DOCTORAL THESIS

Error Analysis of the Ensemble Square Root Filter for Dissipative Dynamical Systems

Department of Mathematics, Kyoto University

Kota Takeda*

Supervisors

Professor Takashi Sakajo, Kyoto University

Professor Takemasa Miyoshi, Kyoto University & RIKEN

Professor Sebastian Reich, University of Potsdam

*Department of mathematics, Kyoto University, Kitashirakawa Oiwake-cho, Sakyo-ku, Kyoto, JAPAN, email: takeda.kota.53r@st.kyoto-u.ac.jp

Acknowledgements

I am grateful to my supervisor, Professor Takashi Sakajo at Kyoto University, for his valuable advice, extensive time spent in discussions, continual encouragement, and support throughout my research. I also appreciate that he provided many valuable opportunities to discuss my research with other researchers from various fields.

I would also like to express my sincere gratitude to Doctor Takemasa at RIKEN for his gracious hosting of me as a junior research associate of the data assimilation research team at RIKEN Center for Computational Science (R-CCS). I spent valuable time discussing practical topics for my thesis with him and the other team members.

I wish to convey my deep appreciation to Professor Sebastian Reich at University of Potsdam for his gentle hospitality during my stay in Potsdam, which was supported by SGU program of "The Japan Gateway: Kyoto University Top Global Program". His constructive suggestions significantly broadened my research perspective.

Finally, I deeply appreciate the invaluable experiences I gained during my PhD course at Kyoto University. I would like to thank all staff in the Department of Mathematics at Kyoto University and in the data assimilation research team at RIKEN R-CCS for supporting me in various situations.

Contents

1. Introduction

In modern weather prediction, numerical simulations of atmospheric models play a crucial role. However, even if the initial state is estimated with sufficient resolution, the simulated state can be separated from the true state over time due to the chaotic nature of the atmosphere. In other words, a small perturbation to the initial state can grow exponentially fast along the deterministic time evolution [50, 66]. This gives rise to the unpredictability of atmospheric motions over long-term periods. Therefore, it is necessary to correct the simulated state using observation data from the true state. This approach is known as data assimilation. See the comprehensive textbooks [34, 50] for further motivations and issues of data assimilation from meteorology and oceanography.

A simple way to correct the simulated state is to replace it with the observation data. However, this approach is not feasible due to two reasons. First, the observations have insufficient resolution to initialize the numerical simulation [28]. Thus, we must reconstruct the high-resolution state from the low-resolution observation data. Second, the observations always contain measurement noises. Therefore, we need to quantify uncertainty in the state estimation from the noisy observations [80]. For these reasons, the state estimation is often formulated as the Bayesian inverse problems [26, 76]. In this approach, the estimated state is represented by a probability distribution. In addition, the numerical model provides the prior distribution, and it is updated into the posterior distribution using the Bayes' formula with the observation data. See the standard textbooks [63, 76, 78] for mathematical formulations of data assimilation as the Bayesian inverse problems.

The Monte Carlo method is a fundamental approach to the Bayesian inverse problems [63, 76], which recovers the posterior distribution with infinitely many samples. From the perspective of computational costs, it is impossible to evaluate the full posterior distribution in a high-dimensional state space. One approach to avoid this issue is estimating a maximizer of the probability distribution function of the posterior, known as the maximum a posteriori (MAP) estimate or the variational method [63, 76]. Another approach is estimating the mean and covariance of the posterior distribution. These two approaches are equivalent and correspond to the least squares method if the evolution of the state and the observation operator are represented by linear maps, and all noises follow the Gaussian distribution. The corresponding sequential data assimilation algorithm is the Kalman filter (KF) [4, 24, 49]. However, for instance, the motions of geophysical flows are usually modeled by nonlinear dynamics. Then, the ensemble Kalman filter (EnKF) is proposed as a nonlinear extension of the KF [34], which approximates the mean and covariance by a set of states known as an ensemble. The EnKF is a hybrid approach between the Monte Carlo method and the least

squares method, and it represents the uncertainty in the state estimation using a small number of ensemble members. The EnKF is effectively applied to a wide range of data assimilation problems for high-dimensional and nonlinear systems in geophysics [20, 34, 50].

Two major implementations of the EnKF are known. The perturbed observation (PO) method [19, 33] is a stochastic one, and the ensemble square root filter (ESRF) [5, 14, 86] is a deterministic one. Theoretical studies have revealed some fundamental properties common to the PO method and the ESRF [51, 58, 70, 88, 89], such as the boundedness of the ensemble and the convergence to the KF in a large ensemble limit for linear systems with the Gaussian noises. However, the error analyses have only been established for the PO method [54], not for the ESRF method. Therefore, to advance the theoretical analysis of the EnKF, an error analysis for the ESRF is necessary.

