, $\sigma_x = 67$ km and $\sigma_y = 133$ km. We also set the Coriolis frequency to $f_0 = 1.028 \times 10^{-4} s^{-1}$ and the stratification to $N_{strat} = 3f_0$.
In this experiment, we study the differences of efficiency of the localized Ensemble Square Root Filter (cf Section 1.3.5) with inflated deterministic forecast and non inflated stochastic simulations. In both cases, the initial ensemble is generated as follows. Starting from the initial condition, the stochastic dynamics is simulated using the SVD noise for 3 days (meaning 1800 time steps for the SPDE), without data assimilation. This way, a random ensemble is generated, and the performances of the SVD noise indicate that this ensemble is well-spread around the truth, which will be confirmed in Figure 3.6. An observation is provided each day (i.e. every 600 time steps of the SPDE (or PDE)), with an observation error covariance set to $r_i = 10^{-5}$ for $i = 1, ..., D_Y$ in equation (3.1), which corresponds to a weak (but not negligible, 1% of the maximum amplitude of the initial buoyancy field) noise on the observation. For the rest of the simulation (100 days), this ensemble is used for two experiments:

— **Experiment 1 (deterministic dynamics with inflation):** The stochasticity is shut down after day 3, the forecast is then generated by the deterministic PDE, and prior multiplicative inflation is used to artificially increase the ensemble spread. Namely, before the assimilation step, given a forecast ensemble X^f, the inflated ensemble $\tilde{X}^f$ is defined, for all n, by

$$\tilde{X}^{(n),f} = \overline{X}^f + \alpha(X^{(n),f} - \overline{X}^f) \tag{3.6}$$

where $\alpha > 1$ is the inflation parameter. Then the localized ESRF is applied to $\tilde{X}^f$ and the same procedure is done each day at each new observation time.

— **Experiment 2 (stochastic dynamics without inflation):** The same SPDE drives the particles for the whole simulation, and the localized ESRF is applied with an observation each day.

The localization radius was set to l_{obs} here, where $l_{obs} \simeq 60km$ denotes the distance between two neighboring observational sites, as it provided the best results for both cases. The metric used here is the mean square error (MSE) on the buoyancy fields, defined for

any time t by $MSE(t) = \frac{1}{N}\sum_{i=1}^{N} ||b^{truth}(\cdot,t) - b^{(i)}(\cdot,t)||_2^2$, where b^{truth} is the 512×512 SQG reference, adequately filtered and decimated. The inflation procedure is very sensitive to the parameter α. It must be finely tuned to have the best results. If it is too small, then the spread is not large enough. If it is too large, then it could entail a divergence of the filter (cf figure 3.2). For our SQG configuration, it turns out that the range of validity of

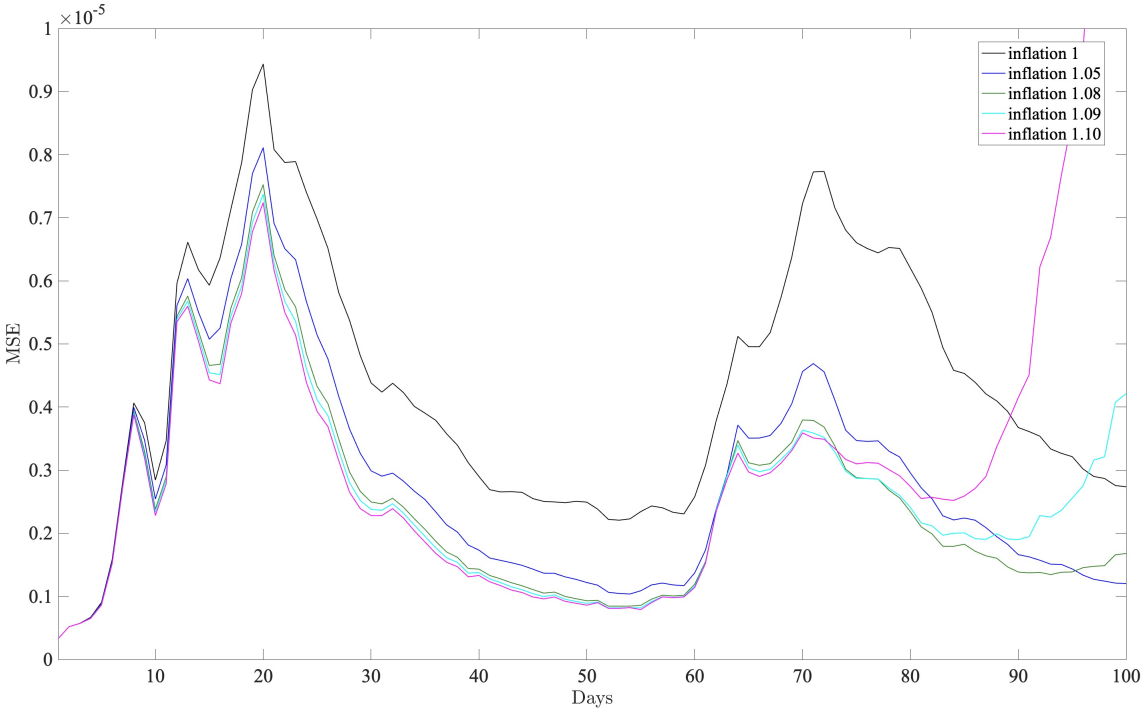


Figure 3.2 – Buoyancy Mean Square Error curves for different values of the inflation parameter α in the deterministic case (case 1): magenta ($\alpha = 1.10$), cyan ($\alpha = 1.09$), green ($\alpha = 1.08$), blue ($\alpha = 1.05$) and black (no inflation: $\alpha = 1.0$). Inflation is overall very beneficial compared to the black line (no inflation), but too big values of α lead to a long-term divergence of the filter (magenta and cyan).

α is between 1 and 1.08 approximately. Starting from $\alpha = 1.09$, the filter starts diverging in the long term. In the context of our model, this small range for the inflation parameter and the tuning procedure is in itself a drawback of the inflation method. Still this tuning can be very different depending on the model at hand, so we do not claim that this range is small in general. Moreover, when α is badly chosen, it brings in the additional problem of strong divergence of the ensemble members for the localized ESRF, as shown in Figure 3.3. In this example ($r_{loc} = 3l_{obs}$ and $\alpha = 1.15$), starting from day 50 or so, strong gradients between neighboring grid points can be observed, and progressively lead to non

physical predictions. Obviously the same kind of behaviour also occurs in our setting with a smaller value of r_{loc}. This was theoretically expected as, in the localization procedure, the posterior ensemble members are combinations by block of linear combinations of the prior ensemble members (cf equation (1.29) in Section 1.3.4). This formula relies on the strong assumption that such a block recombination remains a solution of the underlying PDE (or SPDE), which suggests that the global equation is equivalent to a combination of local ones. Let us note that for instance for fluid dynamics equations, the presence of a pressure term, solution of an elliptic problem, is intrinsically global and theoretically prevents such a local modeling. A careless application of localization may lead to the appearance of some discontinuities or gradients due to very different decisions taken by the filter at neighboring points. This may entail, after subsequent iterations of the underlying dynamics, major errors, numerical divergence of some ensemble members, leading eventually to unphysical realizations.

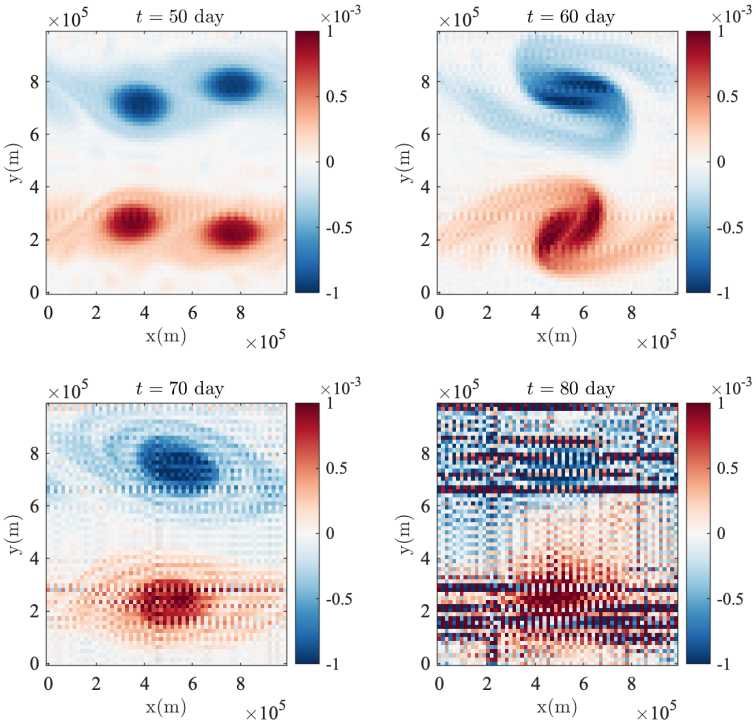


Figure 3.3 – Examples of diverging realizations of buoyancy resulting from LESRF with $r_{loc} = 3l_{obs} \simeq 180km$ and $\alpha = 1.15$.

The long-term instability of inflation is not a good sign, especially considering we deal here with a rather coarse-scale diffusive simulation. For finer resolutions with less diffusion

and much more pronounced non-linear features, divergence is likely to occur sooner for the same inflation value. This instability is theoretically expected, as the artificial variance injection entailed by inflation is never counterbalanced in any way in the model. At a coarse resolution, even for low inflation parameters, the typical slow growth of error at the end of the simulation seems to indicate that long-time divergence, as in figures 3.4 and 3.3, is likely to be observed extending the simulation time (Figure 3.5 equivalently shows that increasing the inflation parameter makes the divergence occur earlier). The divergence caused by variance inflation can be attenuated by temporal adaptive schemes with the introduction of diagnostic criteria (Lee et al., 2017; Raanes et al., 2019). Another alternative consists in changing the hypothesis on the prior distribution accounting for the sampling errors in the ensemble, which was shown in Bocquet (2011) to make multiplicative inflation optional on low dimensional Lorenz models.

An increase of variance without controlling the global energy by a balanced dissipation raises question on the mathematical well-posedness of the numerical scheme but also on the modified physics undergone by this artificial forcing. These two questions are far from cosmetic. The first one is related to the generality of the numerical scheme (i.e. it must be valid for any noise and at any resolution, etc.). The second point is about the error terms introduced, they should not change dramatically the targeted underlying physical system (Chapron et al., 2018). These two points can hardly be met by an artificial increase of variance or by non physical multiplicative noise (see Chapron et al. (2018) for an example on the Lorenz63 model, in which an artificial multiplicative noise of low amplitude was shown to change radically the statistics of the underlying system).

On the opposite, the LU setting brings a natural balance between the energy brought by the noise and the amount dissipated by the stochastic diffusion. In addition, as shown in Figure 3.4, the simulation is stable while bringing an equivalent spread as a relatively strongly inflated deterministic simulation value (with respect to the SQG dynamics studied here). It leads also to better MSE results than the deterministic setting for the different values of inflation experimented here.

Figure 3.5 shows the same comparison with a greater inflation parameter set to $\alpha = 1.20$. The time window is reduced to 50 days because, as expected, increasing the inflation parameter leads to an earlier divergence, here starting from days 40-50. Still it has comparable results with LU in the first 30-35 days. This means that LU has short-term MSE performances comparable with a very strong deterministic inflation parameter, and avoids long-term divergence as well.

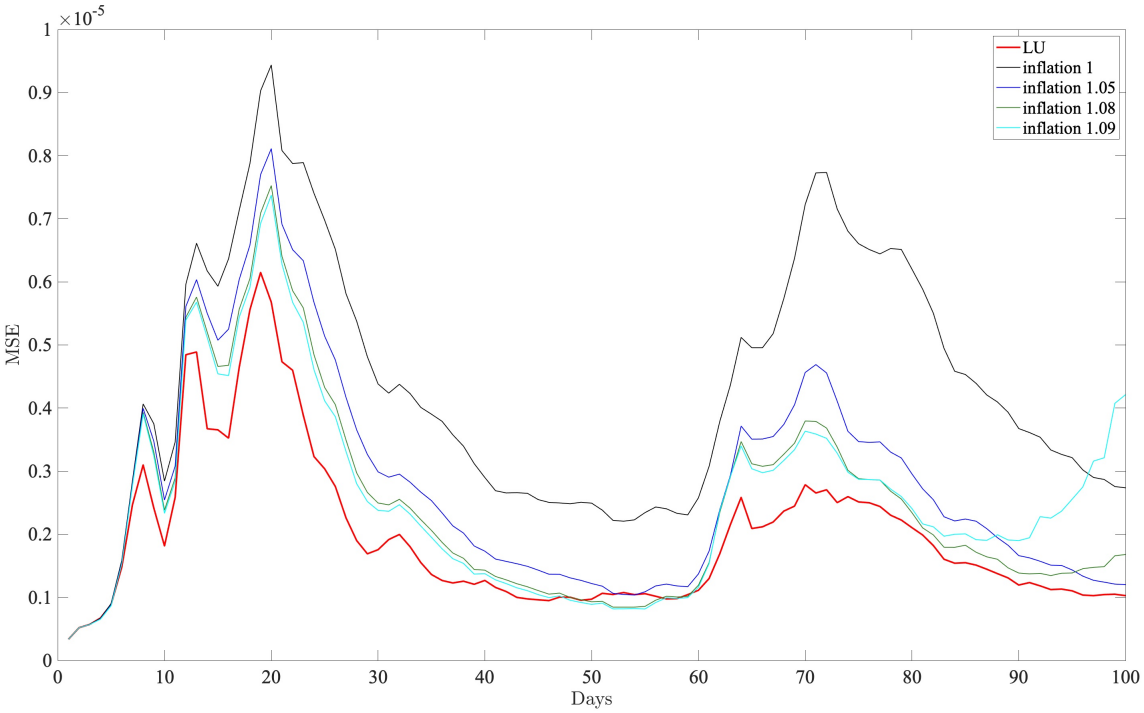


Figure 3.4 – The LU framework (in red) performs better, in terms of buoyancy MSE, than the deterministic cases for all the reasonable inflation parameters tested and plotted in cyan ($\alpha = 1.09$), green ($\alpha = 1.08$), blue ($\alpha = 1.05$) and black (no inflation: $\alpha = 1.0$) color.

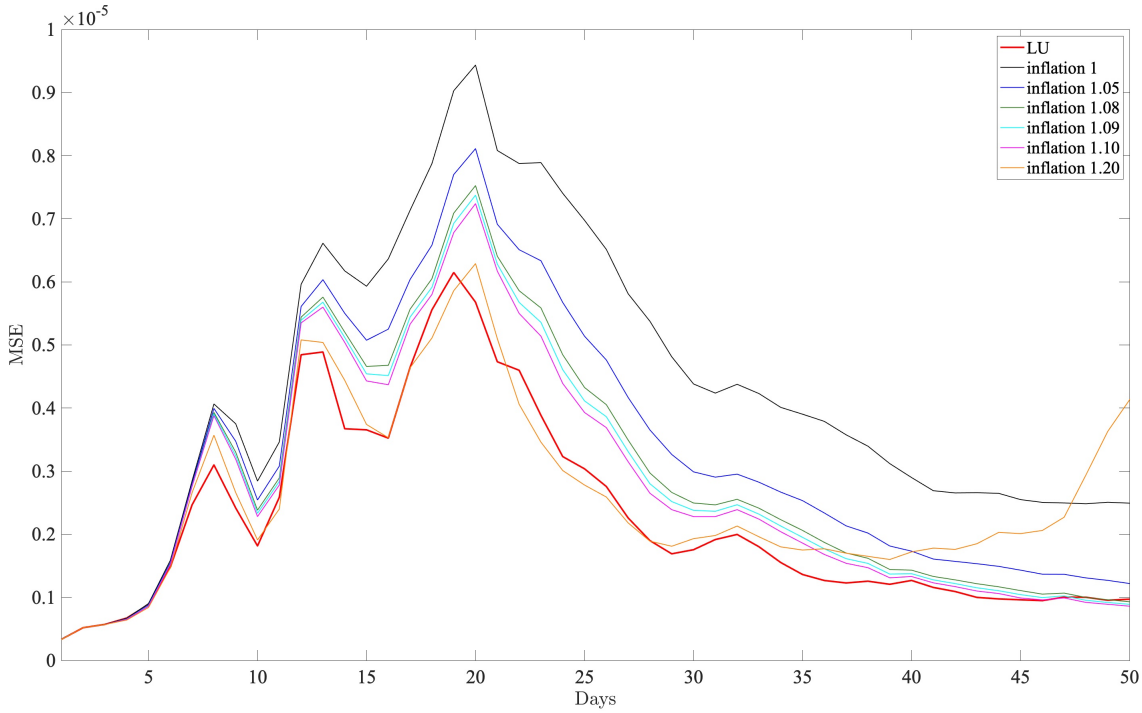


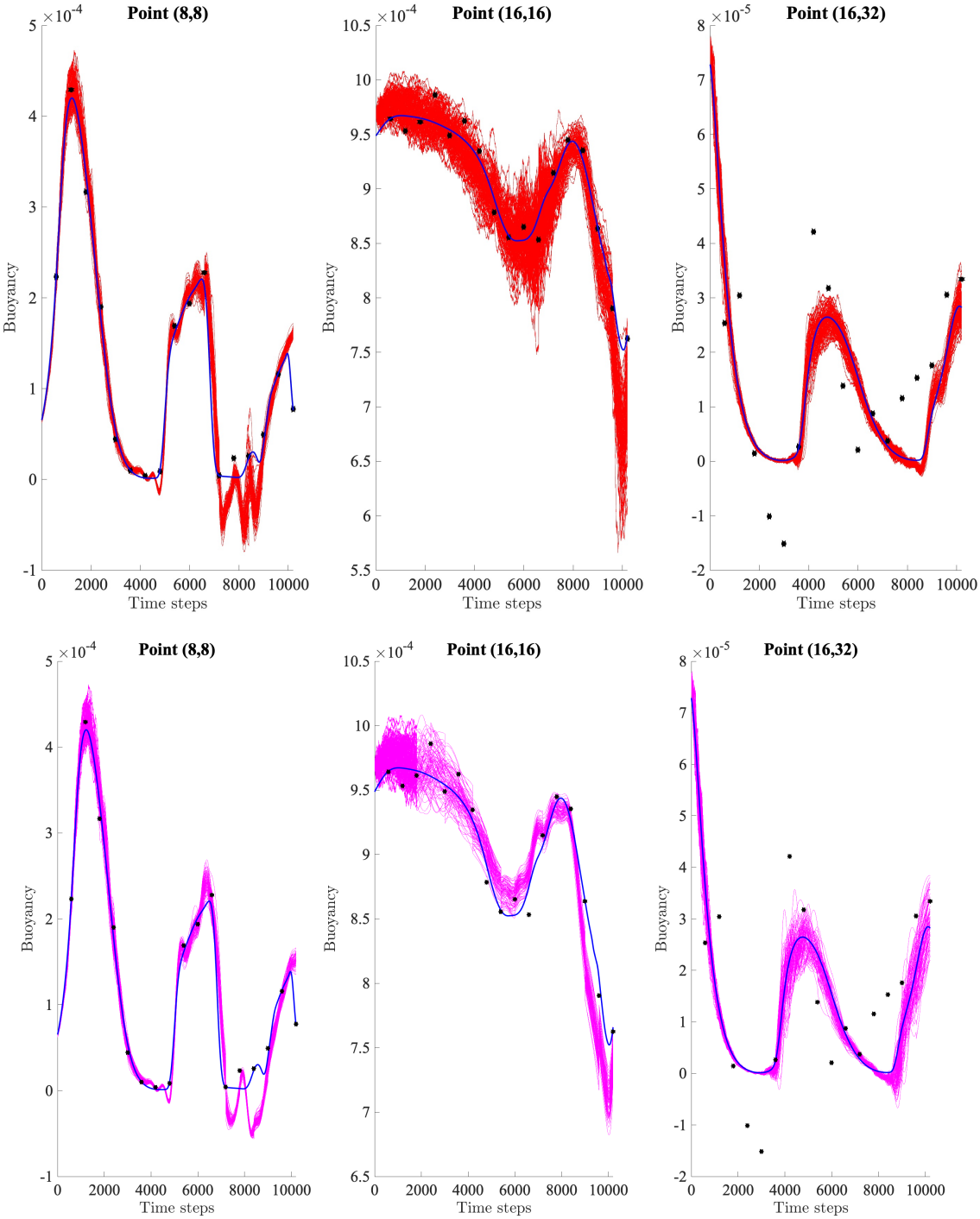
Figure 3.5 – Same figure as 3.4 with additional inflation parameters $\alpha = 1.10$ (in magenta) and $\alpha = 1.20$ (in orange). The time window is reduced to 50 days as the orange curve diverges much sooner than the others.

Additionally, we examine the spread of the ensemble members around the truth and observation points. We chose three characteristic grid points, corresponding to the center, north and southwest of the southwest warm vortex of the initial condition (Figure 3.1). In the following figure (Figure 3.6), we compare the behavior of the spread of LU ensemble members (in red) and deterministically inflated ones with parameter $\alpha = 1.10$ (in magenta) and $\alpha = 1.20$ (in orange) around the truth (blue dots) and observation points (black dots). Note that the two chosen inflation parameters are quite strong and lead both to divergence of the filter, with the second one exhibiting a divergence much sooner than the first one.

Figure 3.6 shows this comparison for the first 17 days of simulation. Although both spreads of subfigures (a) and (b) seem very similar for the figures of left and right columns, we observe on the center column figure that deterministic inflation does not provide enough spread to contain the truth, while LU does (see for instance time window 8,000-10,000). It can also be noticed that the truth at time step 8400 is completely skipped by magenta and orange spreads, while LU manages to reach it.
We see that increasing the inflation parameter does not counteract the flaws of the smaller inflation parameter $\alpha = 1.10$ (magenta spread, Figure 3.6), meaning that increasing α does not entail a richer ensemble. To that extent, we can also notice that when the stochasticity is shut down at day 3, the deterministic sets of trajectories immediately become less dense than the stochastic one. This means that the neighborhood of the truth/observation is visited much more often (in time) by stochastic trajectories. This improved "density" of the stochastic ensemble in the state space is an interesting feature, not exploited yet but it could be relevant in a particle filter framework.

Deterministic inflation sometimes offers more spread than LU. We observe this situation for example at the time window 9,000-10,000 on the right column figure. But this spread seems bigger than necessary, as LU has a smaller spread, but already well-centered around the truth. This supports the idea that deterministic inflation behaves blindly compared to LU, in the sense that it is partly decorrelated from the dynamics and the current state of the ensemble. LU does not only provide a spread of the ensemble equivalent to a strong inflation parameter, it brings an ensemble of better quality. This ensemble is constructed from the large-scale fluctuations in a way to fit the physics encoded by the original dynamics.

Another indicator of this is the following spread-error consistency shown in Figure 3.7. We compared the ensemble bias absolute value $e(x,t) = |\bar{b}(x,t) - b^{truth}(x,t)|$ with the


Point (8,8)
Point (16,16)
Point (16,32)
Buoyancy
Time steps
×10^-4
×10^-5

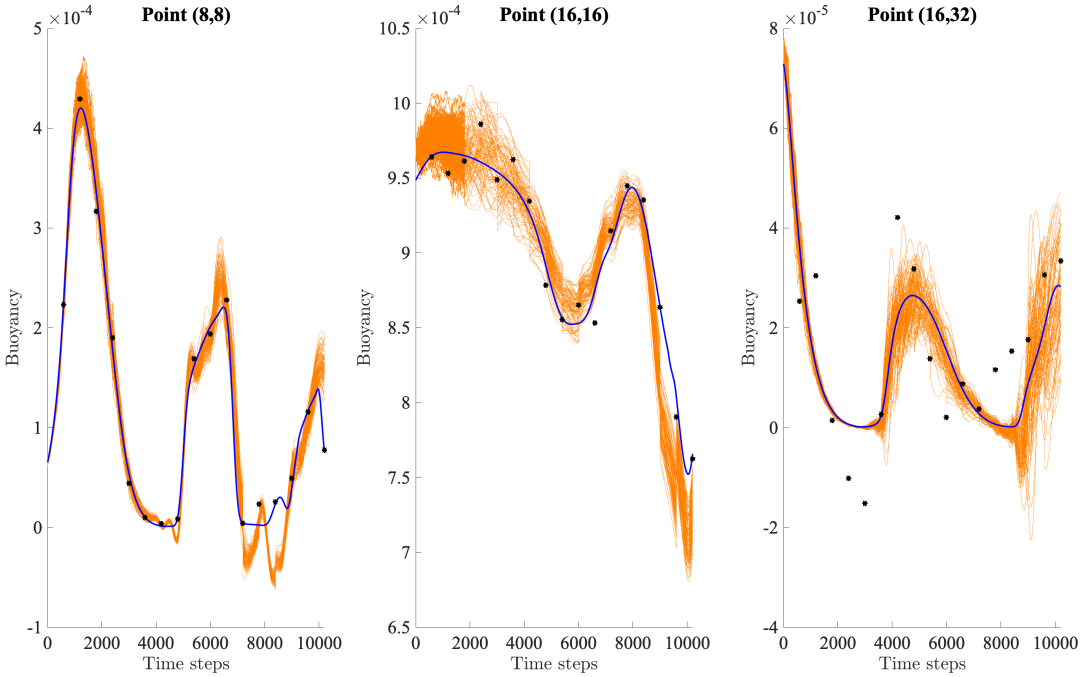
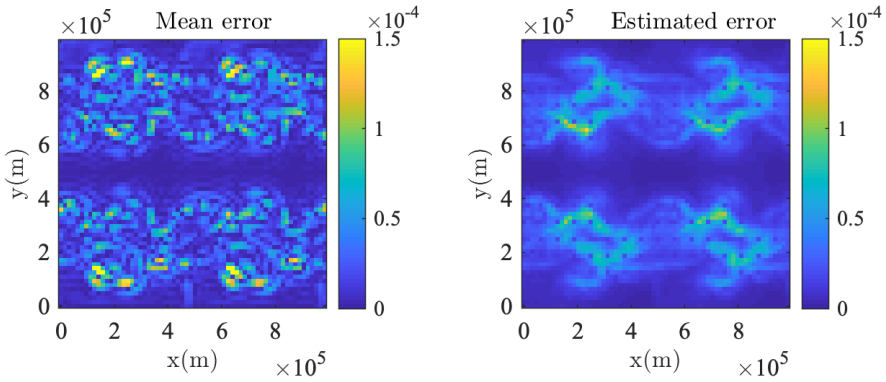
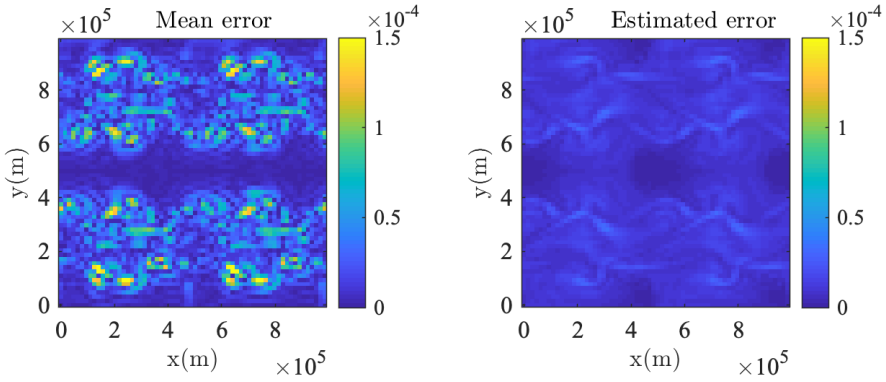


Figure 3.6 – Comparison of spread between LU and deterministic inflation on the first 17 days of the dynamics for three points located at the center, top and bottom left of the bottom left warm vortex of the initial condition. The upper row shows in red the buoyancy values at these points for the stochastic ensemble. The two lower ones show the buoyancy values at the same points for the deterministically inflated ensembles for $\alpha = 1.10$ (in magenta) and $\alpha = 1.20$ (in orange). The black dots are the observations; the blue line stand for the truth.