Another issue in the mathematical analysis of data assimilation is formulating the model dynamics. The motions of geophysical flows are often modeled by dissipative partial differential equations (PDE) such as the Navier-Stokes equations [23, 38], which are defined on infinite-dimensional state spaces. Moreover, the chaotic nature of the geophysical flows is an essential research subject in weather prediction. Therefore, we consider dissipative dynamical systems on Hilbert spaces to incorporate such chaotic properties into the model dynamics. The typical solution has a bounded trajectory and exhibits chaotic behavior on a compact limit set, known as a global attractor. After formulated in the infinite-dimensional space, the model dynamics is discretized in a finite-dimensional space as mentioned in [27].

We have introduced two issues in the mathematical analysis of the EnKF above. According to them, this thesis aims to discuss the following three topics. Firstly, we review analytical results for the EnKF in various formulations of the model dynamics. Secondly, we prove the error analysis of the ESRF applied to dissipative dynamical systems. This is the main contribution of this thesis. Finally, we indicate future directions for mathematical analysis of the EnKF, comparing the results among various formulations.

The thesis is constructed as follows. Section 2 introduces some notations and recalls from the theories of functional analysis, measures, Bayesian inference, and dynamical systems on Hilbert spaces. In Section 3, we define data assimilation problems in various formulations, classified by stochastic/deterministic, finite/infinite-dimensional, and discrete/continuous-time systems. Section 4 reviews the sequential data assimilation algorithms, the KF, and the variants of the EnKF. We then examine them from the perspective of practical implementations. Section 5 provides the analysis of dissipative dynamical systems and some examples appearing in geophysics. In Section 6, the mathematical analysis of the EnKF is explained, where we review the error analysis of the PO method and other fundamental analyses of the EnKF. Section 7 explains the main result of this thesis. We establish the error analysis of the ESRF for the dissipative

dynamical systems. We also validate the analysis with a numerical example. Section 8 is a summary and discussion of future directions.

2. Preliminary

2.1 Notations

We use uppercase letters, e.g., U, for random variables and lowercase letters, e.g., u, for their realizations or deterministic variables. For $n \in \mathbb{N}$, $u \in \mathbb{R}^n$ is assumed to be a column vector. We use the notation u^i for its i th element, and u^* denotes the transpose of u. We use bold letters, e.g., $\boldsymbol{U}$ or $\boldsymbol{u}$, for a set of vectors.

2.2 Functional analysis

We recall some facts on the theory of functional analysis to handle the state estimation problems on Hilbert spaces. For the details and proofs in this section, see the introductory textbook [25].

2.2.1 Basic theory

Let $\mathcal{H}$ be a Hilbert space endowed with the inner product $\langle \cdot, \cdot \rangle$ and the associated norm $|\cdot|$. We use the same notations even when $\mathcal{H}$ is the Euclidean space. We assume Hilbert spaces are separable. By $\mathcal{L}(\mathcal{H}, \mathcal{G})$, we denote the space of bounded linear operators from $\mathcal{H}$ to another Hilbert space $\mathcal{G}$. Let $I_{\mathcal{H}}$ denote the identity operator on $\mathcal{H}$. For $A \in \mathcal{L}(\mathcal{H}) \coloneqq \mathcal{L}(\mathcal{H}, \mathcal{H})$, $|A|_{\mathcal{L}}$ represents the operator norm of A, $\mathrm{Ran}(A)$ denotes the range of A, and A^* is the adjoint of A. For $u, v \in \mathcal{H}$, we define their product $u \otimes v \in \mathcal{L}(\mathcal{H})$ by $u \otimes v : \mathcal{H} \ni w \mapsto u \langle v, w \rangle \in \mathcal{H}$, which is equivalent to $uv^* = u \otimes v$. We call $U \in \mathcal{L}(\mathcal{H})$ unitary if $U^*U = UU^* = I_{\mathcal{H}}$. Let $\mathcal{L}_{sa}(\mathcal{H})$ denote a set of self-adjoint operators in $\mathcal{L}(\mathcal{H})$, i.e., $A^* = A$ for $A \in \mathcal{L}_{sa}(\mathcal{H})$. We define important concepts for $\mathcal{L}_{sa}(\mathcal{H})$ as follows.