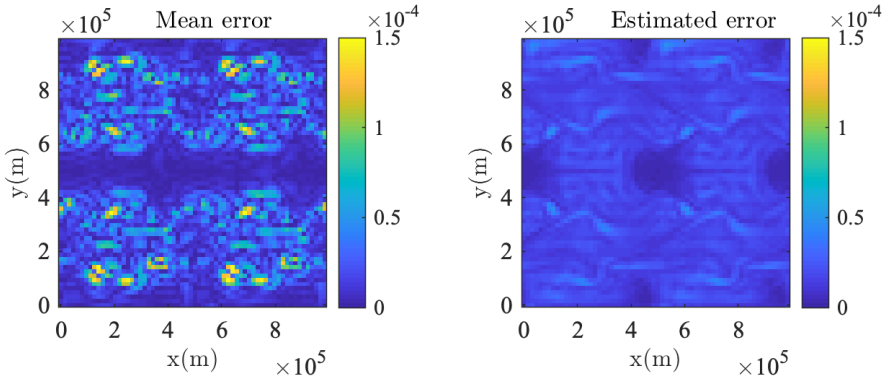
error estimated from the ensemble spread $\varepsilon(x,t) = 1.96\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(b^{(i)}(x,t) - \overline{b}(x,t))^2}$. These two quantities do not necessarily have the same amplitude, but should have the same spatial structure if the ensemble variability is relevant. We did this for LU and for the deterministic setting with inflation parameters $\alpha = 1.10$ and $\alpha = 1.20$ at day 17, when the three models are very close in terms of global MSE (see Figure 3.5). The results clearly point out that the LU framework provides much more spread and a better estimation of the error. The spread induced by inflation has similar structures around the vortices, even if its amplitude is much lower, confirming the insights of Figure 3.6. However, inflation puts non zero spread in areas that are irrelevant, typically at ordinates $0, 5$ and 10 ($\times 10^5$), whereas LU only focuses the noise around the vortices.



(a) LU



(b) Deterministic inflation $\alpha = 1.10$



(c) Deterministic inflation $\alpha = 1.20$

Figure 3.7 – Comparison between the ensemble bias absolute value $e(x) = |\bar{b}(x) - b^{truth}(x)|$ (left maps) compared to the estimate error (1.96× the standard deviation of the ensemble) evaluated at each grid point (right maps) at day 17. The upper row shows this comparison for LU, the other two show the same comparison with the deterministic setting respectively for inflation parameters of $\alpha = 1.10$ (central row) and $\alpha = 1.20$ (bottom row).

3.3 Noise calibration for ensemble data assimilation

As mentioned in the beginning of Chapter 1, data assimilation filters are composed of a forecasting step of the ensemble to provide a sampling of the forecast distribution, and an analysis step correcting the departure from the observations. Until now, observations essentially took part in the analysis step when it comes to correct the predictions of the numerical model. This section suggests a methodology that actually incorporates in some way the observations into the forecasting step.

3.3.1 Change of measure

The purpose of the proposed noise calibration is to modify the forecast distribution, taking into account the upcoming observation, in order to guide the forecast towards it. In the context of transport equations such as in the SQG model, this extra guiding term is an added drift in the noise $\sigma \mathrm{d}B_t$, which was initially built to have zero mean. Allowing $\sigma \mathrm{d}B_t$ to have a non-zero mean entails a modification of the transport equation in order to rewrite it in terms of a centered noise. This is called the Girsanov transform, and it consists in a change of underlying measure so that a non-centered noise becomes centered under a new probability measure, up to a drift term accounting for this change of measure. For now, $\sigma \mathrm{d}B_t$ is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and we define $(\mathcal{F}_t)_t$ the filtration adapted to $\sigma \mathrm{d}B_t$.

Theorem 4 (Girsanov theorem)**.** *Assume that $(Y_t)_{0\leq t\leq T}$ is a stochastic process such that:*

— *$(Y_t)_{0\leq t\leq T}$ is adapted to the Wiener filtration $(\mathcal{F}_t)_{0\leq t\leq T}$.*

— *For the current probability measure $\mathbb{P}$, we have, $\mathbb{P}$-almost surely,*

$$\int_0^T ||Y_t||^2 \mathrm{d}t < \infty.$$

— *The process $(Z_t)_{0\leq t\leq T}$ defined by*

$$Z_t = \exp\left(\int_0^t Y_s \mathrm{d}B_s - \frac{1}{2}\int_0^t ||Y_s||^2 \mathrm{d}s\right) \tag{3.7}$$

is a $\mathcal{F}_t$-martingale.

Then, there exists a probability measure $\tilde{\mathbb{P}}$ under which:

— *The process* $(\tilde{B}_t)_{0\leq t\leq T}$ *defined by*

$$\tilde{B}_t = B_t - \int_0^t Y_s \mathrm{d}s \tag{3.8}$$

is a standard cylindrical Wiener process.

— *The Radon-Nikodym derivative of* $\tilde{\mathbb{P}}$ *with respect to* $\mathbb{P}$ *is* Z_T.

We refer to Prato and Zabczyk (1992) for details on this theorem. The assumption of the martingale property of Z can be substituted by a sufficient condition, called Novikov's condition, stating that having

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T ||Y_t||^2 \mathrm{d}t\right)\right] < \infty \tag{3.9}$$

is enough to ensure that $(Z_t)_{0\leq t\leq T}$ is a $\mathcal{F}_t$-martingale. We also refer to Delyon and Hu (2006) for a version of the theorem for locally bounded drifts.

Let us denote by $(\Gamma_t)_{0\leq t\leq T}$ the drift we intend to add to the noise. With such a change of measure, let us see how equation (2.8) is modified. According to equation (3.8), we have

$$\mathrm{d}B_t = \mathrm{d}\tilde{B}_t + \Gamma_t \mathrm{d}t, \tag{3.10}$$

so the stochastic transport operator rewrites

$$\mathbb{D}_t b = \mathrm{d}_t b + (v^* \mathrm{d}t + \sigma[\mathrm{d}\tilde{B}_t + \Gamma_t \mathrm{d}t]) \cdot \nabla b - \frac{1}{2}\nabla \cdot (a \nabla b)\mathrm{d}t \tag{3.11a}$$

$$= \mathrm{d}_t b + (v^* \mathrm{d}t + v_\Gamma \mathrm{d}t + \sigma \mathrm{d}\tilde{B}_t) \cdot \nabla b - \frac{1}{2}\nabla \cdot (a \nabla b)\mathrm{d}t, \tag{3.11b}$$

where

$$v_\Gamma = \sum_{k=1}^{K} \gamma_k \phi_k \tag{3.12}$$

is the velocity drift entailed by the Girsanov transform and we assume that $\Gamma_t = \Gamma = (\gamma_1, ..., \gamma_K)$ is constant on a small time step $\mathrm{d}t$, which will be the case for the discretized numerical scheme that we use.

As a result, under the probability measure $\tilde{\mathbb{P}}$, (3.11) presents the same form as equation (2.8) since $\tilde{B}$ is indeed a centered cylindrical Wiener process under $\tilde{\mathbb{P}}$, but with an added drifted advection velocity.

3.3.2 Computation of the Girsanov drift

We now describe how to compute Γ in order to guide the forecast towards the next observation.
Let us start from a given time t_1 where a complete buoyancy and velocity field is available. The next observation $b^{obs}(\cdot, t_2)$ is available at time t_2 and L numerical time steps are performed until then ($t_2 - t_1 = L\delta_t$, where δ_t is the time discretization step).
At time t_1, a rough prediction of the buoyancy at time t_2 can be estimated with the current velocity (which, more precisely, comes from previous stochastic iterations, but is $\mathcal{F}_{t_1}$-measurable), namely

$$b^{obs}\left(x + v(x, t_1)L\delta_t, t_2\right) := \tilde{b}(x, t_2), \tag{3.13}$$

that stands for the backward-registered observation with respect to the current deterministic velocity. This way the error made is

$$\Delta_t\tilde{b}(x) = \tilde{b}(x, t_2) - b(x, t_1). \tag{3.14}$$

So $\tilde{b}(x, t_2)$ is a value taken in a modified observation field, because b^{obs} is advected by the current velocity $v(\cdot, t_1)$. For this reason we consider that the backward-registered obervation used for the calibration does not have the same nature as the raw observation used for data assimilation. It constitutes a pseudo-observation, for which we can consider that the error due to the imprecision of the backward-registration (ensuing in particular from successive bilinear interpolations) is way bigger than the observation noise, and almost uncorrelated to the latter. In the second case, only the raw observation is used for the Kalman filter, corresponding only to the observation noise. The aim is now to calibrate the current velocity by adding a Girsanov drift $v_\Gamma = \sum_{k=1}^{K} \gamma_k\phi_k$, such that the solution of the following transport equation

$$b\left(x + v(x, t_1)L\delta_t + v_\Gamma L\delta_t + \sum_{k=1}^{K}(\sqrt{\delta_t}\phi_k)(\sqrt{L\delta_t}\beta_k), t_2\right) = b(x, t_1). \tag{3.15}$$

is approximated in a least square sense. In other words, we solve the following minimization problem:

$$\min_\Gamma \int_\Omega \mathbb{E}\left[b^{obs}\left(x + v(x, t_1)L\delta_t + v_\Gamma L\delta_t + \sum_{k=1}^{K}(\sqrt{\delta_t}\phi_k)(\sqrt{L\delta_t}\beta_k), t_2\right) - b(x, t_1)\right]^2 \mathrm{d}x. \tag{3.16}$$

This can be rewritten as

$$\min_{\Gamma} \int_{\Omega} \left[\Delta_t \tilde{b} + \nabla \tilde{b} \cdot v_\Gamma L\delta_t - \frac{1}{2} \nabla \tilde{b} \cdot \nabla a L\delta_t - \frac{1}{2} \nabla \cdot (a \nabla \tilde{b}) L\delta_t \right]^2 \mathrm{d}x.$$

Using the identities

$$\nabla \cdot a = \sum_{k=1}^{K} (\phi_k \cdot \nabla)\phi_k \quad ; \quad \nabla \cdot (a \nabla b) = \sum_{k=1}^{K} (\phi_k \cdot \nabla)(\phi_k \cdot \nabla b), \tag{3.17}$$

we rewrite the minimization problem as

$$\min_{\Gamma} \int_{\Omega} \left[\Delta_t \tilde{b} + \nabla \tilde{b} \cdot \left(\sum_{k=1}^{K} \gamma_k \phi_k \right) L\delta_t - \frac{1}{2} \sum_{k=1}^{K} (\nabla \tilde{b} \cdot F_k + G_k(\tilde{b})) L\delta_t \right]^2 \mathrm{d}x \tag{3.18}$$

where

$$F_k = (\phi_k \cdot \nabla)\phi_k \quad ; \quad G_k(\tilde{b}) = (\phi_k \cdot \nabla)(\phi_k \cdot \nabla \tilde{b}).$$

Denoting by J the integrand, we have

$$\frac{\partial J}{\partial \gamma_i} = 2 \int_{\Omega} (\nabla \tilde{b} \cdot \phi_i) L\delta_t \left[\Delta_t \tilde{b} + \nabla \tilde{b} \cdot \left(\sum_{k=1}^{K} \gamma_k \phi_k \right) L\delta_t - \frac{1}{2} \sum_{k=1}^{K} (\nabla \tilde{b} \cdot F_k + G_k(\tilde{b})) L\delta_t \right] \mathrm{d}x. \tag{3.19}$$

Finally, we add a regularization term $\alpha ||v_\Gamma||_2^2 = \alpha \sum_{k=1}^{K} \gamma_k^2 \lambda_k$, where λ_k is the eigenvalue of the Q-eigenfunction ϕ_k in equation (3.18) to ensure the uniqueness of the solution of the proposed minimization problem, where α needs to be tuned properly.
As a result, the minimization problem can be written as an inverse problem

$$\mathbf{A}\Gamma = c \tag{3.20}$$

where

$$\mathbf{A}_{ik} := 2 \int_{\Omega} (\nabla \tilde{b} \cdot \phi_i)(\nabla \tilde{b} \cdot \phi_k) + 2\alpha \lambda_k \delta_{ik} \tag{3.21a}$$

$$c_i := \int_{\Omega} (\nabla \tilde{b} \cdot \phi_i) \left[2\Delta_t \tilde{b} - \sum_{k=1}^{K} (\nabla \tilde{b} \cdot F_k + G_k(\tilde{b})) \right] \mathrm{d}x. \tag{3.21b}$$

The parameter α is a priori fixed in order to control the resulting euclidian norm of v_Γ, $||v_\Gamma||_2$. Large values of α lead to very small corrections (Γ tends to $(0, ..., 0)$ when α goes to $+\infty$) whereas small values yield very strong and noisy drifts, as we get closer to an

ill-posed problem. For now, we use an empirical iterative way to tune α, we increase it until the resulting norm of v_Γ is under a given threshold

3.3.3 Multi-resolution procedure

In the previous procedure, the Girsanov drift is computed only by solving the regularized inverse problem (3.20) at the full resolution of the system. In this section, we explain an alternative multi-resolution framework to compute v_Γ. Let ρ_s and ρ_c respecively stand for the simulation grid resolution and a coarser resolution fixed.
For any resolution $\rho_c \leq \rho \leq \rho_s$, we assume to have at disposal an observation $b_\rho^{obs}(\cdot, t_2)$ (obtained from low-pass filtering and decimation as explained in section 3.1) and subsampled velocity fields for all ensemble members $v_\rho^{(n)}(\cdot, t_1), n = 1, ..., N$ matching this resolution. The POD procedure is also applied at each resolution to get the adequate noise modes $(\phi_k^\rho)_{k=1,...,K}$. Let us now detail the multi-resolution computation of v_Γ.

For each resolution ρ, we compute the associated backward registered observation with respect to the subsampled velocity field at resolution ρ:

$$b_\rho^{obs}\left(x + v_\rho(x, t_1)L\delta t, t_2\right) := \tilde{b}_\rho(x, t_2), \tag{3.22}$$

similarly to what is done in equation (3.13). We define accordingly

$$\Delta_t \tilde{b}_\rho(x) = \tilde{b}_\rho(x, t_2) - b_\rho(x, t_1). \tag{3.23}$$

At this resolution, the aim is to calibrate the current velocity by a Girsanov drift term

$$v_\Gamma^\rho = \sum_{k=1}^{K} \gamma_k^\rho \phi_k^\rho, \tag{3.24}$$

which is decomposed on the adequate noise modes for resolution ρ. This Girsanov drift is solution of the equivalent system of equations (3.15-3.20), substituting the full-resolution fields by the subsampled ones at resolution ρ.
Once the minimization problem is solved, the goal is now to export the information brought by v_Γ^ρ towards the finer resolution $\rho + 1$. In order to do that, the velocity field at

resolution $\rho+1$ should be incrementally modified accordingly by

$$\tilde{v}_{\rho+1} = v_{\rho+1} + \mathbf{E}(v_\Gamma^\rho), \tag{3.25}$$

where $\mathbf{E}$ is an extension operator from the coarser to the finer grid (see next subsection for the description of $\mathbf{E}$ in our test case). By doing so, we enrich the structure of the velocity field at resolution $\rho+1$ with a corrected principal component inherited from a noise calibration at a coarser scale. This can be seen as a Gauss-Newton incremental resolution of the initial minimization problem (3.18). This optimization is akin to classical multi-resolution incremental fluid motion estimation (Cai et al. (2018)).

Then, equations (3.21-3.24) are iterated until the simulation grid resolution is reached.

3.3.4 Numerical experiments

The goal is to study the benefits brought by a noise-calibrated forecast in an up-to-date version of a localized ensemble Kalman filter. In particular we wish to observe whether or not the noise calibration brings by itself an efficient and practical improvement of the assimilation step.

One important assumption of the classical EnKF is to consider that the observation and model noise are uncorrelated. This observation-calibrated forecast could imply that the latter assumption no longer holds. Still, the discussion following equation (3.14) on the observation nature explains why we can consider the uncorrelation between the forecast and observation noise. If this assumption appears to be not valid, we refer to the work made in Arnaud et al. (2005) to rigorously justify the introduction of an observation-dependent forecast. In this work, both Kalman and particle filter equations were rewritten to be slightly generalized in terms of the conditional expectation with respect to the underlying sequence of current and past observations.

We refer to section 3.1 for the generation of the observations and the numerical setup. Just note that, in the single-resolution framework of subsection 3.3.2, we have $G_o = G_s$ and $\mathbf{P} = \mathbf{Id}$. In the multi-resolution one, observations Y_ρ and subsample velocity fields are obtained by applying a decimation operator $\mathbf{D}_\rho$ from G_s to G_ρ, similarly to what is done by the convolution-decimation operator $\mathbf{D} \circ \mathbf{G}_\sigma$. The extension operator $\mathbf{E} : G_\rho \to G_{\rho+1}$ used in equation (3.25) is a duplication operator that is in some way the inverse of the decimation operator, copying the values of every point on the coarse grid to its 4 corresponding pixels on the finer grid.

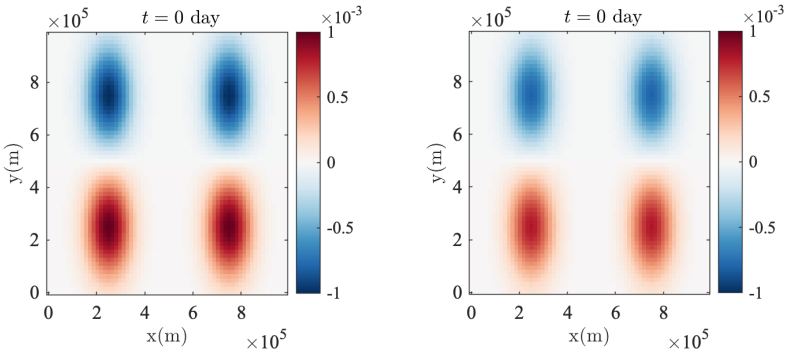


Figure 3.8 – Initial conditions of buoyancy for the truth (on the left) and for each stochastic run (on the right, common to all ensemble members). We enforce an underestimation of the amplitude of the initial vortices of 20%.

The test case considered in this study is the following: an ensemble of $N = 100$ ensemble members is started from the very same initial condition at day 0, which consists in two cold vortices to the north and two warm vortices to the south. However, and very importantly, the amplitude of the initial vortices was decreased compared to the initial condition used for the deterministic run (considered as the truth) by 20%, as shown in Figure 3.8. We imposed thus a strong bias in the initial condition compared to the true one. We refer to Section 3.2 for a mathematical expression of this field.

In this experiment, we study the differences of efficiency of the localized Ensemble Square Root Filter with both noise-calibrated forecast and classical stochastic simulations. We also refer to Raanes et al. (2015) for the extension of the square root filter for additive forecast noise based on covariance transformation, where the advantages of additional model error in the forecast step are shown.

In both cases, starting from the underestimated initial condition, the stochastic dynamics is simulated using the POD noise with $K = 10$ modes. An observation is provided each day (i.e. every 600 time steps of the SPDE), with an observation error covariance set to $r = 10^{-5}$ in (3.1), which corresponds to a weak (but not negligible, 1% of the maximum amplitude in the initial buoyancy field) noise on the observation. The localization radius is set to l_{obs} here, where $l_{obs} \simeq 60km$ denotes the distance between two neighboring observational sites, as it provided the best results for both cases.

In our configuration, the typical behavior of the vortices, at least at the beginning of the simulation, is to spin with no translation of the cores. With the bias we imposed on the initial condition, the true (observed) vortices will spin much faster than those in

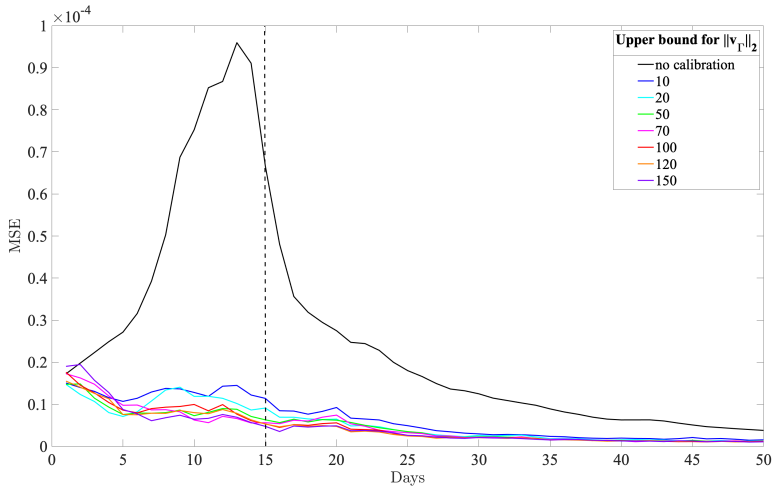


Figure 3.9 – Comparison of buoyancy MSE along time between the non calibrated forecast (in black) and all the different values of the regularization parameter tested here for the noise calibration. The snapshots shown in Figure 3.11 are taken at day 15 (black dashed line).

the biased stochastic runs. The goal of the calibration is then to speed these (simulated) vortices up in order to get them closer to the truth. The forecast is calibrated at a single resolution (Section 3.3.2) at each time step of the SPDE, using the upcoming observation to do it. Multiple values were tried for the regularization parameter α, or alternatively for the upper bound allowed for the L^2-norm of the Girsanov drift v_Γ. Figure 3.9 compares the MSE along time for all the range of values tested here, with also the same experiment without noise calibration. For this latter, the LESRF faces a very difficult task, as it tries to find linear combinations of the prior ensemble members, which all have an underestimated velocity, to get closer to the observation. This is a general issue for ensemble methods (as well as for particle filters), which are not able, by design, to correct the bias if this correction is not made in the forecast. By contrast, the LU calibration offers an additional degree of freedom to guide the ensemble towards the observation. This procedure significantly improves the results in terms of MSE. At day 13, when the MSE is maximal for the usual case, we observe an improvement from 85% to 93% depending on the values tested for the regularization parameter. In Figure 3.10, we provide the MSE results of the multi-resolution procedure with a two-resolution setting. We compare it to the best result obtained with the single resolution procedure, and we doubled the number of ensemble members in the latter so that the computational costs match. At day 1, when the MSE is maximal for both cases, the multi-resolution scheme provides an additional

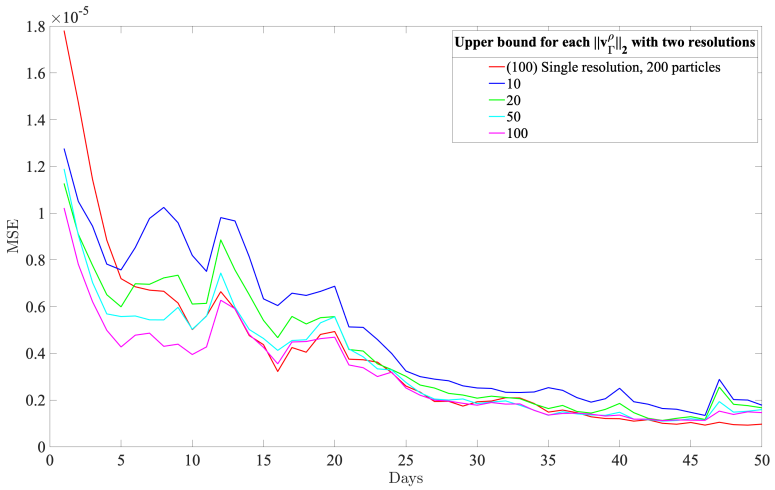


Figure 3.10 – Comparison of buoyancy MSE along time between the best result of single-resolution calibration forecast (in red) and all the different values of the regularization parameter tested here for the multi-resolution noise calibration (with two resolutions). The single-resolution calibration is done with twice as many ensemble members so that all experiments have equivalent computational costs.

improvement of the MSE results from 28% to 45% depending on the values tested.

The case of the underestimation is an example, but we expect this procedure to be efficient in any situation in which all ensemble members have a similar problem of bias, bad amplitude estimation, artefacts, unsymmetrical features, etc. With a reasonably small ensemble size, which is generally the case in practice, this is likely to occur if the initial conditions have such features.

As explained previously, the regularization term α controls the amplitude of the allowed correction drift. In our experiments, all values tested yield significant improvements compared to the classical case, still a good trade-off seems to be found with a control of $||v_\Gamma||_2$ between 70 and 150. Starting from 150, we observe higher MSE in the very first days, certainly due to a lack of constraint on the inverse problem. In addition to the MSE results, we show in Figure 3.11 a more visual example of what calibration does. At day 15, the configuration of the truth is that all four vortices are along the x-axis. Without calibration (first row), the vortices are slanted because of the initial underestimation of the velocity. The velocity field has not been properly corrected. On the other hand, the LU calibration offers a more reliable prediction, as we recovered the global shape of the vortices, with additional spread around the mean. Similarly to Figure 3.11, we also show in Figure 3.12 the comparison between single and multi-resolution procedures at day 1.

The two-resolution setting provides a smoother mean field, certainly due to the fact that v_Γ is partially expanded on coarse-resolution modes, which are expected to be smoother. Finally, we show in Figure 3.13 an insight of how the Girsanov correction v_Γ behaves in time. As the structure of the noise is stationnary, so is the structure of v_Γ because it relies on the same modes as the noise. What is interesting is the evolution of the amplitude of this field, which decreases in time, meaning that most of the calibration work is done in the very first days of simulation, potentially entailing non-physical features as it can be observed in the mean fields in Figure 3.12, but crucially feeding the forecast with the information of the upcoming observation. Once the forecast manages to get closer to the truth, the need for calibration is less crucial and the Girsanov correction gets weaker.

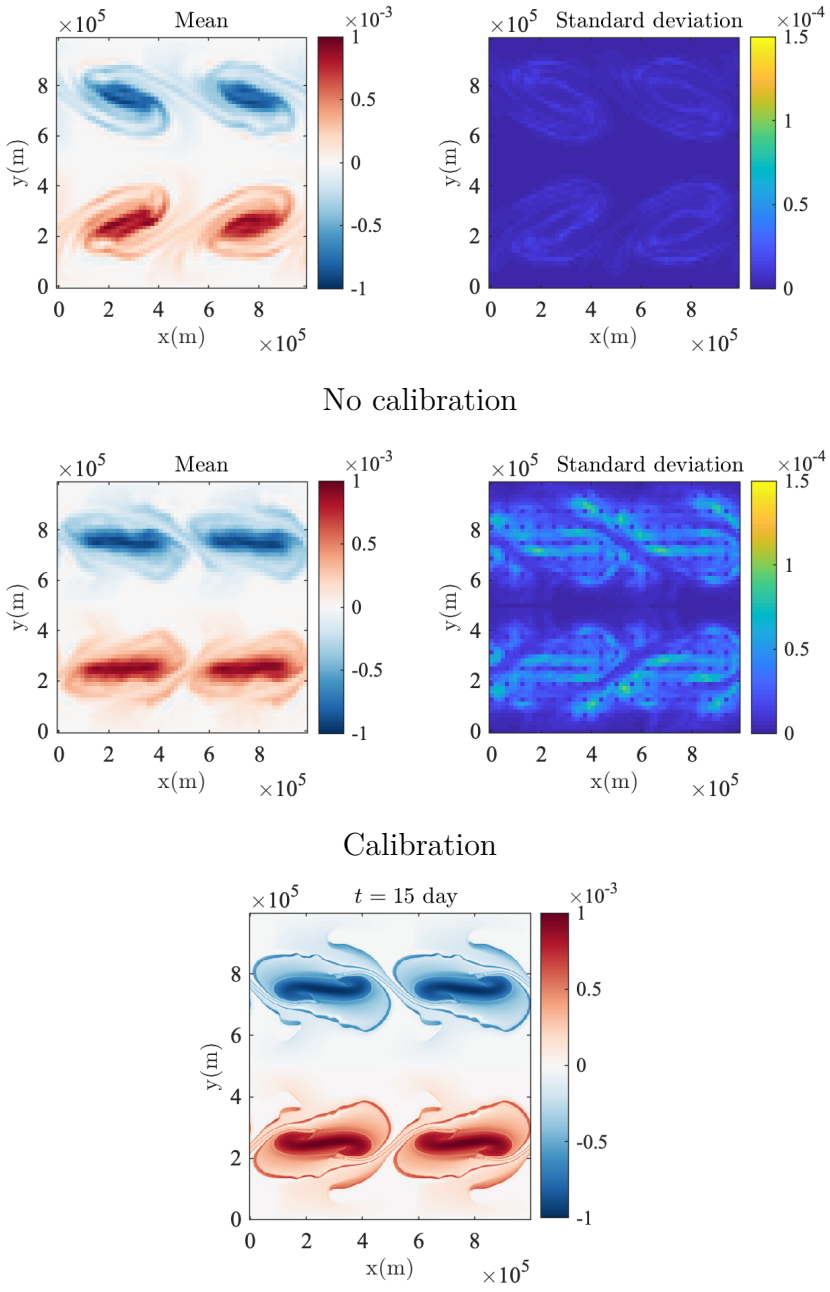
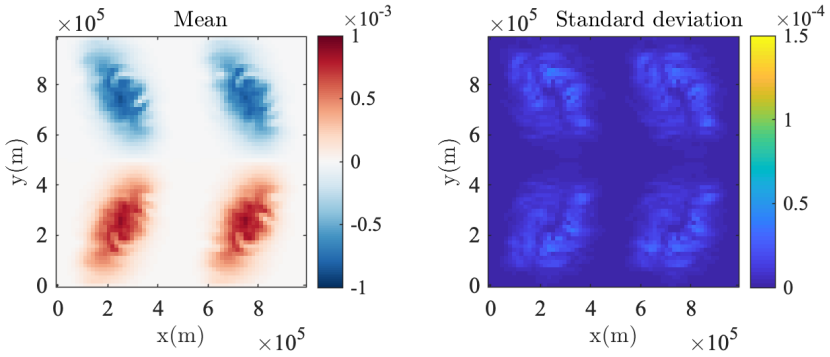
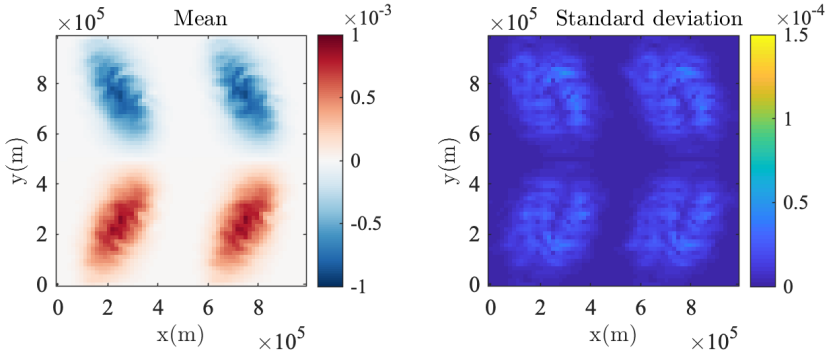


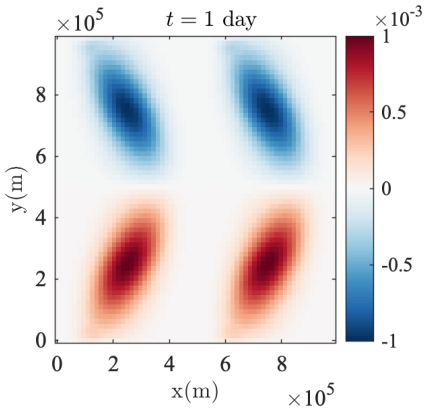
Figure 3.11 – Comparison between the ensemble mean (left) and the ensemble standard deviation (right) buoyancy maps, with and without calibration, at day 15 with the high-resolution truth.



Single-resolution calibration

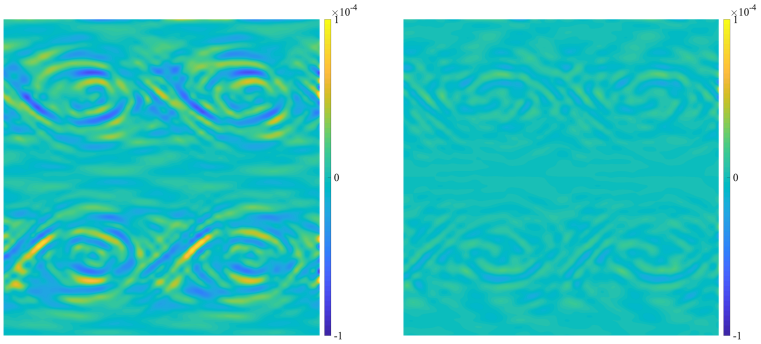


Two-resolution calibration



High-resolution truth

Figure 3.12 – Comparison between the ensemble mean (left) and the ensemble standard deviation (right) buoyancy maps, with single and two-resolution calibration, at day 1 with the high-resolution truth.



Day 0 Day 17

Figure 3.13 – Vorticity of the Girsanov drift v_Γ computed for one ensemble member at the first time step after the initial condition (left) and at the first time step after day 17 (right).

3.4 Synthesis

This chapter recaps our findings on the application of ensemble data assimilation methods to the LU-SQG model. The results of Section 3.2 show that LU seems to be a more efficient and secure way to replace multiplicative inflation for ensemble Kalman filters in the SQG system. It prevents explosion in finite time that occurs for deterministic inflation coupled with localization, and performs better both in terms of MSE and spread relevance. Section 3.3 describes a methodology, based on LU and Girsanov transforms, that aims at modifying the forecast step in order to guide the ensemble towards a desired region close to the observation. In the case of underestimation of the initial conditions, the numerical results show significant improvements brought by this additional technique. A multi-resolution scheme is also presented and yields additional improvements both in terms of MSE and field smoothness in the early stages of the filter.

Chapter 4

Data assimilation for ensemble forecast in RKHS

This chapter investigates the link between data assimilation methods and kernel-based approaches. The framework of reproducing kernel Hilbert spaces (RKHS) is stated in connection with the Koopman operator in order to get these spaces to be intrinsically related to the dynamics at play. A spectral representation of the Koopman operator is formulated in this setting. Within this framework, some classical data assimilation techniques are adapted and enriched with the theoretical properties of RKHS. Among them, the superposition principle assumeded by EnKFs is interpreted as a consequence of a fundamental property of RKHS, and its use is legitimated in this setting.

4.1 Reproducing kernel Hilbert spaces

A RKHS $\mathcal{H}$, is a Hilbert space of smooth complex functions $f : E \mapsto \mathbb{C}$ defined over a non empty set E on which a positive definite kernel and a kernel-based inner product, $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ can be defined. Throughout this work, we will consider the set E to be a locally convex topological compact set as we will work with an integral compact operator from which a convenient functional description can be set and as we will deal with a RKHS of functions that are Gateaux differentiable – through an assumption of smooth enough kernels. RKHS possess remarkable properties, which make their use very appealing in statistical machine learning applications and interpolation problems (Berlinet and Thomas-Agnan, 2001; Cucker and Smale, 2001). The kernels from which they are defined have a so called « reproducing property ».

Definition 1 (Reproducing kernel)**.** *Let $\mathcal{H}$ be a Hilbert space of $\mathbb{C}$-valued functions defined on a non-empty compact locally convex topological space E. A map $k : E \times E \mapsto \mathbb{C}$ is called a reproducing kernel of $\mathcal{H}$ if it satisfies the following principal properties: $\forall x \in E$,*

- *membership of the evaluation function* $k(\cdot, x) \in \mathcal{H}$,
- *reproducing property* $\forall f \in \mathcal{H}$, $\left\langle f, k(\cdot, x) \right\rangle_{\mathcal{H}} = f(x)$.

The last property provides an expression of the kernel as $k(x, y) = \overline{k(y, x)} = \left\langle k(\cdot, y), k(\cdot, x) \right\rangle_{\mathcal{H}}$, which is Hermitian – with $\bar{\bullet}$ denoting complex conjugate –, positive definite and associated with a continuous evaluation function $\delta_x f = \left\langle f, k(\cdot, x) \right\rangle_{\mathcal{H}} = f(x)$. The continuity of the Dirac evaluation operator is indeed sometimes taken as a definition of RKHS. By the Moore-Aronszajn theorem (Aronszajn, 1950), the kernel k defines uniquely the RKHS, $\mathcal{H}$, and vice versa. The set spanned by the feature maps $\mathrm{Span}\{k(\cdot, x), x \in E\}$, is dense in $(\mathcal{H}, \|\cdot\|_{\mathcal{H}})$. We note also that useful kernel closure properties enable to define kernels through operations such as addition, Schur product, and function composition (Berlinet and Thomas-Agnan, 2001). Besides, RKHS can be meaningfully characterized through integral operators, leading to an isometry with $L^2_{\mathbb{C}}(E, \nu)$ the space of square integrable functions defined on a compact metric space E with finite measure ν Cucker and Smale (2001).

Integral kernel operators

Let $k : E \times E \mapsto \mathbb{C}$ be $C^{(1,1)}(E \times E)$ (i.e. one time differentiable with respect to each argument), Hermitian, and positive definite, and let the map $\mathcal{L}_k : L^2_{\mathbb{C}}(E, \nu) \mapsto L^2_{\mathbb{C}}(E, \nu)$ be defined as:

$$\Big(\mathcal{L}_k f\Big)(x) = \int_E k(x, y) f(y)\, \nu(dy). \tag{4.1}$$

This operator, which must be understood within the composition with the continuous inclusion $i : C^0(E, \mathbb{C}) \hookrightarrow L^2_{\mathbb{C}}(E, \nu)$, can be shown to be well defined, positive, compact and self-adjoint (Cucker and Smale, 2001). The range of this operator is assumed to be dense in $L^2_{\mathbb{C}}(E, \nu)$. From Mercer's theorem (König, 1986), the feature maps $k(\cdot, x)$ span a RKHS defined through the eigenpairs $(\mu_i, \varphi_i)_{i \in \mathbb{N}}$ of the kernel operator $\mathcal{L}_k$:

$$\mathcal{H} := \Big\{ f \in L^2_{\mathbb{C}}(E, \nu),\ f = \sum_{i=0}^{\infty} a_i \varphi_i \ :\ \sum_i \frac{|a_i|^2}{\mu_i} < \infty \Big\}, \tag{4.2}$$

with no null eigenvalues since we have assumed that the kernel range is dense in $L^2_{\mathbb{C}}(E, \nu)$. The rank of the kernel (number of – non-zero – eigenvalues) corresponds to the dimension of $\mathcal{H}$, and will potentially be infinite in this work, as E will be a subspace of functions. The RKHS $\mathcal{H}$ is a space of smooth functions that expand on the eigenfunctions of $\mathcal{L}_k$

with decreasing coefficients. In fact, there exists a constant $C > 0$ such that, for all $u \in E$ and all $f \in \mathcal{H}$, we have $\|\partial_u f\|_{L^2_{\mathbb{C}}(E,\nu)} \leq C \ \|f\|_{\mathcal{H}}$ (Dufée et al., 2023b), where the derivative denotes the Gateaux directional derivative of function f in the direction u defined as

$$\forall x, u \in E, \partial_u f = \lim_{\epsilon \to 0} \frac{1}{\epsilon}\Big(f(x + \epsilon u) - f(x)\Big).$$

In order to properly define the Gateaux derivative, E should be embedded with a local vector space structure, which is for instance the case of differentiable manifolds. We may also define uniquely a square-root symmetric isometric bijective operator $\mathcal{L}_k^{1/2}$ between $L^2_{\mathbb{C}}(E,\nu)$ and $\mathcal{H}$. This operator enables to go from $L^2_{\mathbb{C}}(E,\nu)$ to $\mathcal{H}$ by increasing the functions regularity while its inverse lowers the function regularity by bringing them back to $L^2_{\mathbb{C}}(E,\nu)$. Both operators are bounded. The injection $j : \mathcal{H} \mapsto L^2_{\mathbb{C}}(E,\nu)$ is continuous and compact (Dufée et al., 2023b) and $j(\mathcal{H})$ is assumed to be dense in $L^2_{\mathbb{C}}(E,\nu)$ If $\mathcal{L}_k$ has a non trivial kernel, then $\mathcal{L}_k^{1/2}$ is a bijection between $\mathcal{V} = \mathrm{Ker}(\mathcal{L}_k)^{\perp} \subset L^2_{\mathbb{C}}(E,\nu)$ and $\mathcal{H}$.

Remark 1. *The spaces $\mathcal{H}$ and $L^2_{\mathbb{C}}(E,\nu)$ are of course considered here as infinite dimensional spaces, preventing E to be a finite set.*

4.2 Dynamical systems on a RKHS family

We consider an invertible nonsingular dynamical system $X(t) = \Phi_t(X_0)$, defined from a continuous flow, Φ_t (meaning that, for any $X \in \Omega$, the mapping $t \mapsto \Phi_t(X)$ is continuous), on a compact invariant phase space differentiable manifold, Ω, (i.e. $\Phi_t^{-1}(\Omega) = \Omega, \forall t \in \mathbb{R}^+$) of time evolving vector functions over a spatial support Ω_x. The functions $X : \mathbb{R}^+ \times \Omega_x \mapsto \mathbb{R}^d$ with $X \in C^p$, $p \geq 1$, are solutions of the following d-dimensional differential system:

$$\begin{cases} \partial_t X(t,\cdot) = \mathcal{M}\Big(X(t,\cdot)\Big), \text{ with } X(t,\cdot) \in \Omega, \forall t > 0, \\ X(0,\cdot) = X_0(\cdot). \end{cases} \tag{4.3}$$

The nonlinear differential operator $\mathcal{M} : \Omega \to \Omega$ is assumed to be C^1, and in particular its linear tangent expression defined as the Gateaux derivative:

$$\partial_X \mathcal{M}(X)\delta X = \lim_{\beta \to 0} \frac{1}{\beta}\Big(\mathcal{M}\Big(X(t,\cdot) + \beta\delta X(t,\cdot)\Big) - \mathcal{M}\Big(X(t,\cdot)\Big)\Big)$$

is such that $\sup_{X\in\Omega} \partial_X \mathcal{M}(X) < \infty$ (since Ω is compact). We consider the measure space (Ω_x, Lb) where Lb is the Lebesgue measure on Ω_x. We denote $L^2(\Omega_x, \mathbb{R}^d)$ the space of the square-integrable functions on (Ω_x, Lb) and $L^2(\Omega_x, \mathbb{R}^d) := \{f = (f_1, .., f_d) : \Omega_x \mapsto \mathbb{R}^d \quad : \quad \text{for all } 1 \leq i \leq d \quad f_i \in L^2(\Omega_x, \mathbb{R})\}$. We note $\|\cdot\|_{L^2}$ the norm associated with $L^2(\Omega_x, \mathbb{R}^d)$, which is given by $\|f\|^2_{L^2} := \sum_{i=1}^d \|f_i\|^2_{L^2(\Omega_x,\mathbb{R})}$ for all $f \in L^2(\Omega_x, \mathbb{R}^d)$. The set Ω is included in $L^2(\Omega_x, \mathbb{R}^d)$. The system (4.3) is assumed to admit a finite invariant measure ν on Ω (note that from the invertibility property, the measure is also nonsingular with $\nu(\Phi_t^{-1}(A)) = 0$, $\forall A \subset \Omega$ such that $\nu(A) = 0$). The system's observables are square integrable measurable complex functions with respect to measure ν. They belong to the Hilbert space $L^2_{\mathbb{C}}(\Omega, \nu)$ with the inner product $\langle \cdot\,,\, \cdot \rangle_{L^2_{\mathbb{C}}(\Omega,\nu)}$ given for f and $g \in L^2_{\mathbb{C}}(\Omega, \nu)$ by

$$\langle f\,,\, g\rangle_{L^2_{\mathbb{C}}(\Omega,\nu)} := \int_\Omega f(y)\, \overline{g(y)}\ \nu(\mathrm{d}y).$$

Depending on the context, $X(t, \cdot)$ or X_t will denote either an element of Ω or the function $X(t, \cdot)\colon \Omega_x \to \mathbb{R}^d$ such that $x \mapsto X(t, x)$.

In this work, the set of different initial conditions Ω_0 is an infinite compact subset of $L^2(\Omega_x, \mathbb{R}^d)$. For all time $t \geq 0$, we denote by $\Omega_t \subset \Omega \subset L^2(\Omega_x, \mathbb{R}^d)$ the space defined by $\Omega_t := \Phi_t(\Omega_0)$. Furthermore, the set Ω_0 of initial conditions will be assumed to be composed of points that uniquely characterize all the dynamical system trajectories and to be sufficiently rich so that $\bigcup_{t\geq 0}(\Omega_t) = \Omega$. Hence, by this, each point of the manifold, Ω, is assumed to be uniquely characterized by an initial condition and the integration of the dynamical system over a given time t. In other words, for any $X \in \Omega$, there exist a unique initial condition $X_0 \in \Omega_0$ and a unique time $t \in \mathbb{R}_+$ such that $X = \Phi_t(X_0)$. All the sets Ω_t will be assumed to be locally convex.

Defining at each time t, from the subset Ω_t, a positive Hermitian kernel $k_{\Omega_t} : \Omega_t \times \Omega_t$, there exists a unique associated RKHS $\mathcal{H}_t$. In the following, for the sake of concision, the kernels k_{Ω_t} will be denoted by k_t to refer to the dependence on the set Ω_t. For all $t \geq 0$, we will use the notation $X_t = \Phi_t(X_0)$ and $\left(\mathcal{H}_t, \langle \cdot\,,\, \cdot \rangle_{\mathcal{H}_t}, \|\cdot\|_{\mathcal{H}_t}\right)$ for the RKHS associated with the kernel k_t defined on $\Omega_t \times \Omega_t$. The kernels are assumed to be $C^{(1,1)}(\Omega_t \times \Omega_t)$ and as a consequence, the associated feature maps have derivatives in $\mathcal{H}_t$ (Dufée et al., 2023b). The RKHSs $\mathcal{H}_t$ for all time $t \geq 0$ form a family of Hilbert spaces of complex functions, each of them equipped with their own inner product $k_t(Y_t, X_t)$, for all functions $X_t, Y_t \in \Omega_t$. At any time, a measurable function of the system state, usually often referred to as an observable f, belonging to the RKHS $\mathcal{H}_t$ can be described as a linear combination of the feature maps $\{k_t(\cdot, X_t), X_t \in \Omega_t\}$ As it will be shown, the features maps of this

RKHS family can be expressed on a time evolving orthonormal system of basis functions, connected to each other through an exponential form and given by the eigenfunctions of the infinitesimal generator of a "Koopman-like" operator defined on the RKHS family. The RKHS family is defined by $\mathcal{W} = (\mathcal{H}_t)_{t\geq 0}$. In the following, we present a summary of the mathematical results associated to the Koopman operator in the RKHS family.

Numerically, in practice, the setup is based on N realizations (called ensemble members) of this dynamical system, $\{X_t^{(i)}, i = 1, \ldots, N\}$, generated from a finite set of different initial conditions $\{X_0^{(i)}, i = 1, \ldots, N\} \subset \Omega_0$, and are available up to time T. Still, we underline that, in the following development, the time horizon can be infinite, the sets Ω_t are infinite and the corresponding RKHS $\mathcal{H}_t$ are infinite dimensional. This setting (both practically and theoretically) is quite common for ensemble methods applied to geophysical systems. The ensemble size is small in general, while the phase space is in theory infinite (or at least very high) dimensional.

4.3 Koopman operator in the RKHS family

So far we did not give any precise definition of the kernels associated to the RKHS family $\mathcal{H}_t$ yet. These kernels are defined from a known *a priori* initial kernel, $k_0 : \Omega_0 \times \Omega_0$, as:

Definition 2 ($\mathcal{H}_t$ kernel)**.** *The kernel k_t associated to the RKHS $\mathcal{H}_t$ are defined as*

$$\forall X_t, Y_t \in \Omega_t, \quad k_t\Big(Y_t\,,\, X_t\Big) = k_0\Big(\Phi_t^{-1}(Y_t)\,,\, \Phi_t^{-1}(X_t)\Big), \tag{4.4}$$

where $k_0 : \Omega_0 \times \Omega_0$ is a given kernel.

These kernels can also be equivalently defined introducing the operators $\mathcal{U}_t$ acting on the feature maps. An isometric property of this operator on the RKHS family enables us to fully define the kernels along time, in the same way as in the previous definition. The operator $\mathcal{U}_t : \mathcal{H}_0 \mapsto \mathcal{H}_t$ is defined such that

$$\mathcal{U}_t k_0(\cdot, X_0) = k_t\Big(\cdot, \Phi_t(X_0)\Big), \tag{4.5}$$

and transports the kernel feature maps on the RKHS family by composition with the system's dynamics. This operator, and more specifically its infinitesimal operator, will enable us to define the feature maps of $\mathcal{H}_t$ from the initial feature maps on $\mathcal{H}_0$. As it will

be detailed in section 4.3.1, the operator $\mathcal{U}_t$ is indeed directly related to the restriction on $\mathcal{H}_t$ of the adjoint of the Koopman operator U_t on a bigger RKHS space $\mathcal{H}$, associated to a fixed kernel defined on the whole phase space Ω. As $\mathcal{U}_t$ propagates forward the second argument of the feature maps, it is referred to in the following as the Koopman kernel operator in the RKHS family. Indeed, it will be pointed out that, for any $f \in \mathcal{H}$ and any $X_0 \in \Omega_0$,

$$\begin{aligned} U_t f(X_0) = f(X_t) &= \Big\langle R_t f \,,\, k_t(\cdot\,,\, X_t) \Big\rangle_{\mathcal{H}_t} \\ &= \Big\langle R_t f \,,\, \mathcal{U}_t k_0(\cdot\,,\, X_0) \Big\rangle_{\mathcal{H}_t}, \end{aligned}$$

where U_t denotes the Koopman operator operating on $L^2_{\mathbb{C}}(\Omega, \nu)$ and $R_t f$ denotes the restriction of f on $\Omega_t \subset \Omega$. This expression corresponds to a kernel expression of the Koopman operator definition and formally, at this point, the operator $\mathcal{U}_t$ can hence be thought as a kernel expression of the Koopman operator.

The global kernel k (respectively the associated RKHS) is tightly bound to the time evolving kernels k_t (respectively $(\mathcal{H}_t)_{t\geq 0}$). The restriction on $\mathcal{H}$ of the Koopman operator U_t and its adjoint the Perron-Frobenius P_t exhibit some remarkable properties. As classically, the operators U_t and P_t are unitary in $L^2_{\mathbb{C}}(\Omega, \nu)$ (Prop. 1), but they have the striking property to be uniformly continuous in $L^2_{\mathbb{C}}(\Omega, \nu)$ (i.e. with bounded generators – Theorem 6). As such, they can be expended in an uniformly converging exponential series. Nevertheless, it must be outlined that the fixed kernel $k(x, y)$ and consequently $\mathcal{H}$ are in practice only partially accessible as they are defined on the whole manifold of the dynamics and such expansion cannot be directly used. A local representation of the RKHS family $\mathcal{W}$ is on the other hand much easier to infer in practice through the time evolution of ensembles Ω_t and kernels k_t. As we will see, the operator $\mathcal{U}_t$, on $\mathcal{W}$ inherits a lot of the properties of U_t and, in particular, a related form of exponential series expansion (Theorem 5).

The Koopman kernel operator in the RKHS family$\mathcal{U}_t$ defines an isometry from $\mathcal{H}_0$ to $\mathcal{H}_t$ (Theorem 7)

$$\begin{aligned} \Big\langle k_t\Big(\cdot,\Phi_t(Y_0)\Big), k_t\Big(\cdot,\Phi_t(X_0)\Big) \Big\rangle_{\mathcal{H}_t} &= \Big\langle k_0\Big(\cdot,Y_0\Big), k_0\Big(\cdot,X_0\Big) \Big\rangle_{\mathcal{H}_0}, \\ k_t\Big(\Phi_t(X_0), \Phi_t(Y_0)\Big) &= k_0(X_0, Y_0). \end{aligned} \tag{4.6}$$

This isometry ensues obviously directly from definition 2. But it can also be guessed from

definition (4.5) and the unitarity of $\mathcal{U}_t$ (Theorem 7, Prop.6), inherited from the unitarity in $L^2_{\mathbb{C}}(\Omega, \nu)$ of the Koopman operator and of its adjoint, the Perron-Frobenious operator. This property is of major practical interest as it allows us to define the kernels of the RKHS family from a given initial kernel fixed by the user. The kernels remain constant along the system trajectories. Alternatively, an explicit form of the feature maps can be obtained from an adjoint transport equation associated to the infinitesimal generator of the Koopman operator in the RKHS family. Nevertheless, the isometry is far more straightforward to use to set the kernel evolution. Strikingly, we have even more than this kernel isometry. An evolution operator $A_{U,t} : \mathcal{H}_t \to L^2_{\mathbb{C}}(\Omega_t, \nu)$, associated to the infinitesimal generator of Koopman operator U_t can also be defined as

$$A_{U,t}\, k_t(\cdot, X_t) := \partial_{\mathcal{M}(\cdot)}\, k_t(\cdot, X_t), \tag{4.7}$$

where $\partial_u k_t(\cdot, X_t)$ stands for the Gateaux directional derivative along $u \in \Omega$ of function $k(\cdot, X_t)$. This operator, that will be shown to be bounded (Prop.7) and skew-symmetric (Prop.8) for the inner product of $L^2_{\mathbb{C}}(\Omega, \nu)$, plays the role of an infinitesimal generator on $\mathcal{W}$ and enables us expressing an exponential expansion of $\mathcal{U}_t$.

Theorem 5 (The RKHS family spectral representation)**.** *For a measure preserving invertible dynamical system of the form (4.3), assuming a $C^{(2,2)}$ initial kernel, the evolution operator $A_{U,t} : \mathcal{H}_t \mapsto L^2_{\mathbb{C}}(\Omega_t, \nu)$ defined in (4.7), and which is defined from the infinitesimal generator of the Koopman operator in $L^2_{\mathbb{C}}(\Omega, \nu)$, can be diagonalized, at any time $t \geq 0$, by an orthonormal basis $(\psi^t_\ell)_\ell$ of $\mathcal{H}_t$ such that, for all $X_t \in \Omega_t$,*

$$A_{U,t}\psi^t_\ell(X_t) = \lambda_\ell j \circ \psi^t_\ell(X_t),$$

where j is the injection $j : \mathcal{H}_t \mapsto L^2_{\mathbb{C}}(\Omega, \nu)$. We have in addition the following relation between the orthonormal basis systems along time:

$$\forall t \geq 0, \quad \forall X_0 \in \Omega_0, \quad \psi^t_\ell(X_t) = \exp(t\ \lambda_\ell)\psi^0_\ell(X_0), \tag{4.8}$$

with $X_t = \Phi_t(X_0)$. Furthermore, the purely imaginary eigenvalues $(\lambda_\ell)_\ell$ do not depend on time.

This theorem, which constitutes our main result, provides us a time-evolving system of orthonormal bases of the RKHS family. It brings us the capability to express any observable of the system in terms of bases that are intrinsically linked to the dynamics and

related to each other by an exponential relation. The eigenvalues of $A_{U,t}$ are purely imaginary since this operator is skew-symmetric in $L^2_{\mathbb{C}}(\Omega_t, \nu)$. Remarkably, the eigenvalues of each $A_{U,t}$ do not depend on time and are connected with the same covariant eigenfunctions (in the sense of (4.8)). These eigenfunctions correspond to restrictions of eigenfunctions of the infinitesimal generator of Koopman/Perron-Frobenius operators defined on $\mathcal{H}$. The scheme of the proofs of these results are presented in the next section.

The kernel isometry (4.6) (Theorem 7) and the Koopman spectral representation in the RKHS family (Theorem 5) constitute fundamental results enabling us to build very simple ensemble-based trajectory reconstructions for new initial conditions without the requirement of resimulating the dynamical system. Amazingly, the family of RKHS together with the Koopman isometry allows to define a system's trajectory as a constant-in-time linear combination of the time varying RKHS feature maps. Several of such data assimilation techniques, based on this fully justified superposition principle, will be derived in section 4.4. In the next section, we demonstrate the different properties related to the RKHS family.

4.3.1 Schemes of proof on the properties of the Koopman operator in the RKHS family and of the RKHS family spectral theorem

As explained in the previous section, the RKHS family $\left(\mathcal{H}_t\right)_{t\geq 0}$ does not have a good topology to work with. We first need to define a "big" set with a better topology and which encompasses all the RKHSs $\mathcal{H}_t$. On this big encompassing set, we shall then define a Koopman operator, enabling us to study properly the Koopman operator in the RKHS family.

Construction of the « big » encompassing set $\mathcal{H}$

The phase space Ω corresponds to the set generated by the values of the dynamical system at a given time t. We hence note that Ω_t is a subset of Ω. Each point of Ω is a phase-space point $X = \Phi_t(X_0)$ uniquely defined from time t and initial condition $X_0 \in \Omega_0$. In order to define the RKHS $\mathcal{H}$, let us define, from the kernels $k_t : \Omega_t \times \Omega_t$, a symmetric positive definite map $k : \Omega \times \Omega$.

Definition 3 ($\mathcal{H}$ kernel)**.** *For all $X = \Phi_r(X_0)$, $Y = \Phi_s(Y_0) \in \Omega$, with X_0, $Y_0 \in \Omega_0$, we define*

$$k(X\,,\,Y) = k_0\Big(X_0\,,\,Y_0\Big)\ell(r,s) = k_t\Big(\Phi_t\Big(X_0\Big)\,,\,\Phi_t\Big(Y_0\Big)\Big)\ell(r,s) \quad \forall t \geq 0, \tag{4.9}$$

where $\ell : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ is a symmetric kernel defined, for all $r, s \in \mathbb{R}$ by

$$\ell(r,s) = \varphi(r-s), \tag{4.10}$$

where φ is a twice-differentiable even function such that $\varphi(0) = 1$.

The positivity and symmetry of kernel k ensues from the properties of kernels k_0 and ℓ, which are assumed to be valid kernels. Kernel k inherits the regularity conditions of k_0 and ℓ and is $C^{(1,1)}(\Omega\times\Omega)$ as well. In the trivial case where $\varphi = 1$, then comparing any pair of points on two trajectories would be equivalent to compare the initial conditions, which would result in a quite poor kernel and degeneracy issues. In order to enrich the kernel structure, one can think of φ as a regularized Dirac distribution, or a time Gaussian distribution, that will discriminate the points of the phase space that are reached at different times.

By the Moore-Aronszajn theorem, there exists a unique RKHS $(\mathcal{H}, \langle\,\cdot\,,\,\cdot\,\rangle_{\mathcal{H}}, \|\,\cdot\,\|_{\mathcal{H}})$ with kernel k. We note that, in practice, the full knowledge of the phase-space is completely unreachable. Again, we therefore stress the fact that the setting of this encompassing RKHS $\mathcal{H}$ has only a theoretical purpose. The RKHS $\mathcal{H}$ can be connected to each RKHS of the family $\mathcal{W}$ through extension and restriction operators denoted E_t and R_t respectively, and defined as follows. For all $t \geq 0$, let

$$\begin{array}{rccc} E_t : & \mathcal{H}_t & \mapsto & \mathcal{H} \\ & k_t(\,\cdot\,, X_t) & \mapsto & k(\,\cdot\,,\,X_t) \end{array} \tag{4.11}$$

and extend this definition by linearity on $\mathrm{Span}\{k_t(\,\cdot\,, X_t)\,:\,X_t \in \Omega_t\}$. Then, by density, the function $E_t(f)$ is defined for all $f \in \mathcal{H}_t$. The restriction

$$\begin{array}{rccc} R_t : & \mathcal{H} & \mapsto & \mathcal{H}_t \\ & k(\,\cdot\,,\,X) = k(\,\cdot\,,\,\Phi_r(X_0)) & \mapsto & k(\,\cdot\,,\,X)\restriction_{\Omega_t} = k_t\Big(\,\cdot\,,\,\Phi_t(X_0)\Big)\ell(t,r) \end{array} \tag{4.12}$$

is defined similarly for $g \in \mathcal{H}^{\mathrm{sp}} = \mathrm{Span}\{k\Big(\,\cdot\,, X\Big)\,:\,X \in \Omega\}$ and extended by density in $\mathcal{H}$ by the Moore-Aronszajn theorem Aronszajn (1950). The extension map is built in such

a way that each RKHS $\mathcal{H}_t$ of the family is included in the « big » encompassing RKHS $\mathcal{H}$.

In (Dufée et al., 2023b), several useful properties of the restriction and extension operators are listed. Namely, E_t and R_t are both isometries, they form an adjoint pair and the restriction is continuous in $L^2_{\mathbb{C}}(\Omega_t, \nu)$. We define now the Koopman operator on the encompassing RKHS $\mathcal{H}$.

Koopman operators on $\mathcal{H}$

For all $t \geq 0$, we consider the Koopman operator $U_t : \mathcal{H} \to \mathcal{H}$ defined by

$$U_t(f)(X) := f \circ \Phi_t(X) = f\Big(\Phi_t(X)\Big), \text{ for all } f \in \mathcal{H}. \tag{4.13}$$

Since $\mathcal{H}$ is dense in $(L^2_{\mathbb{C}}(\Omega, \nu), \| \cdot \|_{L^2_{\mathbb{C}}(\Omega,\nu)})$, the operator U_t can be continuously extended on $L^2_{\mathbb{C}}(\Omega, \nu)$ and, to avoid notation inflation, we keep denoting U_t this extension. We first study $U_t : L^2_{\mathbb{C}}(\Omega, \nu) \mapsto L^2_{\mathbb{C}}(\Omega, \nu)$ with the $L^2_{\mathbb{C}}(\Omega, \nu)$ topology. The family $(U_t)_{t\geq 0}$ is a strongly continuous semi-group on $(L^2_{\mathbb{C}}(\Omega, \nu), \| \cdot \|_{L^2_{\mathbb{C}}(\Omega,\nu)})$ since $t \mapsto \Phi_t(\cdot)$ is continuous on $\mathbb{R}^+$. As the feature maps are functions of $\mathcal{H}$, it can be noticed that, for all $X_r \in \Omega$,

$$U_t\,[\,k(\,\cdot\,,\,X_r)\,]\;=\;k\Big(\,\Phi_t(\cdot)\,,\,X_r\Big), \tag{4.14}$$

which justifies the stability of $\mathcal{H}$ by the operator U_t. This corresponds to a natural expression of the Koopman operator for any function $g \in \mathcal{H}^{\text{sp}}$, extended then by density to $\mathcal{H}$. Yet another useful equivalent expression of the Koopman operator is available for the feature maps. We have, for any points $X = \Phi_r(X_0), Y = \Phi_s(Y_0) \in \Omega$,

$$U_t\,[\,k(\,\cdot\,,\,X)\,]\,(Y) = k\Big(\Phi_t(Y)\,,\,X\Big) = k\Big(\Phi_{t+s}\Big(Y_0\Big)\,,\,\Phi_r\Big(X_0\Big)\Big) = k_0(Y_0, X_0)\ell(t+s, r). \tag{4.15}$$

From the properties of the time kernel ℓ, we get

$$\ell(t+s, r) = \varphi(t+s-r) = \varphi(s-(r-t)) = \ell(s, r-t), \tag{4.16}$$

which leads to

$$k\Big(\Phi_t(Y)\,,\,X\Big) = k_0(Y_0, X_0)\ell(t+s, r) = k_0(Y_0, X_0)\ell(s, r-t) = k\Big(\Phi_s\Big(Y_0\Big)\,,\,\Phi_{r-t}\Big(X_0\Big)\Big) = k\Big(Y, \Phi_t^{-1}(X)\Big). \tag{4.17}$$

and hence

$$U_t\,[\,k(\,\cdot\,,\,X)\,]\,(Y) = k\Big(Y, \Phi_t^{-1}(X)\Big). \tag{4.18}$$

As the previous equality is true for all $Y \in \Omega$, this implies that, for all $X \in \Omega$,

$$U_t\,[\,k(\,\cdot\,,\,X)\,] \;=\; k\Big(\,\cdot\,,\,\Phi_t^{-1}(X)\Big). \tag{4.19}$$

This dual formulation of the kernel expression of the Koopman operator is intrinsically linked to the definition of the kernel k. This dual expression will be of central interest in the following, as it enables to formulate the time evolution of the feature maps in terms of the Koopman operator U_t and its adjoint at any time $t \geq 0$.

Remark 2 (Transport of the kernel k)**.** *For all $X = \Phi_r(X_0) \in \Omega$ and $t \geq 0$, by definition of the Koopman operator on the feature maps and (4.19), $U_t\,[\,k(\,\cdot\,,\,X)]$ has two expressions and we obtain*

$$U_t\,[\,k(\,\cdot\,,\,X)] = k\Big(\,\Phi_t(\cdot)\,,\,X\,\Big) = k\Big(\,\cdot\,,\,\Phi_t^{-1}(X)\Big)\,.$$

The next remark provides a useful commutation property between U_t and the kernel integral operator $\mathcal{L}_k$ defined in equation (4.1) or of its unique symmetric square-root $\mathcal{L}_k^{1/2}$ defined from the square-root of the kernel eigenvalues (see Dufée et al. (2023b) for a precise definition in the general case). Note that in the case of $\mathcal{H}$, the kernel integral operator is indeed a complex object which hides a time dependency.

Remark 3 (Commutation between $\mathcal{L}_k^{1/2}$ (or $\mathcal{L}_k$) and U_t)**.** *For all $X \in \Omega$, we have*

$$\mathcal{L}_k^{1/2} \circ U_t\,[k(\,\cdot\,,\,X)] \;=\; U_t \circ \mathcal{L}_k^{1/2}\,[k(\,\cdot\,,\,X)].$$

This commutation property, that ensues directly from the compositional nature of the Koopman operator, allows us to write immediately the equality

$$\mathcal{L}_k \circ U_t\,[k(\,\cdot\,,\,X)] \;=\; U_t \circ \mathcal{L}_k\,[k(\,\cdot\,,\,X)]. \tag{4.20}$$

By linearity these properties extend to all functions of $\mathcal{H}$.

The next proposition shows the Koopman operator defined on RKHS $\mathcal{H}$ is unitary in $L^2_{\mathbb{C}}(\Omega, \nu)$, which is a classical property of the Koopman operator in L^2 for measure preserving invertible systems.

Proposition 1 (Unitarity of the Koopman operator in $L^2_{\mathbb{C}}(\Omega, \nu)$)**.** *The map $U_t : (L^2_{\mathbb{C}}(\Omega, \nu), \|\cdot\|_{L^2_{\mathbb{C}}(\Omega,\nu)}) \to (L^2_{\mathbb{C}}(\Omega, \nu), \|\cdot\|_{L^2_{\mathbb{C}}(\Omega,\nu)})$ is unitary for all $t \geq 0$.*

Proposition 1, shows that the Koopman operator U_t is invertible and that its inverse in $L^2_{\mathbb{C}}(\Omega,\nu)$ is U_t^*. Denoting by P_t the operator defined by $P_t := U_t^*$, for all f and $g \in L^2_{\mathbb{C}}(\Omega,\nu)$ we have $\langle\, U_t(f)\,,\, g\rangle_{L^2_{\mathbb{C}}(\Omega,\nu)} = \langle f\,,\, P_t(g)\,\rangle_{L^2_{\mathbb{C}}(\Omega,\nu)}$. The family $(P_t)_t$ is a strongly continuous semi-group on $\Big(L^2_{\mathbb{C}}(\Omega,\nu), \|\cdot\|_{L^2_{\mathbb{C}}(\Omega,\nu)}\Big)$. This operator is referred to as the Perron-Frobenius operator. For all $t \geq 0$, the Perron-Frobenius operator P_t verifies for all $X \in \Omega$

$$P_t[k\Big(\Phi_t(\cdot), X\Big] = k(\,\cdot\,, X). \tag{4.21}$$

From Remark 1, we can write a more explicit expression of P_t on the featutre maps: for all $X \in \Omega$,

$$P_t\Big[k(\cdot, X)\Big] = k\Big(\cdot, \Phi_t(X)\Big). \tag{4.22}$$

Informally, if we see the function $k(\cdot, X)$ as an atom of the measure, then its expression at a future time is provided by (4.22), which corresponds well to the idea that the Perron-Frobenius operator advances in time the density.

As previously stated, the Koopman operator U_t is an isometry in $L^2_{\mathbb{C}}(\Omega,\nu)$. The next result asserts that U_t is also an isometry in $\mathcal{H}$.

Proposition 2 (Isometric relation of the kernel)**.** *For all X and $Y \in \Omega$, we have* $k\Big(\Phi_t(X)\,,\Phi_t(Y)\Big) = k(X\,,\,Y)$.

This results follows immediately from the kernel definition: for all $X = \Phi_r(X_0), Y = \Phi_s(Y_0) \in \Omega$, we have

$$k\Big(\Phi_t(X)\,,\Phi_t(Y)\Big) = k\Big(\Phi_{t+r}(X_0), \Phi_{t+s}(Y_0)\Big) = k_0(X_0,Y_0)\ell(t+r,t+s) = k_0(X_0,Y_0)\ell(r,s) = k(X,Y). \tag{4.23}$$

Remark 2 on the application of the operator U_t to $k(\,\cdot\,, X)$, with $X \in \Omega$, shows that applying the flow Φ_t on the first variable of $k(\,\cdot\,,\,\cdot\,)$. is equivalent to applying Φ_t^{-1} to the second variable and vice-versa. Consequently, for all $f \in L^2_{\mathbb{C}}(\Omega,\nu)$ and $X \in \Omega$, we have $\langle f\,,\, k\Big(\,\cdot\,,\,\Phi_t(X)\Big)\,\rangle_{L^2_{\mathbb{C}}(\Omega,\nu)} = \langle f\,,\, k\Big(\,\Phi_t^{-1}(\cdot)\,,\,X\Big)\,\rangle_{L^2_{\mathbb{C}}(\Omega,\nu)} = \langle f\ ,\ P_t\,k(\,\cdot\,,\,X)\rangle_{L^2_{\mathbb{C}}(\Omega,\nu)}$ and by definition of the adjoint, we obtain

$$\Big\langle f\,,\, k\Big(\,\cdot\,,\,\Phi_t(X)\Big)\,\Big\rangle_{L^2_{\mathbb{C}}(\Omega,\nu)} = \Big\langle U_t(f)\ ,\ k(\,\cdot\,,\,X)\Big\rangle_{L^2_{\mathbb{C}}(\Omega,\nu)}. \tag{4.24}$$

We already knew that this equality was right for $f \in \mathcal{H}$ and for the inner product in

$\mathcal{H}$, namely

$$\Big\langle f, k\Big(\cdot, \Phi_t(X)\Big)\Big\rangle_{\mathcal{H}} = f\Big(\Phi_t(X)\Big) = U_t f(X) = \Big\langle U_t(f), k(\cdot, X)\Big\rangle_{\mathcal{H}}. \tag{4.25}$$

Equation (4.24) provides a weak (in the sense that the $L^2_{\mathbb{C}}(\Omega,\nu)$-inner product against a feature map is no longer the evaluation function) formulation of the transport of any observable in $L^2_{\mathbb{C}}(\Omega,\nu)$ by the flow.

In order to further exhibit several analytical results on the Koopman operator $(U_t)_{t\geq 0}$ in $L^2_{\mathbb{C}}(\Omega,\nu)$, we introduce in the following its infinitesimal generator.

Koopman infinitesimal generator

We will note as $\Big(A_{U}, \mathcal{D}(A_{U})\Big)$ the infinitesimal generator of the strongly continuous semigroup $(U_t)_{t\geq 0}$ on $L^2_{\mathbb{C}}(\Omega,\nu)$ and its domain. As the Koopman and Perron-Frobenius operators are adjoint in $L^2_{\mathbb{C}}(\Omega,\nu)$ their infinitesimal generators are also adjoint of each other with possibly their own domain. The following lemma characterizes first the Perron-Frobenius infinitesimal generator A_{P} and its domain $\mathcal{D}\left(A_{P}\right)$.

Lemma 1 (Perron-Frobenius infinitesimal generator)**.** *The Perron-Frobenius infinitesimal generator is the unbounded operator,* $(A_{P}, \mathcal{D}\left(A_{P}\right))$ *defined by*

$$\begin{aligned}\mathcal{D}\left(A_{P}\right) &= \Big\{ f \in L^2_{\mathbb{C}}(\Omega,\nu) \ : \ x \mapsto \partial_{\mathcal{M}(x)} f(x) \ \in L^2_{\mathbb{C}}(\Omega,\nu) \Big\} \ \textit{and} \\ A_{P} f &:= -\partial_{\mathcal{M}(\cdot)}\, f(\cdot) \quad \textit{for} \quad f \in \mathcal{D}\left(A_{P}\right),\end{aligned}$$

where $\mathcal{M}$ *denotes the differential operator of the system dynamics* (4.3) *and* $\partial_u f$ *stands for the directional derivative of* f *along* $u \in \Omega$*.*

For all $t \geq 0$, P_t is the adjoint of U_t in $L^2_{\mathbb{C}}(\Omega,\nu)$. The Koopman infinitesimal generator in $L^2_{\mathbb{C}}(\Omega,\nu)$ is consequently given by

$$\mathcal{D}(A_{U}) = \mathcal{D}(A_{P}^*) \quad \text{and} \quad A_{U} = A_{P}^*. \tag{4.26}$$

The Koopman and Perron-Frobenius operators are unitary in $L^2_{\mathbb{C}}(\Omega,\nu)$, as $\mathcal{D}\left(A_{U}\right)$ is dense in $L^2_{\mathbb{C}}(\Omega,\nu)$, by Stone's theorem, their infinitesimal generators are therefore skew symmetric for $\langle \cdot\,, \cdot \rangle_{L^2_{\mathbb{C}}(\Omega,\nu)}$ and we have

$$\mathcal{D}\left(A_{P}^*\right) = \mathcal{D}\left(A_{P}\right) \quad \text{and} \quad A_{P}^* = -A_{P}, \tag{4.27}$$

$$\mathcal{D}\left(A_{U}^{*}\right)=\mathcal{D}\left(A_{U}\right) \quad \text{and} \quad A_{U}^{*}=-A_{U}. \tag{4.28}$$

The adjoint has to be understood in the topology of $L^2_{\mathbb{C}}(\Omega,\nu)$. The two next propositions – well known in $L^2_{\mathbb{C}}(\Omega,\nu)$ for invertible measure preserving systems –, summarize (4.26), (4.27) and (4.28).

Proposition 3 (Koopman infinitesimal generator)**.** *The Koopman infinitesimal generator of $(U_t)_t$ in $L^2_{\mathbb{C}}(\Omega,\nu)$ is the unbounded operator, $\left(A_{U},\mathcal{D}(A_{U})\right)$, defined by*

$$\begin{aligned}\mathcal{D}\left(A_{U}\right) &= \left\{f\in L^2_{\mathbb{C}}(\Omega,\nu)\ :\ x\mapsto \partial_{\mathcal{M}(x)}f(x)\ \in L^2_{\mathbb{C}}(\Omega,\nu)\right\} \text{ and}\\ A_{U}f &= \partial_{\mathcal{M}(\cdot)}f(\cdot) \quad \text{for} \quad f\in\mathcal{D}\left(A_{U}\right).\end{aligned}$$

Proposition 4 (Skew symmetry of the generators)**.** *The Koopman infinitesimal generator A_U is skew-symmetric in $L^2_{\mathbb{C}}(\Omega,\nu)$.*

We set now a continuity property of the Koopman infinitesimal generator on $\mathcal{H}$, which is a subspace of $\mathcal{D}(A_{U})$ (Dufée et al., 2023b).

Theorem 6 (Continuity of the Koopman generator on $\mathcal{H}$)**.** *The restriction of the Koopman infinitesimal generator*

$$A_{U}:\left(\mathcal{H}\,,\,\|\cdot\|_{\mathcal{H}}\right)\mapsto\left(L^2_{\mathbb{C}}(\Omega,\nu)\,,\,\|\cdot\|_{L^2_{\mathbb{C}}(\Omega,\nu)}\right)$$

is continuous.

Through the dual expression of the Koopman operator (remark 2), the infinitesimal generator A_U provides a dynamical system specifying the time evolution of the feature maps as, for all $X\in\Omega$, the feature map $k(\,\cdot\,,\Phi_t(X))$ verifies

$$\partial_t\, U_t\, k(\,\cdot\,,\,X)=\partial_t\, k\Big(\,\Phi_t(\cdot)\,,\,X\Big)=A_U\, k\Big(\,\Phi_t(\cdot)\,,\,X\Big)\,. \tag{4.29}$$

We have the following useful differentiation formulae.

Proposition 5 (Differentiation formulae)**.** *For all $f\in\mathcal{H}$ (with $f=E_tg$, $g\in\mathcal{H}_t$) and $X_t\in\Omega_t$, we have*

$$\langle\mathcal{L}_k\circ\partial_t f\ ,\ k(\,\cdot\,,\,X_t)\rangle_{\mathcal{H}}=-\langle f\ ,\ \mathcal{L}_k\circ\partial_t\, k(\,\cdot\,,\,X_t)\rangle_{\mathcal{H}}.$$

As already outlined, the whole phase space Ω and the global embedding RKHS $\mathcal{H}$ defined on it are both completely inaccessible for high dimensional state spaces. Instead of seeking to reconstruct this global RKHS, we will work in the following with time-varying "local" RKHSs built from a small (with respect to the phase space dimension) ensemble of initial conditions. The set of these time-varying spaces forms the RKHS family. In order to express the time evolution of the features maps associated to the RKHS family, we now define an appropriate expression of the Koopman operator on this family.

Derivation of Koopman operator expression in $\mathcal{W}$

To fully specify the Koopman operators in the RKHS family, we rely on the family of extension, restriction mappings $(E_t)_{t\geq 0}$ and $(R_t)_{t\geq 0}$ (4.11, 4.12) relating the "big" encompassing RKHS $\mathcal{H}$ to the family of time-evolving RKHS $\mathcal{H}_t$ and use the Koopman operators $U_t : L^2_{\mathbb{C}}(\Omega, \nu) \mapsto L^2_{\mathbb{C}}(\Omega, \nu)$ for $t \geq 0$. The adjoints P_t will also be very helpful as well as Remark 2 on the dissymmetry of the application of the flow in the global kernel k.

From now on, we note $X_t := \Phi_t(X_0)$ for all $X_0 \in \Omega_0$ with X_t belonging to Ω_t. We define the Koopman operator in the RKHS family by $\mathcal{U}_t := R_t \circ P_t \circ E_0$ for all $t \geq 0$. For all $X_0 \in \Omega_0$, we have

$$\mathcal{U}_t \left[k_0(\,\cdot\,, X_0) \right] \;=\; R_t \circ P_t \left[k(\,\cdot\,, X_0) \right] \;=\; R_t\, k(\,\cdot\,, X_t) \;=\; k_t(\,\cdot\,, X_t)\,, \tag{4.30}$$

where the second equality is due to (4.22), and the third equality holds true from the definition of R_t and the fact that $\ell(t,t) = 1$. The Koopman operator in the RKHS family, $\mathcal{U}_t : \mathcal{H}_0 \mapsto \mathcal{H}_t$, transports the kernel feature maps on the RKHS family by composition with the system's dynamics. It inherits some of the nice properties of the Perron-Frobenius operator defined on the encompassing global RKHS. As shown by the following theorem, proposition 2 remains valid for the family of kernels $(k_t)_{t\geq 0}$ and $\mathcal{U}_t$ is still unitary in the sense of the following theorem.

Theorem 7 (Koopman RKHS isometry)**.** *The Koopman operator on the RKHS family defines an isometry from $\mathcal{H}_0$ to $\mathcal{H}_t$: for all X_0 and $Y_0 \in \Omega_0$,*

$$\Big\langle \mathcal{U}_t\, k_0(\,\cdot\,, Y_0)\;,\; \mathcal{U}_t\, k_0(\,\cdot\,, X_0) \Big\rangle_{\mathcal{H}_t} = \Big\langle k_0(\,\cdot\,, Y_0)\;,\; k_0(\,\cdot\,, X_0) \Big\rangle_{\mathcal{H}_0}\,.$$

The range of $\mathcal{U}_t : \mathcal{H}_0 \mapsto \mathcal{H}_t$ is dense in $\mathcal{H}_t$.

Let us now determine the adjoint of the Koopman operator in the RKHS family. Let $\mathcal{P}_t := R_0 \circ U_t \circ E_t$ for all $t \geq 0$ and let $X_t = \Phi_t(X_0) \in \Omega_t$. With the same arguments as for $\mathcal{U}_t$, we have

$$\mathcal{P}_t \left[k_t(\,\cdot\,, X_t) \right] = k_0(\cdot\,, X_0)\,. \tag{4.31}$$

The mapping $\mathcal{P}_t : \mathcal{H}_t \mapsto \mathcal{H}_0$ with $t \geq 0$ constitutes the Perron-Frobenius family of operators in the RKHS family. The mapping $\mathcal{P}_t : \mathcal{H}_t \mapsto \mathcal{H}_0$ is unitary for the RKHS family topology (isometry from $\mathcal{H}_t$ to $\mathcal{H}_0$ and the range of $\mathcal{P}_t$ is dense in $\mathcal{H}_0$). The next proposition justifies that $\mathcal{U}_t$ and $\mathcal{P}_t$ have inverse actions on the feature maps.

Proposition 6 (Koopman Perron-Frobenius duality)**.** *For all $X_0 \in \Omega_0$ and $Y_t = \Phi_t(Y_0) \in \Omega_t$, we have*

$$\langle \mathcal{U}_t\, k_0(\,\cdot\,, X_0)\,,\; k_t(\,\cdot\,, Y_t) \rangle_{\mathcal{H}_t} = \langle\, k_0(\,\cdot\,, X_0)\,,\; \mathcal{P}_t\, [k_t(\,\cdot\,, Y_t)] \,\rangle_{\mathcal{H}_0}\,.$$

In order to derive the Koopman and Perron-Frobenius operators' spectral representation in the RKHS family, we exhibit now two propagation operators that will allow us to express the evolution of the feature maps in the RKHS family.

The RKHS family spectral representation

We specify hereafter a family of operators $(A_{U,t})_{t\geq 0}$ related to the Koopman infinitesimal generator A_U, and that give rise to an evolution equation on Ω_t akin to (4.29). They will play, in that sense, the role of infinitesimal generators on the RKHS family.

For all $t \geq 0$, let $A_{U,t}$ be defined by $A_{U,t} := R_t \circ A_U \circ E_t$ with $E_t : \mathcal{H}_t \mapsto \mathcal{H}$ and $R_t : L^2_{\mathbb{C}}(\Omega, \nu) \mapsto L^2_{\mathbb{C}}(\Omega_t, \nu)$.

Proposition 7 (Continuity of $A_{U,t}$)**.** *The mapping $A_{U,t} : (\mathcal{H}_t, \|\cdot\|_{\mathcal{H}_t}) \mapsto (L^2_{\mathbb{C}}(\Omega_t, \nu), \|\cdot\|_{L^2_{\mathbb{C}}(\Omega_t,\nu)})$ is well-defined and continuous.*

In a very similar way as the infinitesimal generator on $\mathcal{H}$, the operator $A_{U,t}$ can be understood as an evolution equation of the feature maps defined on Ω_t and associated to $\mathcal{H}_t$. For $X_t \in \Omega_t$, by proposition 3, it can be noticed that

$$A_{U,t}\; k_t(\,\cdot\,, X_t) = R_t \circ A_U\, [k(\,\cdot\,, X_t)] = R_t \left[\partial_{\mathcal{M}(\cdot)}\, k(\,\cdot\,, X_t) \right]$$

and, in particular on Ω_t,

$$A_{U,t}\; k_t(\,\cdot\,, X_t) = \partial_{\mathcal{M}(\cdot)}\, k_t(\,\cdot\,, X_t). \tag{4.32}$$

Remark 4. *Operator $A_{U,t}$ is alike the Koopman infinitesimal generator A_U. As a matter of fact, by proposition 3 and theorem 6, we have, for all $X \in \Omega$,*

$$A_U k(\cdot, X) = \partial_{\mathcal{M}(\cdot)} k(\cdot, X),$$

and for all $X_t \in \Omega_t$,

$$A_{U,t}\, k_t(\cdot, X_t) = \partial_{\mathcal{M}(\cdot)} k_t(\cdot, X_t).$$

Notice that these two operators are different and act on different domains.

Remark 5. *Operators $A_{U,t}$ can be understood as an evolution operator for the feature maps. As a matter of fact through (4.29) and the definition of $A_{U,t}$, we have, for all $X \in \Omega_{t_0}$*

$$\partial_t\, k\Big(\cdot, \Phi_t(X)\Big)_{|_{\Omega_t}} = A_{U,t}\, k_t\Big(\cdot, \Phi_t(X)\Big)\,. \tag{4.33}$$

Through the above remark, the operator $A_{U,t}$ inherits the properties of operator A_U defined on the global encompassing RKHS $\mathcal{H}$. As shown in the next proposition, it remains in particular skew-symmetric.

Proposition 8 (Skew-symmetry of $A_{U,t}$)**.** *The operator $A_{U,t}$ is skew-symmetric in $L^2_{\mathbb{C}}(\Omega_t, \nu)$: for all f and $g \in \mathcal{H}_t$*

$$\langle A_{U,t}\, f\ ,\ g\rangle_{L^2_{\mathbb{C}}(\Omega_t,\nu)} \ = \ -\,\langle f\ ,\ A_{U,t}\, g\rangle_{L^2_{\mathbb{C}}(\Omega_t,\nu)}.$$

Everything is now set to prove the RKHS family spectral representation theorem, which states that the bounded operator $A_{U,t} : \mathcal{H}_t \mapsto L^2_{\mathbb{C}}(\Omega_t, \nu)$ is diagonalizable for all $t \geq 0$.

Scheme of proof of Theorem 5 [The RKHS family spectral representation]

The full proof, organized in two main steps, is thoroughly detailed in Dufée et al. (2023b). For this proof, we consider the infinitesimal generator $A_U : \mathcal{H} \mapsto L^2_{\mathbb{C}}(\Omega, \nu)$, which is connected to each $A_{U,t}$ for $t \geq 0$ through the restriction operator R_t. In the first step, the diagonalization of the operator $A_U : \mathcal{H} \mapsto L^2_{\mathbb{C}}(\Omega, \nu)$ in $L^2_{\mathbb{C}}(\Omega, \nu)$ is first performed. To that end, we introduce an intermediate (approximating) operator denoted $\widetilde{A}_U$, directly related to A_U and whose inverse is shown to be compact and self-adjoint. The second step of the proof consists in deducing the diagonalization of each $A_{U,t}$ from the diagonalization of A_U obtained at step 1.

4.3.2 Tangent linear dynamics

A result of practical interest concerns the establishment of a rigorous ensemble expression of the tangent linear dynamics operator. Recall that we note $X_t = \Phi_t(X_0)$ for all $t \geq 0$. We define $\delta X(t,\cdot) := \Phi_t[X_0(\cdot) + \delta X_0(\cdot)] - \Phi_t[X_0(\cdot)]$ for all $t \geq 0$ and almost all $x \in \Omega_x$, where $\delta X_0(x)$ is a perturbation of the initial condition at point x. The function $\delta X(t,\cdot) \in L^2(\Omega_x, \mathbb{R}^d)$ is the perturbation of the flow at time t with respect to the initial condition X_0. We have, for almost any $x \in \Omega_x$,

$$\begin{aligned}\Phi_t(X_0 + \delta X_0) - \Phi_t(X_0) &= \int_0^t \mathcal{M}[\Phi_s(X_0 + \delta X_0)] - \mathcal{M}[\Phi_s(X_0)]\,\mathrm{d}s + o(\delta X_0) \\ &= \int_0^t \mathrm{d}(\mathcal{M} \circ \Phi_s(X_0))\delta X_0\,\mathrm{d}s + o(\delta X_0)\end{aligned}$$

and we obtain also that

$$\begin{aligned}\Phi_t(X_0 + \delta X_0) - \Phi_t(X_0) &= \int_0^t \mathrm{d}\mathcal{M}[\Phi_s(X_0)]\ [\mathrm{d}\Phi_s(X_0)\delta X_0]\,\mathrm{d}s + o(\delta X_0) \\ &= \int_0^t \partial_X \mathcal{M}[\Phi_s(X_0)]\ \delta X(s,\cdot)\,\mathrm{d}s + o(\delta X_0).\end{aligned}$$

The variation of the flow verifies $\delta X(t,\cdot) = \int_0^t \partial_X \mathcal{M}[X_s]\delta X(s,\cdot)\,\mathrm{d}s$ almost everywhere on Ω_x. Recall that $\mathcal{M}$ is assumed to be C^1, in particular $\sup_{X\in\Omega} \partial_X \mathcal{M}(X) < \infty$ (since Ω is compact). The function $\delta X(t,\cdot)$ belongs to $L^2(\Omega_x, \mathbb{R}^d)$ and verifies, for all $t \geq 0$,

$$\partial_t\, \delta X(t,\cdot) \;=\; \partial_X \mathcal{M}[X_t]\delta X(t,\cdot). \tag{4.34}$$

Each component $[\delta X(t,\cdot)]_i$ of $\delta X(t,\cdot)$ belongs to $L^2(\Omega_x, \mathbb{R})$. For all $x \in \Omega_x$, let us define a map $g_x : \Omega \to \mathbb{R}^d$ such that, for all $X \in \Omega$,

$$g_x(X) = X(x).$$

Then, for $1 \leq i \leq d$, the map $[g_x(\delta X_t)]_i : \Omega \mapsto \mathbb{C}$ defined for all $x \in \Omega_x$ by

$$[g_x(\delta X_t)]_i := [\delta X_t(x)]_i$$

is an element of $L^2(\Omega, \nu)$. For any $x \in \Omega_x$, the quantity $g_x(X_t)$ corresponds to a vector of

observables of the dynamical system and we may immediately write for all $t \geq 0$

$$\partial_t \, g_x(\delta X_t) \;=\; \partial_X \mathcal{M}(X_t) g_x(\delta X_t). \tag{4.35}$$

We are now ready to exhibit a kernel expression of the tangent linear operator $\partial_X \mathcal{M}(X_t)$. For all $1 \leq i \leq d$, the function $\mathcal{L}_k \, [g_x]_i \in \mathcal{H}$ verifies, for all $\delta X_t \in \Omega$,

$$\langle \mathcal{L}_k \circ \partial_t \, \mathcal{L}_k \, [g_x]_i \; , \; k(\,\cdot\,, \, \delta X_t) \rangle_{\mathcal{H}} = \mathcal{L}_k \; (\partial_t \, \mathcal{L}_k \, [g_x]_i)(\delta X_t). \tag{4.36}$$

Upon applying the differentiation formulae of proposition 5 and the evolution equation (4.29), we have

$$\begin{aligned}
\langle \mathcal{L}_k \circ \partial_t \, \mathcal{L}_k \, [g_x]_i \, , \, k(\,\cdot\,, \, \delta X_t) \, \rangle_{\mathcal{H}} &= -\langle \mathcal{L}_k [g_x]_i \, , \, \mathcal{L}_k \circ \partial_t \, k(\,\cdot\,, \, \delta X_t) \rangle_{\mathcal{H}} \\
&= -\langle \mathcal{L}_k [g_x]_i \, , \, \mathcal{L}_k \, [A_{_U} \, k(\,\cdot\,, \, \delta X_t] \, \rangle_{\mathcal{H}} \\
&= -\langle \, \mathcal{L}_k^{1/2} \, [g_x]_i \; , \; \mathcal{L}_k^{1/2} \, [A_{_U} \, k(\,\cdot\,, \, \delta X_t] \, \rangle_{L^2_{\mathbb{C}}(\Omega, \nu)}.
\end{aligned}$$

The operator $\mathcal{L}_k^{1/2}$ being self-adjoint in $L^2_{\mathbb{C}}(\Omega, \nu)$ and $A_{_U}$ being skew-symmetric for the inner product of $L^2_{\mathbb{C}}(\Omega, \nu)$, we have

$$\begin{aligned}
\langle \mathcal{L}_k \circ \partial_t \, \mathcal{L}_k \, [g_x]_i \; , \; k(\,\cdot\,, \, \delta X_t) \, \rangle_{\mathcal{H}} &= \langle A_{_U} \, \mathcal{L}_k [g_x]_i \; , \; k(\,\cdot\,, \, \delta X_t) \, \rangle_{L^2_{\mathbb{C}}(\Omega, \nu)} \\
&= \langle \mathcal{L}_k^{1/2} \, A_{_U} \, \mathcal{L}_k [g_x]_i \; , \; \mathcal{L}_k^{-1/2} \, k(\,\cdot\,, \, \delta X_t) \, \rangle_{L^2_{\mathbb{C}}(\Omega, \nu)} \\
&= \langle \mathcal{L}_k \, A_{_U} \, \mathcal{L}_k [g_x]_i \; , \; k(\,\cdot\,, \, \delta X_t) \, \rangle_{\mathcal{H}} \\
&= \mathcal{L}_k \, A_{_U} \, \mathcal{L}_k \, [g_x]_i (\delta X_t).
\end{aligned}$$

Combining the right-hand side of the above expression with (4.36), we have

$$\mathcal{L}_k(\, \partial_t \, \mathcal{L}_k [g_x]_i \,)(\delta X_t) = \mathcal{L}_k \, (A_{_U} \, \mathcal{L}_k [g_x]_i \,)(\delta X_t). \tag{4.37}$$

As $\mathcal{L}_k$ is injective, the kernel of the tangent linear operator in the RKHS $\mathcal{H}$ reads:

$$(\partial_t \, \mathcal{L}_k [g_x]_i) \, (\delta X_t) = A_{_U} \, \mathcal{L}_k \, [g_x]_i (\delta X_t).$$

As Ω is bounded, the function $t \mapsto \delta X(t, \cdot)$ belongs to $L^\infty([0, T])$ and therefore $\delta X_t \mapsto g_x(\delta X_t)$ belongs to $L^\infty\Big([0, T], L^2(\Omega, \nu)\Big)$. By (4.35), we have also that $\delta X_t \mapsto \partial_t \, g_x(\delta X_t)$

belongs to $L^\infty\Big([0,T], L^2(\Omega,\nu)\Big)$ since $\sup_{Y\in\Omega} \partial_X \mathcal{M}(Y) < \infty$. We have therefore on Ω

$$\partial_t \, \mathcal{L}_k[g_x]_i(\delta X_t) \;=\; \int_\Omega k(\delta X_t \,,\, z)\; \partial_t[g_x]_i(z)\, \nu(\mathrm{d}z) \;=\mathcal{L}_k \left(\partial_t \, [g_x]_i \,\right)(\delta X_t),$$

and thus

$$\partial_t \, [g_x]_i(\delta X_t) \;=\; \mathcal{L}_k^{-1}\, A_{_U}\, \mathcal{L}_k \, [g_x]_i(\delta X_t).$$

By the commutation property of remark 3, together with (4.35) and as $\mathcal{L}_k^{1/2}\,[g_x]_i \in \mathcal{H}$, the following equalities hold for all $\delta X_t \in \Omega$:

$$\partial_t \, [g_x]_i(\delta X_t) \;=\; \mathcal{L}_k^{-1/2}\, A_{_U}\, \mathcal{L}_k^{1/2} \, [g_x]_i(\delta X_t),$$

$$\partial_X \mathcal{M}(X_t)\, g_x\,(\delta X_t) = \mathcal{L}_k^{-1/2}\, A_{_U}\, \mathcal{L}_k^{1/2}\, g_x(\delta X_t). \tag{4.38}$$

Let us specify now the kernel expression of the tangent linear dynamics in Ω_s for all $s \geq 0$. The function $\mathcal{L}_k^{1/2}\,[g_x]|_{\Omega_s}$ belongs to $\mathcal{H}_s$. We have as well $\mathcal{L}_k^{1/2}(\partial_t g_x) = \mathcal{L}_{k_s}^{1/2}(\partial_t g_x)$ on Ω_s. For all Y_s and $X_s' \in \Omega_s$, we have that $A_{_U}\, k(\,\cdot\,, X_s')(Y_s) = A_{U,s}\, k_s(\,\cdot\,, X_s')(Y_s)$. In particular, we get $A_{_U}\, \mathcal{L}_k^{1/2}\,[g_x](Y_s) = A_{U,s}\, \mathcal{L}_{k_s}^{1/2}\,[g_x](Y_s)$. We obtain hence the kernel expression for all $Y_s \in \Omega_s$:

$$\partial_X \mathcal{M}(X_s)\, g_x\,(Y_s) = \mathcal{L}_{k_s}^{-1/2} A_{U,\,s}\, \mathcal{L}_{k_s}^{1/2}\, g_x\,(Y_s). \tag{4.39}$$

Note that the domains of the kernel expressions of the tangent linear (4.38) and (4.39) are different. The right-hand side of (4.39) provides a convenient kernel expression of the tangent linear operator, enabling us to evaluate the tangent linear dynamics from an ensemble of feature maps. The adjoint of the tangent linear dynamics is straightforwardly given by

$$\partial_X \mathcal{M}^*(X_s) g_x\,(Y_s) = -\mathcal{L}_{k_s}^{-1/2} A_{U,\,s}\, \mathcal{L}_{k_s}^{1/2} g_x\,(Y_s). \tag{4.40}$$

Remark 6 (Projection observables)**.** *The point observable functions g_x used above can be extended to other functions defined from a basis $(\Psi_j)_{j\geq 0}$ of $L^2(\Omega_x, \mathbb{R}^d)$, with*

$$g_x^\psi X_t := \sum_j^\infty \langle X_t \,,\, \Psi_j \rangle_{L^2(\Omega_x)} \Psi_j(x),$$

$$g_x^\psi : \Omega \mapsto \sum_j^\infty \langle \,\cdot\,, \Psi_j \rangle_{L^2(\Omega_x)} \Psi_j(x) \in \mathbb{C}.$$

Remark 7 (Dependence on X_t)**.** *It should be noted that in* (4.39) *and* (4.40) *the right-hand side does not depend on X_s whereas the tangent linear operator or its adjoint on the left-hand side does. Function X_s indicates around which function of Ω_s the nonlinear system is linearized. The infinitesimal generator of the Koopman operator, representing the dynamics' linear tangent operator on the RKHS family, depends necessarily also on this function. This dependence is here implicit and induced by the considered sampled functions used to define the RKHS $\mathcal{H}_t$. If the set of members $(X_t^{(i)})_{i\geq 1}$ are centered around a particular function X_s, the infinitesimal generators of the Koopman operator can be interpreted as a representation of the tangent linear operator around function X_s. In ensemble methods, X_s is in general taken as the ensemble mean, and $(X_s^{(i)})_{i\geq 1}$ is an ensemble of time-dependent perturbations around this mean.*

In the perspective of evaluating the tangent linear dynamics, the following remark provides an even more convenient expression than (4.39).

Remark 8 (More regularity on δX_t)**.** *For all $1 \leq i \leq d$, if we suppose that the function $[g_x]_i$ belongs to $\mathcal{H} \subset L^2_{\mathbb{C}}(\Omega, \nu)$, the proof can be simplified and* (4.39) *is replaced by*

$$\partial_X \mathcal{M}(X_s)\, g_x\,(\delta X_s) = A_{U,\,s}\, g_x\,(\delta X_s). \tag{4.41}$$

It can be pointed out that the expression above corresponds to the approximation of the tangent linear dynamics used in ensemble methods if we work in a finite dimensional space such that $\Omega \subset \mathbb{R}^n$ and assume that $k_t(\cdot, X_t)$ is defined as $(N-1)^{-1/2}\langle (X_t - \overline{X}_t), \cdot\, - \overline{X}_t\rangle_{\mathbb{R}^n}$, with $\overline{X}_t$ the empirical ensemble mean. With that definition, we have $A_{U,t}k_t(\cdot, X^{(i)}) = \sum_j \partial_{X_j} k_t(\cdot, X_t^{(i)})\mathcal{M}(\cdot)_j = \langle (\delta_{X_t^{(i)}} - \frac{1}{N}\sum_\ell \delta_{X_t^{(\ell)}})\mathcal{M}(\cdot), \cdot - \overline{X}_t\rangle_{\mathbb{R}^n}$ that reads $(N-1)^{-1/2}\langle \mathcal{M}(X_t^{(i)}) - \frac{1}{N}\sum_j \mathcal{M}(X_t^{(j)}), \cdot - \overline{X}_t\rangle_{\mathbb{R}^n}$, for $i = 1, \ldots, N$ and for which, when associated to the $\mathbb{R}^n$ Euclidean inner product on a resolution grid of size n, the left-hand side of this latter expression corresponds to the so-called $(N \times n)$ anomaly matrix built from N ensemble members of the dynamical system. The tangent linear approximation provided by ensemble methods can be thus immediately interpreted as a particular instance of feature maps together with a given choice of specific inner product to define the reproducing kernel. Keeping a finite dimensional approximation but working without assuming that the functions $[g_x]$ belong to $\mathcal{H}$, and thus with now expression (4.38) for the ensemble tangent linear expression, corresponds to the case in which a localization procedure identified to the square root operator $\mathcal{L}_{k_t}^{1/2}$ has been considered. These two choices embed the problem within a particular RKHS family of functions. The relation between

the tangent linear dynamics and the anomaly matrix is in our case exact and does not correspond to a finite difference approximation as classically presented in ensemble methods. The RKHS family can be seen as a way of linearizing locally a nonlinear system in a convenient sequence of spaces of smooth functions.

4.3.3 Finite time Lyapunov exponents

The kernel of the Koopman operator provides also a direct access to the finite time Lyapunov exponents. Recalling from (4.35) that for any punctual observable g_x, as defined previously, we have

$$g_x(\delta X_t) \ = \int_0^t \partial_X \mathcal{M}(X_s) g_x(\delta X_s) \mathrm{d}s + o(\delta X_0). \tag{4.42}$$

With the expression of the tangent linear operator in terms of the Koopman infinitesimal generator (4.38) on Ω_s, we have then, at first order,

$$g_x(\delta X_t) = \int_0^t \mathcal{L}_{k_s}^{-1/2} A_{U,s} \, \mathcal{L}_{k_s}^{1/2} g_x(\delta X_s) \, \mathrm{d}s. \tag{4.43}$$

For all $s \in [0,t]$, we consider a perturbation $g_x(\delta X_s) = \mathcal{L}_{k_s}^{-1/2} \psi_\ell^s(\delta X_s)$ along a Koopman generator eigenfunction associated to the eigenvalue of maximal modulus $|\lambda_\ell|$. By Theorem 5, we have

$$\partial_t g_x(\delta X_s) = \lambda_\ell \, \mathcal{L}_{k_s}^{-1/2} \psi_\ell^s = \lambda_\ell \, g_x(\delta X_s).$$

Therefore we get

$$|g_x(\delta X_t)| = e^{|\lambda_\ell| t} \, |g_x(\delta X_0)|, \tag{4.44}$$

and the finite time Lyapunov exponent is consequently defined as

$$\sigma = |\lambda_\ell|. \tag{4.45}$$

For regular perturbations $g_x \in \mathcal{H}_s$ with unitary perturbation $g_x(\delta X_s) = \psi_\ell^s(\delta X_s)$, the derivation is even simpler as we obtain from remark 8

$$\partial_t g_x(\delta X_s) \ = A_{U,t} \, g_x(\delta X_s),$$

which yields directly to expression (4.44) and to the same expression for the Lyapunov exponent.

The modulus of the larger Koopman eigenvalue in the RKHS family provides thus an estimate of the Lyapunov exponent. It can be outlined that the computation of Lyapunov exponents for large scale systems is computationally very demanding as it requires the construction of the linear tangent dynamics operator and the solution of an eigenvalue problem of very big dimension. The construction of the exact numerical tangent linear operator is in general a tedious task when expressed in $L^2(\Omega_x)$, as in equation (4.34). The ensemble-based method provided by our formalism is on the contrary very simple by expressing the tangent linear operator in $L^2_{\mathbb{C}}(\Omega_t, \nu)$, as in equation (4.35). It can be noticed that, by this change of norm in the definition, the computed values are not the same.

Three distinct values can then be defined for practical computations. First, the Lyapunov spectrum expressed in $L^2_{\mathbb{C}}(\Omega_t, \nu)$ can be determined by computing the singular values of $\mathcal{L}_{k_t}^{-1/2} A_{U,t}\, \mathcal{L}_{k_t}^{1/2}$. The time integral is dropped since the evaluations are constant along trajectories. It can be viewed as an advantage of working in $L^2_{\mathbb{C}}(\Omega_t, \nu)$ instead of $L^2(\Omega_x)$ as performed classically. The time independence is due to the fact that the Koopman operator is intrinsic to the dynamical system. However, from a numerical point of view, as the computation is performed in practice through an ensemble with a limited number of members, the learned spectrum is representative only of the local dynamics at the time ($t = t_0$) at which the kernel has been evaluated. As an alternative, modal Lyapunov exponents can be defined by the square root of the first singular values of $\mathcal{L}_{k_t}^{-1/2} \psi_\ell |\lambda_\ell|^2 \psi_\ell^* K_t^{-1} \mathcal{L}_{k_t}^{1/2} = \mathcal{L}_{k_t}^{-1/2} \psi_\ell |\lambda_\ell|^2 \psi_\ell^* \mathcal{L}_{k_t}^{-1/2}$ with $K_t(i,j) = k_t(X_t^{(i)}, X_t^{(j)})$. We call these singular values the *Koopman modal Lyapunov exponents* (KMLE). Finally, as a third option, equation (4.45) can simply be considered to evaluate modal exponents. We can notice that the two modal Lyapunov exponents definitions are very similar; the former being expressed in $L^2_{\mathbb{C}}(\Omega_t, \nu)$ and the latter in $\mathcal{H}_t$.

4.3.4 Practical considerations

Let us stress again that in practice, we only have access to the mappings k_t and $A_{U,t}$ with $t \geq 0$. The mapping k and A_U are completely inaccessible for high-dimensional systems, as they require the complete knowledge of the phase space or at least of a long enough orbit with a density assumption in the whole phase space. This last assumption is associated to strong requirements of the dynamical system and is not necessarily valid for a given time series of a particular observable. Instead of working with an infinite (dense) trajectory, Theorem 5 enables us to estimate the eigenvalues and eigenfunctions

of the Koopman operator locally in the RKHS family, which can locally conveniently be accessed from an ensemble of finite time trajectories. As it will be described in the following, operators $A_{U,t}$, can be discretized as an ensemble matrix – itself related, as we saw it previously, to the tangent linear dynamics operator. This matrix is then diagonalized to get access to Koopman eigenvalues and their associated eigenfunctions. In theory, the diagonalization of $A_{U,t}$ needs to be performed only once, at a given time, to access the Koopman eigenpairs $(\psi_\ell^t)_\ell$ and $(\lambda_\ell)_\ell$. However, the exponential relation between distinct instants allows us also to consider averaging strategies to eventually robustify the estimation in practice.

Diagonalization in practice

For all $t \geq 0$, let m_t be the kernel expression of the operator $A_{U,t}$ given by $m_t(X_t\,,\,Y_t) := A_{U,t}^*\Big[\,k_t(\,\cdot\,,\,Y_t)\,\Big](X_t)$ for all X_t and $Y_t \in \Omega_t$. By Proposition 8, we have $m_t(X_t\,,\,Y_t) = -A_{U,t}\Big[\,k_t(\,\cdot\,,\,Y_t)\,\Big](X_t)$.

Let us denote by $\{X_t^{(i)} \,:\, 1 \leq i \leq N\}$ an ensemble of members generated by the dynamical system and by $\{k_t(\,\cdot\,,\,X_t^{(i)}) \,:\, 1 \leq i \leq N\}$ the N associated feature maps. For all $t \geq 0$, these N feature maps enable us to build a kernel expression of the operator $A_{U,t}$ as the $N \times N$ matrix $\mathbb{M}_t = \Big(\,m_t(X_t^{(i)}\,,\,X_t^{(j)})\,\Big)_{1\leq i,j\leq N}$ with:

$$(\mathbb{M}_t)_{ij} \;:=\; -A_{U,t}\Big[\,k_t(\cdot\,,\,X_t^{(j)})\,\Big]\,(X_t^{(i)}) = \Big[-\partial_{\mathcal{M}(\cdot)}\,k_t(\,\cdot\,,\,X_t^{(j)})\Big]\;(X_t^{(i)}).$$

As shown in the following, this matrix enables us to access to the Koopman generator eigenvalues and to the evaluation of the eigenfunctions at the ensemble members.

By definition of $A_{U,t}$, we have $A_{U,t}[\,k_t(\cdot\,,\,X_t^{(j)})\,] = A_{U}[\,k(\cdot\,,\,X_t^{(j)})\,]$. Moreover, as $-\mathcal{L}_k^{1/2}\circ A_{U}[\,k(\,\cdot\,,\,X_t^{(j)})\,]$ belongs to $\mathcal{H}$, by the proof of Theorem 5 (Dufée et al., 2023b) and denoting $(\psi_\ell)_\ell$ an orthonormal basis of $\mathcal{H}$ set from the eigenfunctions of A_{U}, we get

$$\begin{aligned}
-\mathcal{L}_k^{1/2}\circ A_{U}\Big[\,k(\cdot\,,\,X_t^{(j)})\,\Big](X_t^{(i)}) &= -\sum_{\ell=0}^{N}\langle \mathcal{L}_k^{1/2}\circ A_{U}\,\Big[\,k(\cdot\,,\,X_t^{(j)})\,\Big]\,,\,\psi_\ell\,\rangle_{\mathcal{H}}\;\psi_\ell,\\
&= -\sum_{\ell=0}^{N}\langle A_{U}\,k(\,\cdot\,,\,X_t^{(j)})\;,\;\mathcal{L}_k^{-1/2}\,\psi_\ell\,\rangle_{L^2_{\mathbb{C}}(\Omega,\nu)}\;\;\psi_\ell.
\end{aligned}$$

Denoting β_ℓ the eigenvalues of $\mathcal{L}_k$, we have also $\mathcal{L}_k j(\psi_\ell) = \beta_\ell\, j(\psi_\ell)$ and

$$-\mathcal{L}_k^{1/2} \circ A_{_U}\Big[\, k(\boldsymbol{\cdot}\,,\, X_t^{(j)})\,\Big] = -\sum_{\ell=0}^{N} \beta_\ell^{-1/2}\, \langle A_{_U}\, k(\,\boldsymbol{\cdot}\,,\, X_t^{(j)})\;,\; j(\psi_\ell)\,\rangle_{L^2_{\mathbb{C}}(\Omega,\nu)}\;\; \psi_\ell.$$

By the skew symmetry of the generators (proposition 4), we have that

$$\begin{aligned}-\mathcal{L}_k^{1/2} \circ A_{_U}\Big[\, k(\boldsymbol{\cdot}\,,\, X_t^{(j)})\,\Big] &= \sum_{\ell=0}^{N} \lambda_\ell\, \beta_\ell^{-1/2}\, \langle k(\boldsymbol{\cdot}\,,\, X_t^{(j)})\,,\, j(\psi_\ell)\,\rangle_{L^2_{\mathbb{C}}(\Omega,\nu)}\;\; \psi_\ell \\ &= \sum_{\ell=0}^{N} \lambda_\ell\, \beta_\ell^{-1/2}\, \overline{\mathcal{L}_k j(\psi_\ell)(X_t^{(j)})}\; \psi_\ell,\end{aligned}$$

which leads to

$$-\mathcal{L}_k^{1/2} \circ A_{_U}\Big[\, k(\boldsymbol{\cdot}\,,\, X_t^{(j)})\,\Big] = \sum_{\ell=0}^{N} \lambda_\ell\, \beta_\ell^{1/2} \overline{\psi_\ell(X_t^{(j)})}\; \psi_\ell,$$

and, upon applying $\mathcal{L}_k^{-1/2}$ on both sides, we get

$$-A_{_U}\Big[\, k(\boldsymbol{\cdot}\,,\, X_t^{(j)})\,\Big] = \sum_{\ell=0}^{N} \lambda_\ell\, \beta_\ell^{1/2}\, \overline{\psi_\ell(X_t^{(j)})}\;\; \mathcal{L}_k^{-1/2}\psi_\ell = \sum_{\ell=0}^{N} \lambda_\ell\, \overline{\psi_\ell(X_t^{(j)})}\; j(\psi_\ell).$$

By the restriction expression, we finally obtain the following equality for all $t \geq 0$:

$$m_t(X_t^{(i)}\,,\, X_t^{(j)}) = \sum_{\ell=0}^{N} \lambda_\ell\; \psi_\ell^t(X_t^{(i)})\; \overline{\psi_\ell^t(X_t^{(j)})},$$

which shows that the diagonalization of $\mathbb{M}_t$ provides a set of the Koopman generator eigenpairs in the RKHS family.

In practice, the skew-symmetric matrix $\mathbb{M}_t := \Big(m_t(X_t^{(i)}\,,\, X_t^{(j)})\Big)_{1\leq i,j\leq N}$ is assembled from the definition of $A_{U,t}$ and a given choice of the kernel. As explained in the previous section, this matrix corresponds to a kernel expression of $A_{U,t}$ with

$$\mathbb{M}_t(X_t^{(i)}\,,\, X_t^{(j)}) \;:=\; -A_{U,t}\Big[\, k_t(\boldsymbol{\cdot}\,,\, X_t^{(j)})\,\Big](X_t^{(i)}) = \left[-\partial_{\mathcal{M}(.)}\, k_t(\,\boldsymbol{\cdot}\,,\, X_t^{(j)})\right](X_t^{(i)}).$$

This matrix can be interpreted in the RKHS setting as resulting from the matrix multiplication:

$$\mathbb{M}_t(X_t^{(i)}\,,\, X_t^{(j)}) \;:=\; \sum_{\ell} \mathbb{F}_t(j,\ell) K_t(\ell, i),$$

with $\mathbb{F}_t(i,\ell) = -A_{U,t}\Big[\, k_t(Z^{(\ell)}\,,\, X_t^{(j)})\Big]$. This indeed corresponds to a discretization of the kernel expression of operator $A_{U,t}$ through the empirical Dirac measure. Numerically, instead of working with matrix $\mathbb{M}_t$ (i.e. the evaluation of $A_{U,t}k_t(\cdot, X)$ at several discrete points), we will work directly with matrix $\mathbb{F}_t = \mathbb{M}_t K_t^{-1}$. This has the advantage of directly working with an implicit discretization of operator $A_{U,t}$, and to relax somehow its dependency on the kernel choice. The skew-symmetric matrix $\mathbb{M}_t K_t^{-1}$ is then diagonalized through a direct numerical procedure (using LAPACK library and working numerically on the anti-symmetric part of $\mathbb{M}_t K_t^{-1}$) and can be written $\mathbb{M}_t K_t^{-1} = V_t \Lambda V_t^T$, with V_t a unitary matrix and Λ a diagonal matrix. The matrix V_t gathers eigenvectors V_j^t of $\mathbb{F}$, which is a discretization of $A_{U,t}k_t(\cdot, X)$ giving access to the values of the Koopman eigenfunctions $\psi_j^t = \mathcal{L}_k^{1/2} V_j^t$ at the n ensemble members points $X_t^{(i)}$. The matrix Λ is composed of Koopman eigenvalues with conjugate pairs of pure imaginary eigenvalues.

As previously mentioned, this diagonalization can be performed at a single time or at several instants accompanied with an averaging procedure. Theorem 5 and (4.8), give access to the eigenvectors evaluation along a trajectory for all time instants.

We provide below, as examples, expressions of the evaluation of $\mathbb{M}_t(X_t^{(i)}, X_t^{(j)}) = \left(-\partial_{\mathcal{M}(\cdot)}\, k_t(\cdot, X_t^{(j)})\right)\left(X_t^{(i)}\right)$ for the empirical covariance kernel (k_E) and the Gaussian kernel (k_G). The empirical covariance kernel is defined through the kernel isometry property as

$$k_t(X_t^{(i)}, X_t^{(j)}) = k_E(X_0^{(i)}, X_0^{(j)}) = \left\langle X_0^{(i)}, X_0^{(j)} \right\rangle_{\Omega_x}, \tag{4.46}$$

with $(\cdot,\cdot)_{\Omega_x}$ the inner product of $L^2(\Omega_x, \mathbb{R}^d)$ and where Ω_x denotes the physical domain of the considered dynamics. We obtain

$$\mathbb{M}_t(X_t^{(i)}, X_t^{(j)}) = \left\langle \frac{\partial}{\partial t} X_0^{(j)}, X_0^{(i)} \right\rangle_{\Omega_x}. \tag{4.47}$$

In this expression, we see that the time derivative of the ensemble members at the initial time is required. Similarly, the Gaussian kernel is defined as

$$k_t(X_t^{(i)}, X_t^{(j)}) = k_G(X_0^{(i)}, X_0^{(j)}) = \exp\left(-\frac{1}{{\ell_G}^2}\left\|X_0^{(i)} - X_0^{(j)}\right\|_{\Omega_x}^2\right). \tag{4.48}$$

This leads to

$$\mathbb{M}_t(X_t^{(i)}, X_t^{(j)}) = -\frac{2}{{\ell_G}^2}\left\langle \frac{\partial}{\partial t}X_0^{(j)}, \left(X_0^{(i)} - X_0^{(j)}\right)\right\rangle_{\Omega_x} \exp\left(-\frac{1}{{\ell_G}^2}\left\|X_0^{(i)} - X_0^{(j)}\right\|_{\Omega_x}^2\right). \quad (4.49)$$

Thanks to the isometry property, this matrix needs to be evaluated only at a single time.

4.4 Data assimilation for ensemble forecast in RKHS

This subsection aims at providing data assimilation techniques enriched with the RKHS structure in their formulations.

On a time horizon T, related to the Lyapunov exponents of the system, we are given the ensemble of trajectories $(X_t^{(i)})_{i=1,...,N;t=0,...,T}$ resulting from the dynamical system (4.3) started from initial conditions $(X_0^{(i)})_{i=1,...,N}$. At discrete times $t_1, ..., t_J \in [0, T]$, we are given observations

$$Y_{t_j} = \mathbf{H}(X_{t_j}) + \varepsilon_j,$$

where $(X_t)_{t=0,...,T}$ is the underlying truth or reference of the system, $\mathbf{H}$ is the observation operator which, for all t, maps Ω_t to an observation-related space $\mathbb{R}^{D_Y}$, and $\varepsilon_j \sim \mathcal{N}(0, \mathbf{R}_{t_j})$, where $\mathbf{R}_{t_j} \in \mathbb{R}^{D_Y \times D_Y}$ is the observation error covariance matrix.
In a classical variational data assimilation framework, one is looking for the estimator $\hat{X} = \sum_{i=1}^N \beta_i X^{(i)}$ minimizing the following cost function:

$$J(X) = \frac{1}{2T}\int_0^T ||\mathbf{H}(X_t) - Y_t||^2_{\mathbf{R}_t^{-1}} \mathrm{d}t + \frac{1}{2}||X - X_b||^2_{\mathbf{B}^{-1}},$$

where $X_b = (X_b^{(1)}, ..., X_b^{(N)}) \in \Omega^N$ is called the background ensemble and $\mathbf{B}$ is the background covariance matrix in Ω. There is a certain degree of freedom for the definition of the penalization term (the second term). The norm could be either computed in the family of phase spaces Ω_t or in the RKHS family built in the previous subsections, with also the choice of the time t at which it is computed (time 0 for the initial condition, time T for the ending point of the trajectory,...) Different leads are explored hereafter.

In the following, the observation covariance matrix is assumed to be constant over time ($\mathbf{R}_t = \mathbf{R}$ for all t). We define the observation space $\mathcal{Y} = (\mathbb{R}^{D_Y})^J$ and denote by $Y = (Y_{t_1}, ..., Y_{t_J}) \in \mathcal{Y}$ the observation discrete trajectory.

4.4.1 Smoothness hypothesis

This subsection provides a simple relation between the superposition principle that is completely justified for the feature maps, and the one applied to the states of the system, which is questionable for the EnKF for nonlinear dynamics. This relation relies on a so called smoothness hypothesis on a particular family of observables.

Let X be an element of the phase space Ω. Then, by the RKHS properties, its feature map $k(\cdot, X)$ can be expanded as a linear combination of the ensemble feature maps. Namely,

$$k(\cdot, X) = \sum_{i=1}^{N} \alpha_i k(\cdot, X^{(i)}). \tag{4.50}$$

Morally, if one thinks of k as a smooth (and signed) version of the Dirac distribution (which is of course not a valid candidate in our framework), the latter equation is precisely the expression of the posterior distribution resulting from a particle filter (1.30) for signed weighths. With this respect, looking for the adequate linear combination of the feature maps can be seen as an extension of particle filtering, for which the superposition principle makes perfect sense. Still, we can notice that, in the particle filter, the linear combination defining δ_X is convex by construction, which is not necessarily the case for the linear combination defining the feature map $k(\cdot, X)$ in the RKHS. Now, the point is to see how the superposition principle, completely legitimate for the feature maps, can translate to the states themselves. This is the object of the rest of this section.

Similarly to Remark 8, the space-wise evaluation functions $g_x : \Omega \to \mathbb{R}^d$ such that $g_x(X) = X(x)$ for all $x \in \Omega_x$ and $X \in \Omega$ will be assumed to have all its components $[g_x]_i$ belonging to the RKHS $\mathcal{H}$ for $1 \leq i \leq d$. For the sake of concision, we will take the shortcut notation that $g_x \in \mathcal{H}$. This hypothesis, called in the following the **smoothness hypothesis**, considerably simplifies this transfer between feature maps and states to see how to establish Kalman-like schemes. An alternative without this assumption will be considered in Section 4.4.4.

Considering the previous hypothesis, we have, for all $x \in \Omega_x$,

$$\begin{aligned} X(x) = g_x(X) &= \langle g_x, k(\cdot, X)\rangle_{\mathcal{H}} \\ &= \left\langle g_x, \sum_{i=1}^{N} \alpha_i k(\cdot, X^{(i)}) \right\rangle_{\mathcal{H}} \\ &= \sum_{i=1}^{N} \alpha_i g_x(X^{(i)}) \\ &= \sum_{i=1}^{N} \alpha_i X^{(i)}(x). \end{aligned}$$

As this holds for all $x \in \Omega_x$, then

$$X = \sum_{i=1}^{N} \alpha_i X^{(i)}, \tag{4.51}$$

which means that the linear combination that defines $k(\cdot, X)$ still holds for X itself. This can be seen a linear interpolation of the kernel. Indeed, with these assumptions we get

$$k\left(\cdot, \sum_{i=1}^{N} \alpha_i X^{(i)}\right) = \sum_{i=1}^{N} \alpha_i k(\cdot, X^{(i)}), \tag{4.52}$$

so k behaves like a right-linear kernel. Moreover, under these hypothesis, the usual euclidian norms that are considered in data assimilation cost functions can be interpreted as norms in the RKHS. Indeed,

$$||X||_{\Omega}^2 = \sum_{x \in \Omega_x} |X(x)|^2 = \sum_{x \in \Omega_x} |\langle g_x, k(\cdot, X)\rangle_{\mathcal{H}}|^2 . \tag{4.53}$$

As a summary, this smoothness assumption transposes the superposition principle, which is completely justified for the feature maps, to the state space, and hence fully justifies the use of Kalman techniques. In short, Kalman filters can be seen as local (in time) linear ensemble interpolations of a given RKHS kernel.

4.4.2 Ensemble Kalman filter for pieces of trajectory in RKHS

The idea of the following modified EnKF is to make the cost function account for the whole set of observations gathered along time, but still looking for the posterior

ensemble members, which are whole pieces of trajectory between times 0 and T, as linear combinations of the prior ones. An usual framework would be to sequentially filter each observation at each time it comes in, resulting in time-varying linear combinations for the posterior ensemble members. The RKHS family spectral representation theorem 5 coupled with the smoothness hypothesis presented above ensures that it is legitimate to consider linear combinations with constant in time coefficients as potential solutions of the system.

A trajectory-wise observation operator is introduced as

$$\begin{aligned}\tilde{\mathbf{H}} : (\Omega_t)_{t\geq 0} &\to \mathcal{Y} \\ X &\mapsto (\mathbf{H}(X_{t_1}), ..., \mathbf{H}(X_{t_J})),\end{aligned}$$

so that

$$Y = \tilde{\mathbf{H}}(X) + \varepsilon,$$

where $\varepsilon \sim \mathcal{N}(0, \tilde{\mathbf{R}})$ with $\tilde{\mathbf{R}} = \mathbf{I}_J \otimes \mathbf{R}$.
Similarly, we define the discrete trajectory

$$\tilde{X} = (X_{t_1}, ..., X_{t_J}),$$

and the discrete anomaly matrix

$$\tilde{\mathbf{A}} = \begin{pmatrix} \mathbf{A}_{t_1} \\ \vdots \\ \mathbf{A}_{t_J} \end{pmatrix} \in \mathbb{R}^{(JD_X)\times N},$$

where $D_X = |\Omega_x| \times d$ denotes the dimension of the state space. The corresponding ensemble covariance matrix is defined accordingly by $\tilde{\mathbf{P}} = \frac{1}{N-1}\tilde{\mathbf{A}}\tilde{\mathbf{A}}^T \in \mathbb{R}^{(JD_X)\times(JD_X)}$.
With these notations, the cost function can be rewritten

$$J(\tilde{X}) = \frac{1}{2}||\tilde{\mathbf{H}}(\tilde{X}) - Y||^2_{\mathcal{Y},\tilde{\mathbf{R}}^{-1}} + \frac{1}{2}||\tilde{X} - \overline{\tilde{X}}||^2_{\Omega^J,\tilde{\mathbf{P}}^{-1}}.$$

This cost function is very similar to the classical one introduced in Section 1.2, but with extended observation and state spaces. Moreover, it is now justified by the smoothness hypothesis that the penalization term can be seen as a RKHS norm and that the superposition principle applies to the states X. Following the Kalman equations that were

presented in Section 1.3, let us now define the Kalman gain matrix

$$\tilde{\mathbf{K}} = \tilde{\mathbf{P}}\tilde{\mathbf{H}}^T(\tilde{\mathbf{H}}\tilde{\mathbf{P}}\tilde{\mathbf{H}}^T + \tilde{\mathbf{R}})^{-1}, \tag{4.54}$$

where $\tilde{\mathbf{H}} = \mathbf{I}_J \otimes \mathbf{H}$. The ensemble square root filter derivation (see Section 1.3.2) provides the equations for the update of the posterior ensemble: for $n = 1, ..., N$,

$$\tilde{X}_a^{(n)} = \sum_{i=1}^{N} w_i \tilde{X}^{(i)} + \sum_{i=1}^{N} (\tilde{X}^{(i)} - \overline{\tilde{X}})\mathbf{S}_{in}, \tag{4.55}$$

where

$$\mathbf{S} = \left(\mathbf{I}_N + \frac{1}{N-1}(\tilde{\mathbf{H}}\tilde{\mathbf{A}})^T\tilde{\mathbf{R}}^{-1}\tilde{\mathbf{H}}\tilde{\mathbf{A}}\right)^{-\frac{1}{2}} \tag{4.56}$$

and

$$w = \frac{1}{N}\mathbf{1} - \frac{1}{N-1}\mathbf{S}^2(\tilde{\mathbf{H}}\tilde{\mathbf{A}})^T\tilde{\mathbf{R}}^{-1}(\tilde{\mathbf{H}}(\overline{\tilde{X}}) - Y). \tag{4.57}$$

The case $J = 1$ corresponds to the standard scheme where an ESRF is applied each time an observation comes in. Contrary to classical smoothing techniques with forward-backward strategy (Anderson and Moore, 2012), the incorporation of several observations ($J > 1$) is here immediate.
Localization techniques can be applied in the very same way as for a classical square root filter (cf Section 1.3.4). Still, localization remains a questionable technique in the general case, and its justification in the RKHS formalism stays unclear.

4.4.3 Ensemble Kalman filter on the initial perturbation

In this subsection, a new cost function accounting for the Koopman formalism is introduced and a Kalman-like scheme is presented, that updates the vectors of coefficients that define the ensemble members as linear combinations instead of the ensemble members themselves. As previously, these coefficients can be considered constant in time, so the method boils down to studying the initial perturbation.

Regularization by the initial perturbation

The initial perturbation is defined as the following linear combination of the ensemble initial conditions

$$X_0 = \sum_{i=1}^{N} w_i X_0^{(i)} = \mathbf{X}_0 w \in \Omega_0, \tag{4.58}$$

where $\mathbf{X}_0 = [X_0^{(1)}, \cdots, X_0^{(N)}]$ and $w = (w_1, ..., w_N)^T$. The initial anomaly X_0' is defined accordingly, with $X' = X - \overline{X}$ and $\overline{X}$ is the empirical ensemble mean. Note that

$$X_0' = \mathbf{A}_0 w, \tag{4.59}$$

where $\mathbf{A}_0$ is the initial anomaly matrix.

As we seek the solution as the perturbation of the initial conditions, we suggest the following regularization term to replace the second term of the cost function by

$$P = \frac{1}{2}||k_0(\cdot, X_0')||^2_{\mathcal{H}_0}. \tag{4.60}$$

which is the initial anomaly squared norm in $\mathcal{H}_0$. From the reproducing property, $k_0(\cdot, X_0')$ can be expanded as a linear combination of the feature maps:

$$k_0(\cdot, X_0') = \sum_{i=1}^{N} \alpha_i k_0(\cdot, X_0^{(i)}). \tag{4.61}$$

Using the smoothness hypothesis (Section 4.4.1), the vector of coefficients α is related to w by:

$$\alpha = w - \frac{1}{N}\mathbf{1} := w'. \tag{4.62}$$

The penalization term in this smooth case, denoted by P^s in the following, writes

$$P^s = \frac{1}{2} w'^T \mathbf{K} w', \tag{4.63}$$

where $\mathbf{K}$ is the kernel matrix defined by $\mathbf{K}_{i,j} = k_0(X_0^{(i)}, X_0^{(j)})$. Then, its Hessian matrix is simply

$$\nabla^2_w P^s = \mathbf{K}. \tag{4.64}$$

Expansion of the observation term on the Koopman eigenfunctions

The goal is now to simplify the main term of the cost function with the Koopman formalism. In what follows, the observation operator will be assumed to have all its components in the RKHS, namely $\mathbf{H} = (H, ..., H) \in \mathcal{H}^{D_Y}$.

The infinitesimal generators of the Koopman operator $A_{U,t}$ were shown to be diagonalizable in the RKHS family spectral representation theorem (Theorem 5) with eigenpairs $(\lambda_\ell, \psi_\ell^t)_\ell$. As a consequence, the functions $R_t H \in \mathcal{H}_t$ can be expanded on the Koopman

eigenfunctions by:

$$R_t H(X_t) = \sum_\ell m_\ell^H \psi_\ell^t(X_t) = \sum_\ell m_\ell^H e^{\lambda_\ell t} \psi_\ell^0(X_0), \tag{4.65}$$

where

$$m_\ell^H = \left\langle R_t H, \psi_\ell^t \right\rangle_{\mathcal{H}_t} = \langle H, \psi_l \rangle_{\mathcal{H}} = \langle R_0 H, \psi_l^0 \rangle_{\mathcal{H}_0}, \tag{4.66}$$

recalling that the time-dependant family of basis $(\psi_l^t)_l$ are built in the proof of Theorem 5 (Dufée et al., 2023b) as restrictions of global basis functions $(\psi_l)_l$ of $\mathcal{H}$. Finally,

$$H(X_t) = \sum_\ell m_\ell^H e^{\lambda_\ell t} \psi_\ell^0(X_0). \tag{4.67}$$

As $\psi_l^0 \in \mathcal{H}_0$, the reproducing property enables to rewrite, for all ℓ,

$$\psi_\ell^0(X_0) = \sum_{i=1}^N \beta_\ell^{(i)} k_0(X_0, X_0^{(i)}) = \beta_\ell^T \mathbf{k}, \tag{4.68}$$

where the components of $\mathbf{k}$ are defined by $\mathbf{k}_i = k(X_0, X_0^{(i)})$. In the end, we have

$$H(X_t) = v_t \mathbf{k}, \tag{4.69}$$

where

$$v_t = \sum_l m_l e^{\lambda_l t} \beta_l^T. \tag{4.70}$$

Denoting $\mathbf{V}_t = \begin{pmatrix} v_t \\ \vdots \\ v_t \end{pmatrix} \in \mathbb{R}^{D_Y \times N}$, we get

$$\mathbf{H}(X_t) = \mathbf{V}_t \mathbf{k}. \tag{4.71}$$

Similarly to equation (4.52), it can be shown with the smoothness hypohtesis that, for all $i = 1, ..., N$,

$$k_0(X_0, X_0^{(i)}) = k_0\left(\sum_{j=1}^N w_j X_0^{(j)}, X_0^{(i)}\right) = \sum_{j=1}^N w_j k_0(X_0^{(j)}, X_0^{(i)}). \tag{4.72}$$

As a consequence, we get

$$\mathbf{k} = \mathbf{K}w. \tag{4.73}$$

The first term of the cost function finally rewrites

$$\frac{1}{T}\int_0^T ||\mathbf{H}(X_t) - Y_t||^2_{\mathbf{R}^{-1}} dt = (\mathbf{V}\mathbf{K}w - Y)^* \tilde{\mathbf{R}}^{-1}(\mathbf{V}\mathbf{K}w - Y), \tag{4.74}$$

where $\mathbf{V} = \frac{1}{J}\begin{pmatrix} \mathbf{V}_{t_1} \\ \vdots \\ \mathbf{V}_{t_J} \end{pmatrix} \in \mathbb{R}^{(JD_Y)\times N}$, $Y = (Y_{t_j})_{j=1,\dots,J}$ and $\tilde{\mathbf{R}} = \mathbf{I}_J \otimes \mathbf{R}$.

The cost function accounting for the RKHS framework can now be fully expressed with respect to the w variable:

$$J(w) = \frac{1}{2}(\mathbf{V}\mathbf{K}w - Y)^* \tilde{\mathbf{R}}^{-1}(\mathbf{V}\mathbf{K}w - Y) + \frac{1}{2} w'^T \mathbf{K} w'. \tag{4.75}$$

Kalman scheme

Now the problem boils down to Kalman-like schemes, interpreting the matrix $\mathbf{VK}$ as the observation operator and $\mathbf{K}^{-1}$ as the prior empirical covariance matrix for the variable w, for which no dynamics is prescribed a priori.

More precisely, the Hessian matrix of the cost function is given by

$$\nabla^2_w J_0 = (\mathbf{V}\mathbf{K})^* \tilde{\mathbf{R}}^{-1}(\mathbf{V}\mathbf{K}) + \mathbf{K}. \tag{4.76}$$

Denoting $\mathbf{P}^a$ the posterior ensemble covariance matrix for w, the Sherman-Morrison-Woodbury formula gives

$$\begin{aligned} \mathbf{P}^a &= [(\mathbf{V}\mathbf{K})^* \tilde{\mathbf{R}}^{-1}(\mathbf{V}\mathbf{K}) + \mathbf{K}]^{-1} \\ &= \mathbf{K}^{-1} - \mathbf{G}(\mathbf{V}\mathbf{K})\mathbf{K}^{-1} \end{aligned}$$

where

$$\mathbf{G} = \mathbf{K}^{-1}(\mathbf{V}\mathbf{K})^* \left(\tilde{\mathbf{R}} + (\mathbf{V}\mathbf{K})\mathbf{K}^{-1}(\mathbf{V}\mathbf{K})^*\right)^{-1} \tag{4.77}$$

is the equivalent of the Kalman gain matrix in the classical covariance matrix update equation (1.19). As previously mentioned, the latter equation corresponds to (1.19), interpreting $\mathbf{VK}$ as the observation operator and $\mathbf{K}^{-1}$ as the prior ensemble covariance matrix for the vectors of coefficients w.

The posterior ensemble of vectors of coefficients $(w^{(i),a})_{i=1,...,N}$ can then be updated in a square root filter way. We first update (cf equation (1.18)) the ensemble mean by

$$\overline{w}^a = \overline{w}^f - \mathbf{G}(\mathbf{VK}\overline{w}^f - Y). \tag{4.78}$$

Then, denoting $\mathbf{A}_w$ the anomaly matrix for the vectors of coefficients w, each ensemble member $i = 1, ..., N$ by (cf equation (1.25)):

$$w^{(i),a} = \overline{w}^a + \sum_{j=1}^{N} (w^{(j),f} - \overline{w}^f)\mathbf{S}_{j,n}, \tag{4.79}$$

where, similarly to equation (1.22)

$$\mathbf{S} = \left(\mathbf{I} + \frac{1}{N-1}(\mathbf{VKA}_w^f)^T\mathbf{R}^{-1}\mathbf{VKA}_w^f\right)^{-\frac{1}{2}}. \tag{4.80}$$

The actual ensemble members in the state space are then updated accordingly:

$$X^{(i),a} = \sum_{j=1}^{N} (w^{(i),a})_j X^{(j),f}. \tag{4.81}$$

Note that, as we have $\mathbf{VK}w = \mathbf{H}X$, the matrix $\mathbf{S}$ defined in equation (4.80) is exactly the same as the one defined for the classical ESRF in equation (1.22).

4.4.4 Relaxation of the smoothness hypothesis

This subsection provides an equivalent of the cost function computed in (4.75) without using the smoothness hypothesis for the space-wise evaluation functions g_x. The following will study once again the two terms of the cost function as in the previous subsection. This is a fundamental change, as the linear combination defining the feature maps (coefficients α) and the states (coefficients w) will no longer be the same, and have a nonlinear relationship through the kernel.

Regularization by the initial perturbation

Starting over from (4.58), the regularization term is still defined as

$$P = \frac{1}{2}||k_0(\cdot, X_0')||^2_{\mathcal{H}_0}. \tag{4.82}$$

From the reproducing property (4.61), we get, for all $j = 1, ..., N$,

$$k_0(X_0^{(j)}, X_0') = \sum_{i=1}^{N} \alpha_i k_0(X_0^{(j)}, X_0^{(i)}). \tag{4.83}$$

So the coefficients α rewrite

$$\alpha = \mathbf{K}^{-1} \overrightarrow{k}', \tag{4.84}$$

where $\overrightarrow{k}'_i = k_0(X_0^{(i)}, X_0')$. The regularization term then rewrites:

$$\begin{aligned} 2P = ||k_0(\cdot, X_0')||^2_{\mathcal{H}_0} &= \left\langle \sum_{i=1}^{N} \alpha_i k_0(\cdot, X_0^{(i)}), \sum_{j=1}^{N} \alpha_j k_0(\cdot, X_0^{(j)}) \right\rangle_{\mathcal{H}_0} \\ &= \sum_{i,j} \alpha_i \alpha_j^* k_0(X_0^{(j)}, X_0^{(i)}) \\ &= \alpha^* \mathbf{K} \alpha \\ &= \overrightarrow{k}'^* \mathbf{K}^{-1} \overrightarrow{k}'. \end{aligned}$$

Expansion of the observation term on the Koopman eigenfunctions

Starting over from equation (4.71), the smoothness assumption no longer ensures that $\overrightarrow{k} = \mathbf{K}w$ (equation 4.73). So the integral term of the cost function writes

$$\frac{1}{T} \int_0^T ||\mathbf{H}(X_t) - Y_t||^2_{\mathbf{R}^{-1}} dt = (\mathbf{V}\overrightarrow{k} - Y)^* \tilde{\mathbf{R}}^{-1} (\mathbf{V}\overrightarrow{k} - Y), \tag{4.85}$$

with the notations of (4.74). The whole cost function then rewrites with respect to the vector $\overrightarrow{k}$ as follows:

$$J(w) = \frac{1}{2}(\mathbf{V}\overrightarrow{k} - Y)^* \tilde{\mathbf{R}}^{-1} (\mathbf{V}\overrightarrow{k} - Y) + \frac{1}{2} \overrightarrow{k}'^* \mathbf{K}^{-1} \overrightarrow{k}'. \tag{4.86}$$

Gradient and Hessian matrix computations

For the gradient computation, we will use the following result: for any $Y \in \Omega_0$, we have

$$\nabla_w k_0(X_0, Y) = \mathbf{X}_0^* \nabla_X k_0(X_0, Y) \tag{4.87}$$

and

$$\nabla_w^2 k_0(X_0, Y) = \mathbf{X}_0^* \nabla_X k_0(X_0, Y) \mathbf{X}_0. \tag{4.88}$$

Defining a 3D tensor $\tilde{\mathbf{X}}_0 = \begin{pmatrix} \mathbf{X}_0 \\ \vdots \\ \mathbf{X}_0 \end{pmatrix} \in \mathbb{R}^{N \times D_X \times N}$ and denoting respectively $\cdot^{*2,3}$ and $\otimes_{2,3}$ the transposition and tensor product with respect to the last two coordinates, we get

$$\nabla_w \overrightarrow{k}' = \tilde{\mathbf{X}}_0^{*2,3} \otimes_{2,3} \nabla_X \overrightarrow{k}' = \begin{pmatrix} \nabla_w k_0(X_0^{(1)}, X_0') \\ \vdots \\ \nabla_w k_0(X_0^{(N)}, X_0') \end{pmatrix} \in \mathbb{R}^{N \times N} \tag{4.89}$$

and

$$\nabla_w^2 \overrightarrow{k}' = \tilde{\mathbf{X}}_0^{*2,3} \otimes_{2,3} \nabla_X^2 \overrightarrow{k}' \otimes_{2,3} \tilde{\mathbf{X}}_0 = \begin{pmatrix} \nabla_w^2 k_0(X_0^{(1)}, X_0') \\ \vdots \\ \nabla_w^2 k_0(X_0^{(N)}, X_0') \end{pmatrix} \in \mathbb{R}^{N \times N \times N} \tag{4.90}$$

We can then compute the gradient of P :

$$2\nabla_w P = (\nabla_w \overrightarrow{k}')^* \mathbf{K}^{-1} \overrightarrow{k}' + \overrightarrow{k}'^* \mathbf{K} \nabla_w \overrightarrow{k}' \tag{4.91}$$

and the hessian matrix

$$2\nabla_w^2 P = (\nabla_w^2 \overrightarrow{k}')^* \mathbf{K}^{-1} \overrightarrow{k}' + \overrightarrow{k}'^* \mathbf{K}^{-1} \nabla_w^2 \overrightarrow{k}' + 2(\nabla_w \overrightarrow{k}')^* \mathbf{K}^{-1} \nabla_w \overrightarrow{k}'. \tag{4.92}$$

4.4.5 Numerical experiments

Similarly to what was done in Section 3.1, we describe the numerical setup for the experiments.

Contrary to the experiments of Chapter 3, only the deterministic SQG will be at stake here, as the stochasticity is not handled yet by the RKHS formalism. The variability comes from the set of initial conditions. They are generated from the exact same procedure that generates the SVD noise (cf Section 2.3.2). As a quick reminder, this procedure consists in building pseudo-observations resulting from the random draws of local velocity fluctuations around a sample mean. Applying this procedure to the reference initial condition for the truth of the system (cf Figure 3.1), the set of initial conditions is taken as the set of pseudo-observations gathered in the matrix V of equation (2.20).

The simulation grid G_s remains of size 64×64, but we now consider a new type

of observations, namely one meridional line of partial observations, moving in time, at the maximum resolution resolved by the simulations. Consequently, the observation grid G_o now has size 64×1, instead of 16×16 in the previous experiments. This change is motivated by a wish to get closer to realistic settings, where lines of observations at finer resolution can be collected much more often than global observed fields at coarser resolution, which are in practice agglomerated and interpolated from past series of sparse observations. Our numerical setting can be seen as a very simplified version of the one chosen for wide-swath altimetry in Le Guillou et al. (2021), where back and forth nudging is coupled with 4D-Var methods for both balanced motions and internal tides. In our setting, the snapshots of two meridional lines will be assumed to be collected 4 times a day (every 6h) instead of one coarse 16×16 observation collected every day.

We first studied the performances of the ensemble square root filter (ESRF) for pieces of trajectory (Section 4.4.2) that agglomerates the $J = 4$ observations over 1 day before filtering with the update formula given by equation (4.55). In Figure 4.1, it is compared with a more classical scheme that consists in applying a standard ESRF every 6h with the same nature of observations considered (two meridional lines moving with time), which in fact corresponds to applying the previous scheme with $J = 1$. It appears to be very beneficial in the long term to wait for multiple observations to be gathered before filtering with a global ESRF.

This statement is not necessarily true for the local ESRF, as shown in Figure 4.2. As the local ESRF corrects much more brutally the ensemble members, waiting too long for upcoming observations can be detrimental for the filtering, as the ensemble may not be corrected soon and strong enough. In Figure 4.2, we tested different numbers J of observations to gather before performing a local ESRF in order to find an adequate trade-off. It appears that waiting for 3 partial observations or more becomes detrimental in this case.

Second, we compared the ESRF for pieces of trajectory with the ESRF on the initial perturbation (Section 4.4.3), which is enriched with a regularization term in the RKHS (equation (4.60)). The kernels considered are the empirical covariance kernel (4.46) and the Gaussian kernel (4.48). As the balance between the observation term and the penalization term in the original cost function and the RKHS enriched cost function (4.75) is clearly modified, we allow ourselves to set a scaling parameter for the penalization term in (4.60).

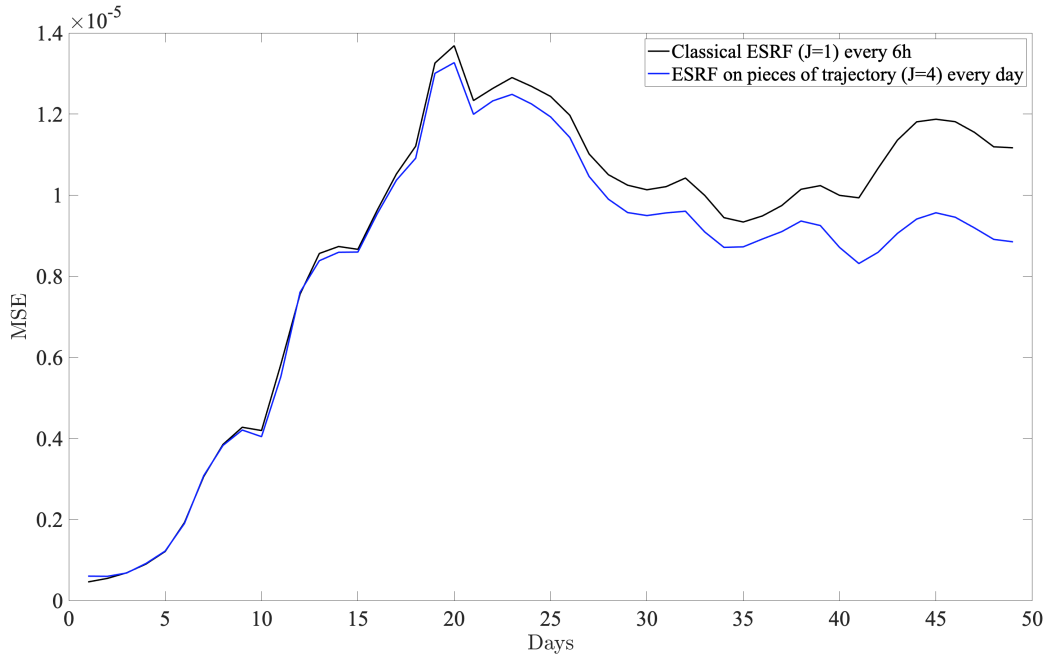


Figure 4.1 – Mean Square Error curves for a classical scheme where a standard ESRF is applied each time (every 6 hours) an observation comes up (in black), compared with the ESRF scheme on pieces of trajectory that gathers one day ($J = 4$) of observations before performing the filtering (in blue).

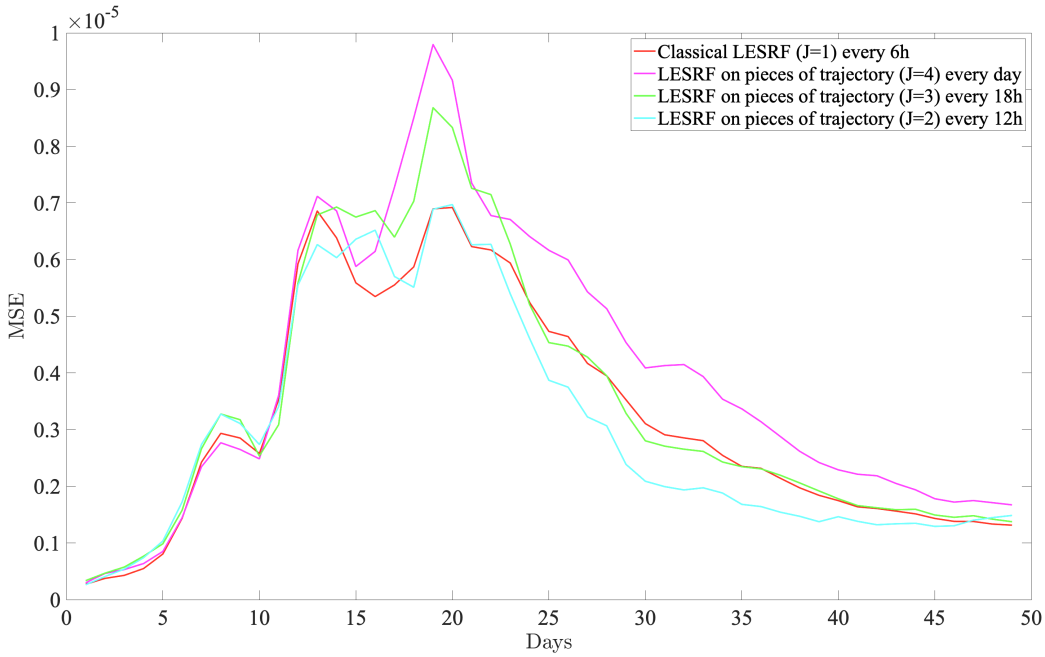


Figure 4.2 – Mean Square Error curves for a classical scheme where a standard LESRF is applied each time (every 6 hours) an observation comes up (in red), compared with the ESRF scheme on pieces of trajectory that gathers multiple observations before performing the filtering: $J = 4$ observations (in magenta), $J = 3$ observations (in green) and $J = 2$ observations (in purple). It appears that waiting for more than two observations becomes detrimental for the filter.

Namely, this term will be computed as

$$P = \frac{C}{2}||k_0(\cdot, X_0')||^2_{\mathcal{H}_0}, \tag{4.93}$$

where C is a scaling parameter to tune and that can highly depend on the kernel considered. Figure 4.3 compares, additionally to Figure 4.1, the ESRF with the RKHS regularization term for both the empirical and Gaussian kernel, with the parameters C that yield the best results for each kernel. We observe that, with the empirical kernel tuned with the right scaling factor $C = 10^5$, the ensemble behaves relatively alike the empirical mean, it is a very decent estimator in the early stages and then diverges in the long term. However, the Gaussian kernel with its well-tuned parameter ($C = 10^4$) brings a significant improvement starting from day 17 approximately. The main ingredient of the success of the Gaussian kernel in the numerics, which is striking in our experiments, is a high production of spread, as shown in Figure 4.4, with the example of day 35, where the Gaussian kernel is at its lower MSE. This is certainly due to a higher rank of the Gaussian kernel matrix compared to the empirical kernel, coupled with non negligible non-diagonal coefficients, penalizing the redundancies in the ensemble.

4.5 Synthesis

This chapter investigates the framework of reproducing kernel Hilbert spaces (RKHS) for the Koopman operator. A spectral representation of the Koopman operator is formulated in this setting. Within this framework, some classical data assimilation techniques are adapted and enriched with the theoretical properties of RKHS. Among them, the superposition principle entailed by EnKFs is interpreted as a consequence of a fundamental property of RKHS, and its use is legitimated in this setting. These modified data assimilation techniques rely on a justified gathering of observations along a trajectory time-horizon before performing the filtering, instead of assimilating each observation sequentially. Numerical experiments show that it can be worth accumulating information from multiple observations in the proposed modified Kalman schemes.

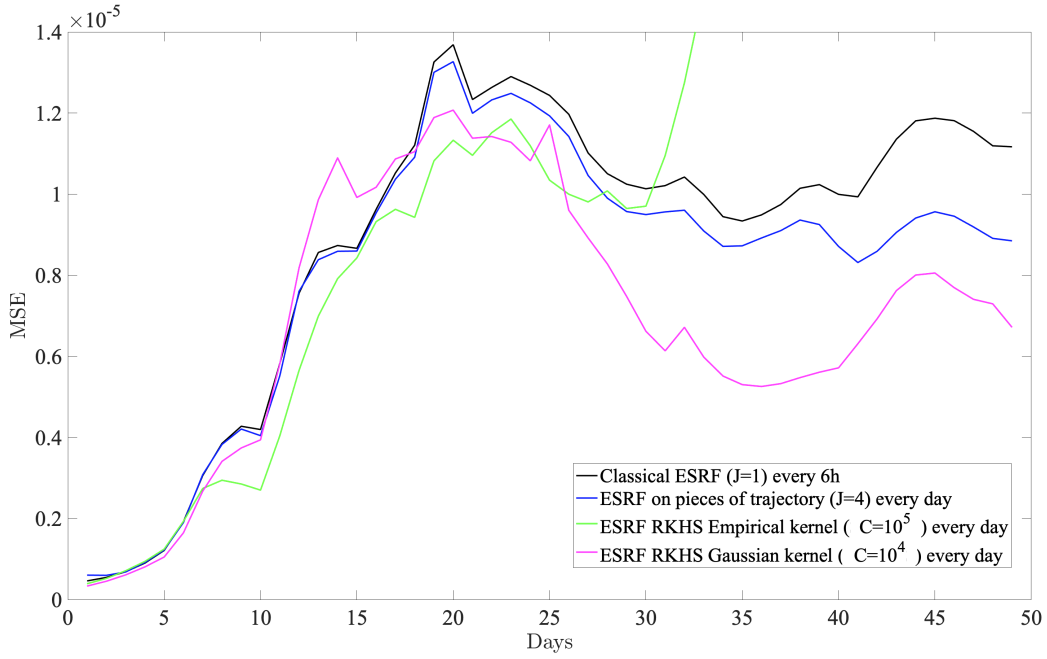
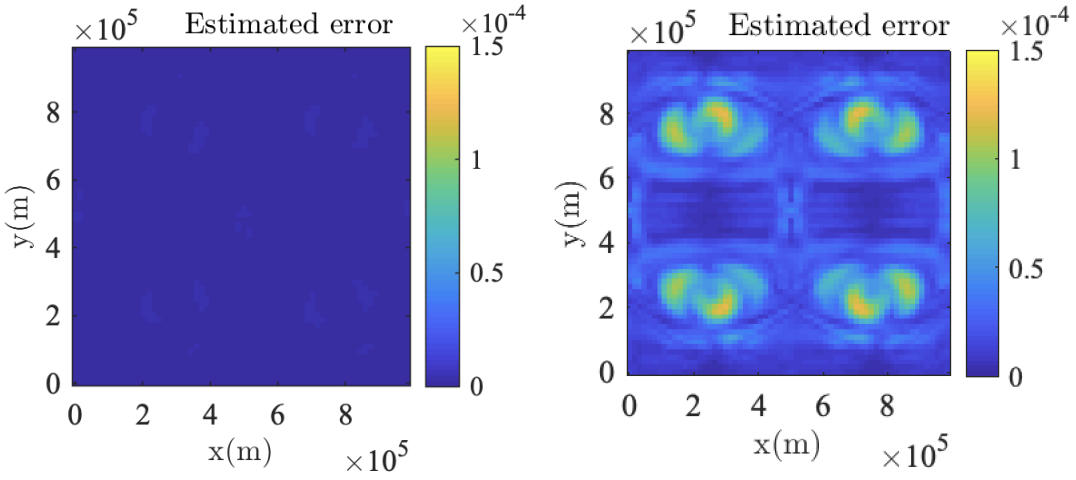


Figure 4.3 – Mean Square Error curves for a classical scheme where a standard ESRF is applied each time (every 6 hours) an observation comes up (in black), compared with the ESRF scheme on pieces of trajectory (Section 4.4.2) that gathers one day ($J = 4$) of observations before performing the filtering (in blue), and ESRF schemes enriched with a RKHS regularization term (Section 4.4.3) with both empirical (in green) and Gaussian (in magenta) kernel, with the best scaling parameter C for each kernel.



ESRF on pieces of trajectory ESRF RKHS Gaussian kernel

Figure 4.4 – Comparison of spread at day 35 between the ESRF on pieces of trajectory and the ESRF with RKHS regularization term with a Gaussian kernel and the scaling parameter $C = 10^4$.

Conclusive chapter

This conclusive chapter aims at putting in perspective the process that led to the work presented in this thesis and explain the choices that were made to present part of the job that was tried and explored.

In this thesis, we investigated data assimilation techniques for ocean models and particularly stochastic ones. The stochastic framework driving these models, called Location Uncertainty, was first presented in its initial principles and its main features were shown. It is based on a decomposition of the Lagrangian velocity into a time-smooth component and a highly oscillating noise term. It satisfies classical physical conservation laws and provides a stochastic equivalent of the deterministic material derivative. This stochastic transport operator was stated in a Surface Quasi-Geostrophic (SQG) case and the different possibilities in order to generate the noise were exposed.

The main objective of this thesis was to handle high-dimensional systems with nonlinear features from a data assimilation point of view. In the LU framework, the cylindrical Wiener process provides a multiplicative non Gaussian noise in the stochastic transport operator. In short, both conditions that are theoretically set to derive classical ensemble Kalman filters are clearly not met in the models at stake. As they constitute a widely spread data assimilation technique, we first implemented them in their different variations, with the localized ensemble square root filter (ESRF) as the method with the best numerical results in our framework. Still, the two main features of this method are theoretically more than questionable. First, these methods look for the posterior ensemble members as linear combinations of the prior ones, which consequently cannot be solutions of the underlying dynamical system anymore. Second, the localization procedure is also likely to introduce important gradients through different decisions taken at neighboring points, potentially destroying the physical balance of the posterior members. We even pointed out that, with deterministic dynamics coupled with inflation, the system blew up in finite time, partially due to the fact that the SQG system is based on an elliptic equation that will not handle unexpectedly high gradients. This issue was stabilized by the introduction of the stochastic parametrization, that improved the latter procedure from many aspects, and these findings were published in Dufée et al. (2022).

In order to handle nonlinear systems with no hypothesis on the model noise, particle filters constitute the theoretical ideal framework. Still, these methods are well-known to struggle with high-dimensional systems. As they are Monte-Carlo methods, the number of realizations to run has to scale exponentially with the dimension of the state space to be effective, which is computationally intractable. The main issue, called ensemble degeneracy, consists in one single particle absorbing the whole ensemble because it is the closest to the observation in an Euclidian norm quantification, which overdiscriminates particles in high-dimension. Some strategies were figured out to handle filter degeneracy (see Section 1.4 and references therein). Tempering consists in guiding the set of particles towards the observation through a Monte-Carlo Markov Chain (MCMC) procedure. Jittering then intends to resimulate multiple realizations in a small neigbourhood of the best particles after a possible filter degeneracy, in order to recover the diversity that was lost in the filtering. Although these methods are very interesting from a theoretical point of view, they did not prove effective enough in our case of interest.

Tempering is a very costly procedure that needs to resimulate whole pieces of trajectory many times. Jittering is not as expensive as tempering, as it just consists in one re-run to perturb around the particles of interest that were picked by the filter. In order to reduce its cost, in our case, we had to be very careful about the initial conditions, and more generally one needs to ensure that the ensemble spread contains the observation so that the method can indeed enhance the standard particle filter.

Numerical results are produced in Figures C.1 and C.2. For this simulation, we sticked to the numerical setting of Section 3.2. Namely, we started a set of 100 particles starting from the same "exact" initial condition of the system with the LU parametrization and a SVD noise. In this setting, the comparison was made between local and global Ensemble Square Root Filter (LESRF and ESRF), the standard particle filter (SIR) and a modified particle filter with tempering and jittering.

Figure C.1 shows the comparison of MSE between these filters over 10 days. In this period, the SQG system is quite predictable at large scales and the LU framework was shown to provide sufficient spread so that the observations are contained within the spread (even the empirical mean is a very decent estimator in this case). When these conditions are fulfilled, we can see that the modified particle filter matches the performances of the local ensemble square root filter (still the comparison is not entirely fair as the computational cost of tempering is much bigger).

However, when the SQG model starts its highly non-linear behavior, typically between

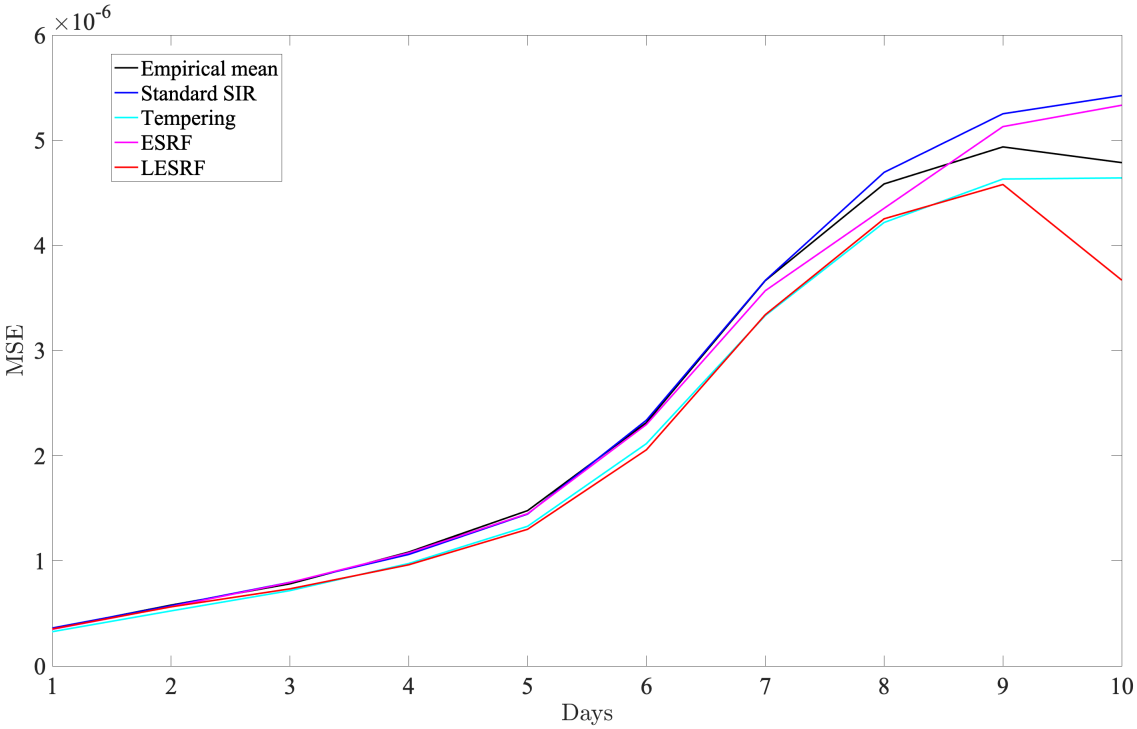


Figure C.1 – Mean Square Error curves for different types of particle and Kalman filters. In this 10 day period, the modified particle filter with tempering and jittering (cyan) performs better than standard SIR (blue) and global ESRF (magenta) and matches the local ESRF performances (red).

day 15 and 20, the enhanced particle filter diverges exactly as its standard version, while the ensemble square root filters remain very stable. The corresponding results are shown in Figure C.2.

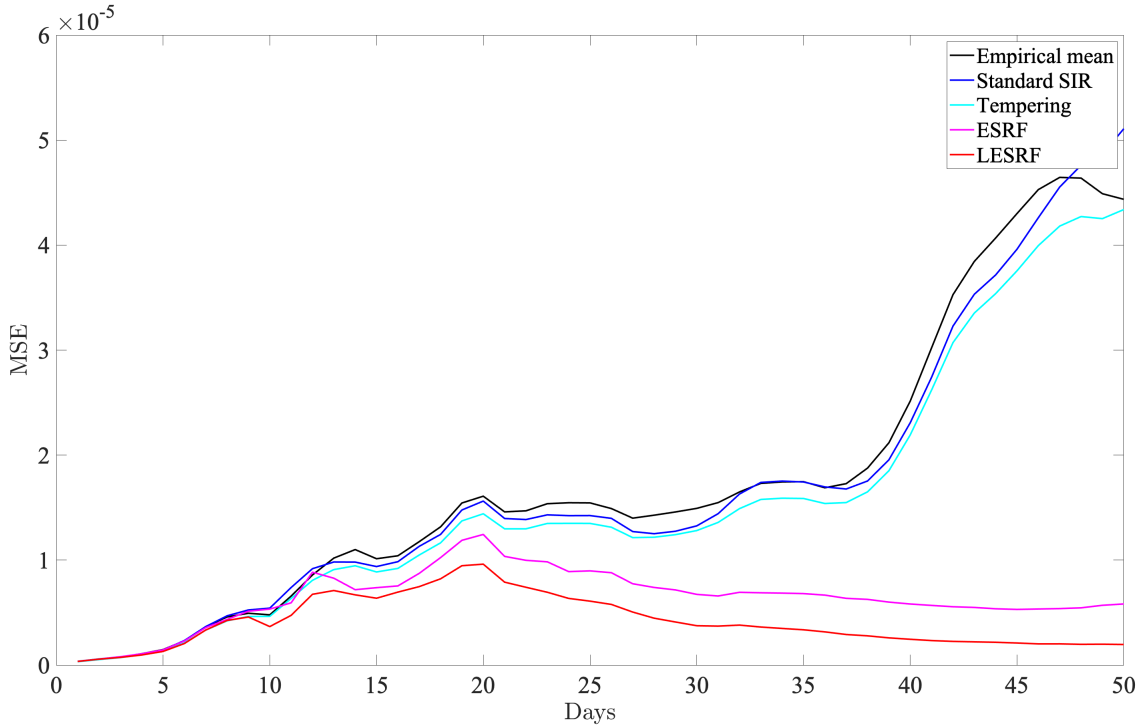


Figure C.2 – Mean Square Error curves for different types of particle and Kalman filters. Contrary to Kalman filters, even the enhanced particle filter diverges at long term.

Local particle filters (Farchi, 2019; Poterjoy, 2016; Shen et al., 2017) or hybrid filters combining Kalman and particle filter methodologies (Papadakis et al., 2010; Shen and Tang, 2015; Reinhardt, 2020) were shown to prove as effective or slightly better than ensemble Kalman filters, with possibly a more crucial need of an ensemble size large enough. Still, localization raises the same theoretical problems for particle filters, as the weights of the posterior distribution of the global state are no longer global. Consequently, the choice that was made in this thesis consists in elaborating on ensemble Kalman filters in the purpose of the study of stochastic geophysical systems, as they constitute the best up-to-date method in the benchmarks that we made on the model at hand, the SQG system.

As previously stated, we have first shown that, for this model, the introduction of the LU stochastic parametrization enables to improve a localized ESRF when compared to an ensemble simulation of the original deterministic dynamics. The stochastic framework was

shown to provide better MSE performance and spread together with an improved physical relevance compared to deterministic dynamics. Finally, it does not require variance inflation, contrary to the latter. The parameter of such inflation is known to be difficult to set and requires a fine tuning to get good performances. The fact that the stochastic framework is spared from inflation is a first clear advantage. Moreover, we have shown that, even for small values of the inflation parameter, the classical deterministic system was exploding in finite time when combined with local filtering, which is likely to be due to the combination of both (localization and inflation) procedures, as explained before.

The stochastic framework also provides an additional degree of freedom, which is the noise in itself. In particular, we explored the idea to give this noise a non-zero expectation. Girsanov transform ensures the possibility to conserve the form of the equation up to changing the underlying probability measure. This non-zero expectation is computed with techniques that are close to optimal flow estimation procedures, and is designed to play a guiding role for the set of realizations, targeting the upcoming observation. This first quite rough approach was shown to be very effective in case of underestimation of the initial condition (for instance resulting from initial estimations relying on regularized inverse problems). A multi-resolution scheme was then designed in order to refine this first approach and to compute smoother guiding terms for the ensemble members.

Finally, the last part of this thesis attempts to propose a bridge with respect to the questions raised about the ensemble Kalman filter methodology. In this section, we aim in particular at justifying the superposition principle, which is the core feature of the ensemble Kalman filters, through the theoretical help of reproducing kernel Hilbert spaces (RKHS). The Koopman operator acting on the RKHS associated with the phase space is studied and a spectral representation theorem is stated. In this RKHS, the feature maps can, by essence, be expressed as linear combinations of ensemble feature maps, which legitimates the goal to look for the posterior ensemble members through a superposition principle.

Part of the perspectives for future work consists in the extension of the Wonderland representation for stochastic dynamical systems like in LU, and to observe how randomness plays a role in a stochastic version of a spectral representation theorem for the stochastic Koopman operator. For such stochastic setting, as two degrees of freedom are now introduced (initial condition and noise realization), one could think of an ensemble (indexed by randomness) of ensembles of realizations starting from different initial conditions, given this randomness (or the associated noise term). In such setting, an interesting

way to proceed could be to use a particle filter that picks the best ensemble(s) in terms of randomness, so the best realizations of the noise, and then let an ensemble Kalman filter assimilate the chosen ensembles. These are just ideas, which would need an adequate stochastic Wonderland spectral theorem to be further developed. As this setting seems to offer new perspectives for data assimilation techniques, we also hope this framework, either to justify the localization in some way or, even better, to get rid of it through more sophisticated approaches.

BIBLIOGRAPHY

Anderson, B., Moore, J., 2012. Optimal Filtering. Dover Books on Electrical Engineering, Dover Publications. URL: https://books.google.fr/books?id=iYMqLQp49UMC.

Anderson, J., 2001. An ensemble adjustment Kalman filter for data assimilation. Monthly Weather Review 129, 2884–2903.

Anderson, J., 2007. An adaptive covariance inflation error correction algorithm for ensemble filters. Tellus 59A, 210–224.

Anderson, J., 2012. Localization and sampling error correction in ensemble kalman filter data assimilation. Monthly Weather Review 140, 2359 – 2371. URL: https://journals.ametsoc.org/view/journals/mwre/140/7/mwr-d-11-00013.1.xml, doi:10.1175/MWR-D-11-00013.1.

Arnaud, E., Mémin, E., Cernuschi-Frias, B., 2005. Conditional filters for image sequence based tracking - application to point tracking. IEEE Trans. on Image Processing 14, 63–79.

Aronszajn, N., 1950. Theory of reproducing kernels. Trans. of the American Math. Society 68.

Asch, M., Bocquet, M., Nodet, M., 2016. Data Assimilation: Methods, Algorithms, and Applications. Fundamentals of algorithms, SIAM.

Bauer, W., Chandramouli, P., Chapron, B., Li, L., Mémin, E., 2020a. Deciphering the role of small-scale inhomogeneity on geophysical flow structuration: a stochastic approach. Journal of Physical Oceanography 50, 983–1003.

Bauer, W., Chandramouli, P., Li, L., Mémin, E., 2020b. Stochastic representation of mesoscale eddy effects in coarse-resolution barotropic models. Ocean Modelling 151, 101646.

Berlinet, T., Thomas-Agnan, C., 2001. Reproducing kernel Hilbert spaces in Probability and Statistics. Kluwer Academic Publishers.

Berner, J., Achatz, U., Batté, L., Bengtsson, L., de la Cámara, A., Christensen, H.M., Colangeli, M., Coleman, D.R.B., Crommelin, D., Dolaptchiev, S.I., Franzke, C.L.E., Friederichs, P., Imkeller, P., Järvinen, H., Juricke, S., Kitsios, V., Lott, F., Lucarini, V., Mahajan, S., Palmer, T.N., Penland, C., Sakradzija, M., von Storch, J.S., Weisheimer, A., Weniger, M., Williams, P.D., Yano, J.I., 2017. Stochastic parameterization: Toward a new view of weather and climate models. Bulletin of the American Meteorological Society 98, 565 – 588. URL: https://journals.ametsoc.org/view/journals/bams/98/3/bams-d-15-00268.1.xml, doi:10.1175/BAMS-D-15-00268.1.

Beskos, A., Crisan, D., Jasra, A., 2014. On the stability of sequential Monte Carlo methods in high dimensions. The Annals of Applied Probability 24, 1396 – 1445. URL: https://doi.org/10.1214/13-AAP951, doi:10.1214/13-AAP951.

Beskos, A., Crisan, D., Jasra, A., Kamatani, K., Zhou, Y., 2017. A stable particle filter for a class of high-dimensional state-space models. Advances in Applied Probability 49, 24–48. URL: http://www.jstor.org/stable/44985412.

Bishop, A.N., Del Moral, P., 2023. On the mathematical theory of ensemble (linear-gaussian) kalman-bucy filtering. arXiv:2006.08843.

Bishop, C., Etherton, B., Majumdar, S., 2001. Adaptive sampling with the ensemble transform Kalman filter. part I: Theoretical aspects. Monthly weather review 129, 420–436.

Blayo, E., Bocquet, M., Cosme, E., Cugliandolo, L.F., 2014. Advanced Data Assimilation for Geosciences: Lecture Notes of the Les Houches School of Physics: Special Issue, June 2012. Oxford University Press. URL: https://doi.org/10.1093/acprof:oso/9780198723844.001.0001, doi:10.1093/acprof:oso/9780198723844.001.0001.

Blumen, W., 1978. Uniform potential vorticity flow: part I. theory of wave interactions and two-dimensional turbulence. Journal of the Atmospheric Sciences 35, 774–783.

Bocquet, M., 2011. Ensemble kalman filtering without the intrinsic need for inflation. Nonlinear Processes in Geophysics 18, 735–750. URL: https://npg.copernicus.org/articles/18/735/2011/, doi:10.5194/npg-18-735-2011.

Boussinesq, J., 1877. Mémoires présentés par divers savants à l'Académie des Sciences, 23 (1): 1–680.

Brecht, R., Li, L., Bauer, W., Mémin, E., 2021. Rotating shallow water flow under location uncertainty with a structure-preserving discretization. arXiv, physics.flu-dyn.

Brunton, S., Brunton, B., Proctor, J., Kaiser, E., Kutz, J., 2017. Chaos as an intermittently forced linear system. Nat. Commun. 8.

Budisic, M. and Mohr, R. and Mezic, I. , 2012. Applied koopmanism. Chaos: An Interdisciplinary Journal of Nonlinear Science 22, 047510. doi:10.1063/1.4772195.

Cai, S., Mémin, E., Dérian, P., Xu, C., 2018. Motion estimation under location uncertainty for turbulent fluid flows. Experiments in Fluids 59, 8.

Carrassi, A., Bocquet, M., Bertino, L., Evensen, G., 2018. Data assimilation in the geosciences: An overview of methods, issues, and perspectives. WIREs Climate Change 9, e535.

Chandramouli, P., Memin, E., Heitz, D., 2020. 4d large scale variational data assimilation of a turbulent flow with a dynamics error model. Journal of Computational Physics 412, 109446. URL: http://www.sciencedirect.com/science/article/pii/S0021999120302205, doi:https://doi.org/10.1016/j.jcp.2020.109446.

Chapron, B., Dérian, P., Mémin, E., Resseguier, V., 2018. Large-scale flows under location uncertainty: a consistent stochastic framework. QJRMS 144, 251–260.

Constantin, P., Lai, M., Sharma, R., Tseng, Y., Wu, J., 2012. New numerical results for the surface quasi-geostrophic equation. Journal of Scientific Computing 50, 1–28.

Constantin, P., Majda, A., Tabak, E., 1994. Formation of strong fronts in the 2-D quasi-geostrophic thermal active scalar. Nonlinearity 7, 1495.

Constantin, P., Nie, Q., Schörghofer, N., 1999. Front formation in an active scalar equation. Physical Review E 60, 2858.

Cotter, C., Crisan, D., Holm, D., Pan, W., Shevchenko, I., 2020a. Data assimilation for a quasi-geostrophic model with circulation-preserving stochastic transport noise. Journal of Statistical Physics 179, 1186–1221. URL: https://doi.org/10.1007/s10955-020-02524-0, doi:10.1007/s10955-020-02524-0.

Cotter, C., Crisan, D., Holm, D., Pan, W., Shevchenko, I., 2020b. A particle filter for stochastic advection by lie transport (salt): A case study for the damped and forced incompressible 2d euler equation. arXiv:1907.11884.

Crisan, D., Doucet, A., 2002. A survey of convergence results on particle filtering methods for practitioners. IEEE Transactions on Signal Processing 50, 736–746. doi:10.1109/78.984773.

Cucker, F., Smale, S., 2001. On the mathematical foundation of learning. Bul. of the Amer. Math. Soc. 39, 1–49.

Das, S., D.Giannakis, 2020. Koopman spectra in reproducing kernel Hilbert spaces. Appl.Comput.Harmon.Anal. 49, 573–607.

Das, S., Giannakis, D., 2019. Delay-coordinate maps and the spectra of Koopman operators. J. Stat. Phys. 175, 1107–1145.

Das, S., Giannakis, D., Slawinska, J., 2021. Reproducing kernel Hilbert space compactification of unitary evolution groups. Applied and Computational Harmonic Analysis 54, 75–136. doi:https://doi.org/10.1016/j.acha.2021.02.004.

Del Moral, P., 1997. Nonlinear filtering: Interacting particle resolution. Comptes Rendus de l'Académie des Sciences - Series I - Mathematics 325, 653–658. URL: https://www.sciencedirect.com/science/article/pii/S0764444297847787, doi:https://doi.org/10.1016/S0764-4442(97)84778-7.

Del Moral, P., Guionnet, A., 1999. Central limit theorem for nonlinear filtering and interacting particle systems. The Annals of Applied Probability 9, 275 – 297. URL: https://doi.org/10.1214/aoap/1029962742, doi:10.1214/aoap/1029962742.

Delyon, B., Hu, Y., 2006. Simulation of conditioned diffusion and application to parameter estimation. Stochastic Processes and their Applications 116, 1660–1675. URL: https://www.sciencedirect.com/science/article/pii/S0304414906000469, doi:https://doi.org/10.1016/j.spa.2006.04.004.

Dufée, B., Mémin, E., Crisan, D., 2023a. Observation-based noise calibration: An efficient dynamics for the ensemble kalman filter, in: Chapron, B., Crisan, D., Holm, D., Mémin, E., Radomska, A. (Eds.), Stochastic Transport in Upper Ocean Dynamics, Springer International Publishing, Cham. pp. 43–56.

Dufée, B., Hug, B., Memin, E., Tissot, G., 2023b. Ensemble forecasts in reproducing kernel hilbert space family: dynamical systems in wonderland. arXiv:2207.14653.

Dufée, B., Mémin, E., Crisan, D., 2022. Stochastic parametrization: An alternative to inflation in ensemble kalman filters. Quarterly Journal of the Royal Meteorological Society 148, 1075–1091. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.4247, doi:https://doi.org/10.1002/qj.4247, arXiv:https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.4247.

Eisner, T., Farkas, B., Haase, M., Nagel, R., 2015. Operator Theoretic Aspects of Ergodic Theory. Graduate Texts in Mathematics, Springer.

Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. Journal of Geophysical Research: Oceans 99, 10143–10162. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/94JC00572, doi:https://doi.org/10.1029/94JC00572, arXiv:https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/94JC00572.

Evensen, G., 2006. Data assimilation: The ensemble Kalman filter. Springer-Verlag, New-york.

Farchi, A., 2019. Localisation des méthodes d'assimilation de donnée d'ensemble. Theses. Université Paris-Est. URL: https://pastel.archives-ouvertes.fr/tel-03149002.

Franzke, C.L.E., O'Kane, T.J., Berner, J., Williams, P.D., Lucarini, V., 2014. Stochastic climate theory and modeling. Wiley Interdisciplinary Reviews: Climate Change 6, 63–78.

Giannakis, D., 2017. Data-driven spectral decomposition and forecasting of ergodic dynamical systems. Appl. Comput. Harmon. Anal. https://doi.org/10.1016/j.acha.2017.09.001.

Giannakis, D., Slawinska, J., Zhao, Z., 2015. Spatiotemporal feature extraction with data-driven Koopman operators, in: Journ. of Mach. Learn. Res.: Workshop and Conference Proceedings, pp. 103–115.

Gottwald, G.A., Crommelin, D.T., Franzke, C.L.E., 2016. Stochastic climate theory. arXiv:1612.07474.

Gratton, S., Lawless, A.S., Nichols, N.K., 2007. Approximate gauss–newton methods for nonlinear least squares problems. SIAM Journal on Optimization 18, 106–132. URL: https://doi.org/10.1137/050624935, doi:10.1137/050624935, arXiv:https://doi.org/10.1137/050624935.

Gürol, S., Weaver, A.T., Moore, A.M., Piacentini, A., Arango, H.G., Gratton, S., 2014. B-preconditioned minimization algorithms for variational data assimilation with the dual formulation. Quarterly Journal of the Royal Meteorological Society 140, 539–556. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2150, doi:https://doi.org/10.1002/qj.2150, arXiv:https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.2150.

Hamill, T., Whitaker, J., Snyder, C., 2001. Distance-dependent filtering of background error covariance estimates in an ensemble kalman filter. Monthly Weather Review 129, 2776 – 2790. URL: https://journals.ametsoc.org/view/journals/mwre/129/11/1520-0493_2001_129_2776_ddfobe_2.0.co_2.xml, doi:10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2.

Held, I., Pierrehumbert, R., Garner, S., Swanson, K., 1995. Surface quasi-geostrophic dynamics. Journal of Fluid Mechanics 282, 1–20.

Houtekamer, P., Mitchell, H., 1998. Data assimilation using an ensemble kalman filter technique. Monthly Weather Review - MON WEATHER REV 126. doi:10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2.

Houtekamer, P., Mitchell, H., 2001. A sequential ensemble kalman filter for atmospheric data assimilation. JANUARY 129. doi:10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2.

Houtekamer, P., Mitchell, H., 2005. Ensemble kalman filtering. Quarterly Journal of the Royal Meteorological Society 131, 3269–3289. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1256/qj.05.135, doi:https://doi.org/10.1256/qj.05.135, arXiv:https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1256/qj.05.135.

Houtekamer, P., Mitchell, H., Pellerin, G., Buehner, M., Charron, M., Spacek, L., Hansen, B., 2005. Atmospheric data assimilation with an ensemble kalman filter: Results with

real observations. Monthly Weather Review - MON WEATHER REV 133, 604–620. doi:10.1175/MWR-2864.1.

Kadri Harouna, S., Mémin, E., 2017. Stochastic representation of the Reynolds transport theorem: revisiting large-scale modeling. Computers and Fluids 156, 456–469. URL: https://hal.inria.fr/hal-01394780, doi:10.1016/j.compfluid.2017.08.017.

Kantas, N., Beskos, A., Jasra, A., 2014. Sequential monte carlo methods for high-dimensional inverse problems: A case study for the navier–stokes equations. SIAM/ASA Journal on Uncertainty Quantification 2, 464–489. URL: https://doi.org/10.1137/130930364, doi:10.1137/130930364, arXiv:https://doi.org/10.1137/130930364.

Kloeden, P., Platen, E., 1999. Numerical Solution of Stochastic Differential Equations. Springer, Berlin.

König, H., 1986. Eigenvalue distribution of compact operators with application to integral operators. Linear Algebra and its application 84, 111–122.

Kwiatkowski, E., Mandel, J., 2015. Convergence of the square root ensemble kalman filter in the large ensemble limit. SIAM/ASA Journal on Uncertainty Quantification 3, 1–17. URL: https://doi.org/10.1137%2F140965363, doi:10.1137/140965363.

Lapeyre, G., Klein, P., 2006. Dynamics of the upper oceanic layers in terms of surface quasigeostrophy theory. Journal of physical oceanography 36, 165–176.

Le Dimet, F., Talagrand, O., 1986. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. Tellus 38A, 97–110.

Le Gland, F., Monbet, V., Tran, V., 2011. Large sample asymptotics for the ensemble kalman filter. Handbook on Nonlinear Filtering .

Le Guillou, F., Lahaye, N., Ubelmann, C., Metref, S., Cosme, E., Ponte, A., Le Sommer, J., Blayo, E., Vidard, A., 2021. Joint estimation of balanced motions and internal tides from future wide-swath altimetry. Journal of Advances in Modeling Earth Systems 13, e2021MS002613. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021MS002613, doi:https://doi.org/10.1029/2021MS002613, arXiv:https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2021MS002613. e2021MS002613 2021MS002613.

Lee, Y., Majda, A.J., Qi, D., 2017. Preventing catastrophic filter divergence using adaptive additive inflation for baroclinic turbulence. Monthly Weather Review 145, 669–682.

Lindeberg, T., 1993. Scale-Space Theory in Computer Vision. The Springer International Series in Engineering and Computer Science, Springer US. URL: https://books.google.fr/books?id=u21OX41EbVkC.

Mandel, J., Cobb, L., Beezley, J., 2011. On the convergence of the ensemble kalman filter. Applications of Mathematics , 533–541.

Mémin, E., 2014. Fluid flow dynamics under location uncertainty. Geophysical & Astrophysical Fluid Dynamics 108, 119–146. URL: https://doi.org/10.1080/03091929.2013.836190, doi:10.1080/03091929.2013.836190, arXiv:https://doi.org/10.1080/03091929.2013.836190.

Mezic, I., 2005. Spectral properties of dynamical systems, model reduction and decompositions. Nonlinear Dynamics 41, 309–325.

Palmer, T.N., 2019. Stochastic weather and climate models. Nature Reviews Physics URL: http://doi.org/10.1038/s42254-019-0062-2, doi:10.1038/s42254-019-0062-2.

Papadakis, N., Mémin, E., Cuzol, A., Gengembre, N., 2010. Data assimilation with the weighted ensemble kalman filter. Tellus A 62, 673–697. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0870.2010.00461.x, doi:https://doi.org/10.1111/j.1600-0870.2010.00461.x, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1600-0870.2010.00461.x.

Pham, D., 2001. Stochastic methods for sequential data assimilation in strongly nonlinear systems. Monthly Weather Review 129, 1194 – 1207.

Poterjoy, J., 2016. A localized particle filter for high-dimensional nonlinear systems. Monthly Weather Review 144, 59–76.

Prato, G.D., Zabczyk, J., 1992. Stochastic equations in infinite dimensions. Cambridge University Press.

Raanes, P.N., Bocquet, M., Carrassi, A., 2019. Adaptive covariance inflation in the ensemble kalman filter by gaussian scale mixtures. Quarterly Journal of the Royal Meteorological Society 145, 53–75. URL: https://rmets.onlinelibrary.

wiley.com/doi/abs/10.1002/qj.3386, doi:https://doi.org/10.1002/qj.3386, arXiv:https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.3386.

Raanes, P.N., Carrassi, A., Bertino, L., 2015. Extending the square root method to account for additive forecast noise in ensemble methods. Monthly Weather Review 143, 3857 – 3873. URL: https://journals.ametsoc.org/view/journals/mwre/143/10/mwr-d-14-00375.1.xml, doi:10.1175/MWR-D-14-00375.1.

Reich, S., Cotter, C., 2015. Probabilistic Forecasting and Bayesian Data Assimilation. Cambridge University Press. doi:10.1017/CBO9781107706804.

Reinhardt, M., 2020. Hybrid filters and multi-scale models. doctoralthesis. Universität Potsdam. doi:10.25932/publishup-47435.

Resseguier, V., Li, L., Jouan, G., Dérian, P., Mémin, E., Chapron, B., 2020a. New trends in ensemble forecast strategy: Uncertainty quantification for coarse-grid computational fluid dynamics. Archives of Computational Methods in Engineering 28. doi:10.1007/s11831-020-09437-x.

Resseguier, V., Li, L., Jouan, G., Derian, P., Mémin, E., Chapron, B., 2020b. New trends in ensemble forecast strategy: uncertainty quantification for coarse-grid computational fluid dynamics. Archives of Computational Methods in Engineering , 1886–1784.

Resseguier, V., Mémin, E., Chapron, B., 2017a. Geophysical flows under location uncertainty, Part I Random transport and general models. Geophys. & Astro. Fluid Dyn. 111, 149–176.

Resseguier, V., Mémin, E., Chapron, B., 2017b. Geophysical flows under location uncertainty, Part II Quasi-geostrophy and efficient ensemble spreading. Geophys. & Astro. Fluid Dyn. 111, 177–208. URL: https://hal.inria.fr/hal-01391476, doi:10.1080/03091929.2017.1312101.

Resseguier, V., Mémin, E., Chapron, B., 2017c. Geophysical flows under location uncertainty, Part III SQG and frontal dynamics under strong turbulence conditions. Geophys. & Astro. Fluid Dyn. 111, 209–227. URL: https://hal.inria.fr/hal-01391484, doi:10.1080/03091929.2017.1312102.

Rowley, C., Mezic, I., Bagheri, S., Schlatter, P., Henningson, D., 2009. Spectral analysis of nonlinear flows. J. Fluid Mech. 641, 115–127.

Rubin, D.B., 1987. A noniterative sampling/importance resampling alternative to the data augmentation algorithm for creating a few imputations when fractions of missing information are modest: the sir algorithm. Journal of the American Statistical Association 82, 543–546.

Sakov, P., Bertino, L., 2011. Relation between two common localisation methods for the enkf. Computational Geosciences 15, 225–237. URL: https://doi.org/10.1007/s10596-010-9202-6, doi:10.1007/s10596-010-9202-6.

Schmid, P., 2010. Dynamic mode decomposition of numerical and experimental data. J. Fluid Mech. 656, 5–28.

Shen, Z., Tang, Y., 2015. A modified ensemble kalman particle filter for non-gaussian systems with nonlinear measurement functions. Journal of Advances in Modeling Earth Systems 7, 50–66. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014MS000373, doi:https://doi.org/10.1002/2014MS000373, arXiv:https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2014MS000373.

Shen, Z., Tang, Y., Li, X., 2017. A new formulation of vector weights in localized particle filters. Quarterly Journal of the Royal Meteorological Society 143, 3269–3278. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3180, doi:https://doi.org/10.1002/qj.3180, arXiv:https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.3180.

Simon, E., Debreu, L., Blayo, E., 2011. 4d variational data assimilation for locally nested models : complementary theoretical aspects and application to a 2d shallow water model. International Journal for Numerical Methods in Fluids 66, 135 – 161. doi:10.1002/fld.2244.

Snyder, C., Bengtsson, T., Bickel, P., Anderson, J., 2008. Obstacles to high-dimensional particle filtering. Monthly Weather Review 136, 4629 – 4640. URL: https://journals.ametsoc.org/view/journals/mwre/136/12/2008mwr2529.1.xml, doi:10.1175/2008MWR2529.1.

Tippett, M., Anderson, J., Bishop, C., Hamill, T., Whitaker, J., 2003. Ensemble square root filters. Monthly Weather Review 131, 1485 – 1490.

Whitaker, J., Hamill, T., 2002a. Ensemble data assimilation without perturbed observations. Monthly Weather Review .

Whitaker, J.S., Hamill, T.M., 2002b. Ensemble data assimilation without perturbed observations. Monthly Weather Review 130, 1913 – 1924. URL: https://journals.ametsoc.org/view/journals/mwre/130/7/1520-0493_2002_130_1913_edawpo_2.0.co_2.xml, doi:10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2.

Yang, Y., Mémin, E., 2019. Estimation of physical parameters under location uncertainty using an ensemble^2-expectation-maximization algorithm. QJRMS 145, 418–433.

COLLEGE DOCTORAL BRETAGNE

MATHS, TELECOMS
INFORMATIQUE, SIGNAL
SYSTEMES, ELECTRONIQUE

Université de Rennes

Titre : Assimilation de données pour des modèles océaniques stochastiques

Mot clés : Assimilation de données, modélisation stochastique, systèmes dynamiques, prévision d'ensemble

Résumé : Cette thèse explore différentes techniques d'assimilation de données pour des modèles océaniques, et en particulier les modèles stochastiques. La méthodologie stochastique utilisée s'appelle l'incertitude de position (LU en anglais) et vise à incorporer un caractère stochastique à des modèles géophysiques via une décomposition de la vitesse en une composante grande échelle lisse en temps, ainsi qu'une composante aléatoire fortement oscillante, modélisée par un processus de Wiener cylindrique. Comme le caractère aléatoire du modèle ainsi que la quantification d'incertitude sont cruciaux pour l'assimilation de données, il est exposé dans cette thèse que le modèle LU comporte des avantages indéniables sur le modèle SQG (Surface Quasi-Geostrophic en anglais).
Nous avons dans un premier temps, pour ce modèle SQG, comparé le modèle stochastique aux techniques déterministes d'inflation pour un filtre de Kalman d'ensemble "square-root" localisé. Nous avons obtenu, dans cette première étude, une validation numérique que l'inflation peut être difficile à régler et peut mener à une divergence du filtre en temps fini, et que le modèle stochastique donne de meilleurs performances que les modèles déterministes avec inflation en terme d'erreurs et de qualité de variance d'ensemble.
Une deuxième étude a consisté à la proposition d'une procédure de calibration du bruit stochastique, visant à guider l'essaim de trajectoires vers une région d'intérêt, proche des observations. Cette procédure s'appuie sur le caractère stochastique inhérent de LU et se base sur les transformations de Girsanov. L'ajout de ce terme de guidage a mené à une amélioration significative des résultats dans le cas d'une mauvaise estimation de la condition initiale.
La dernière partie de cette thèse étudie la prédiction d'ensemble sous le point de vue des espaces de Hilbert à noyau auto-repoduisant (RKHS en anglais). Dans ce contexte, l'opérateur de Koopman associé à la dynamique et son adjoint sont tous deux unitaires et uniformément continus, ce qui conduit à l'énoncé d'un théorème spectral adapté aux RKHS. Des méthodes d'assimilation de données sont conçues pour prendre en compte la structure et les propriétés des RKHS, qui justifient notamment un principe de superposition qui est largement utilisé, bien que sujet à caution, pour les filtres de Kalman d'ensemble.

Title: Data assimilation for stochastic ocean models

Keywords: Data assimilation, stochastic modeling, dynamical system, ensemble forecasting

Abstract: This thesis explores data assimilation techniques for ocean models, and particularly for stochastic models. The stochastic framework of interest is called location uncertainty (LU) and aims at incorporating the stochasticity in geophysical systems through a velocity decomposition into a large-scale smooth in time component, and a highly oscillating random velocity, modelized as a cylindrical Wiener process. As model randomness and uncertainty quantification are of upmost importance for data assimilation, LU is shown, in this thesis, to have undeniable advantages on the model at hand, namely the Surface Quasi-Geostrophic (SQG) model.
We first compared, for this SQG model, the stochastic framework with deterministic inflation techniques for a localized ensemble square-root filter. We found in this first study a numerical validation that inflation can be difficult to tune and lead to filter divergence in finite time, and that the stochastic setting performs better than deterministic ones in terms of MSE and spread relevance.
A second study designed a noise-calibration procedure, aiming at guiding the set of realizations towards a region of interest, close to the observations. This procedure relies on the inherent stochasticity of LU and is based on Girsanov transforms. The addition of this extra guiding term was shown to significantly improve the numerical results in the case of bad estimation of the initial condition.
The last part of the thesis studies ensemble forecasts within the framework of reproducing kernel Hilbert spaces (RKHS). In this framework, the Koopman operator attached to the dynamics and its adjoint are both unitary and uniformly continuous, which leads to a RKHS spectral theorem adapted to this framework. Adequate data assimilation techniques are devised, enriched with the RKHS structure and properties, which can in some sense justify the questionable superposition principle widely used in ensemble Kalman filters.