Definition 2.1. *Let $A \in \mathcal{L}_{sa}(\mathcal{H})$.*

(a) An operator A is said to be positive semi-definite, denoted by $A \succeq 0$, if $\langle u, Au \rangle \geq 0$ for all $u \in \mathcal{H}$.

(b) An operator A is said to be positive definite, denoted by $A \succ 0$, if there exists $c > 0$ such that $\langle u, Au \rangle \geq c|u|^2$ for all $u \in \mathcal{H}$.

Remark 2.2. *Definition 2.1 (b) is different from the conventional definition of the positive definiteness. It is often said to be bounded from below, which implies that A is invertible.*

For a positive semi-definite $A \in \mathcal{L}_{sa}(\mathcal{H})$, a square root $A^{\frac{1}{2}} \succeq 0$ is uniquely well-defined [25]. We thus define a weighted norm $|\cdot|_A = |A^{-1/2} \cdot|$ on $\mathcal{H}$ for $A \succ 0$. For $A, B \in \mathcal{L}_{sa}(\mathcal{H})$, the order $A \succ$ (resp. $\succeq$) B means $A - B \succ$ (resp. $\succeq$) 0. We use the following inequality to estimate operator norms.

Lemma 2.3 ([25]). *If $A \succeq B$, then $|A|_{\mathcal{L}} \geq |B|_{\mathcal{L}}$.*

For a linear operator $A : \mathcal{H} \to \mathcal{H}$, we denote the spectrum of A by

$$\sigma(A) = \{\lambda \in \mathbb{C} \mid (\lambda I - A)^{-1} \notin \mathcal{L}(\mathcal{H})\}.$$

and the resolvent set of A by

$$\rho(A) = \mathbb{C} \setminus \sigma(A).$$

If $A \in \mathcal{L}_{sa}(\mathcal{H})$ then $\sigma(A) \subset \mathbb{R}$. Moreover, if $A \succeq 0$ then $\sigma(A) \subset [0, \infty)$. We also denote the spectral radius of A by

$$r(A) = \sup_{\lambda \in \sigma(A)} |\lambda|.$$

In general, $r(A) \leq |A|_{\mathcal{L}}$, i.e., $|\lambda| \leq |A|_{\mathcal{L}}$ for any $\lambda \in \sigma(A)$. Thus, if $|\lambda| > |A|_{\mathcal{L}}$, then $\lambda \in \rho(A)$ by taking the contrapositive. The following fact is well known.

Proposition 2.4 ([25]). *If $A \in \mathcal{L}(\mathcal{H})$ is normal, i.e., $AA^* = A^*A$, then we have*

$$r(A) = |A|_{\mathcal{L}}.$$

Note that if A is self-adjoint or unitary, then it is normal.

The spectrum of the product of two self-adjoint operators is estimated as follows.

Proposition 2.5 ([43]). *Let $A, B \in \mathcal{L}_{sa}(\mathcal{H})$ and $B \succeq 0$, then we have the following relationships.*

(1) $\sigma(AB) = \sigma(BA) = \sigma(B^{\frac{1}{2}}AB^{\frac{1}{2}})$.

(2) If $A \succeq 0$, $\sigma(AB) \subset [m(A)m(B), M(A)M(B)]$ where $m(A) = \inf \sigma(A), M(A) = \sup \sigma(A)$.

The following lemma is useful to estimate the operator norm of the inverse operator of $A \in \mathcal{L}(\mathcal{H})$.

Lemma 2.6. *For $A \in \mathcal{L}(\mathcal{H})$, if $\sigma_0 = \inf_{\lambda \in \sigma(A)} |\lambda| > 0$, then $A^{-1} \in \mathcal{L}(\mathcal{H})$ and*

$$|A^{-1}|_{\mathcal{L}} \leq \frac{1}{\sigma_0}. \tag{2.1}$$

To prove this, we prepare the following lemma.

Lemma 2.7. *Let $\lambda \in \sigma(A)$ for $A \in \mathcal{L}(\mathcal{H})$. Then,*

$$|Av| \geq |\lambda||v| \tag{2.2}$$

for any $v \in \mathcal{H}$.

Proof. We prove it by contradiction. Suppose that there is $v \in \mathcal{H}$ such that

$$|Av| < |\lambda||v|. \tag{2.3}$$

This implies $|A|_{\mathcal{L}} < |\lambda|$, hence $\lambda \in \rho(A)$. This contradicts the assumption that $\lambda \in \sigma(A)$, hence (2.3) is false. We thus have (2.2). □

Proof of Lemma 2.6. From $\sigma_0 > 0$, we have $0 \in \rho(A)$, hence $A^{-1} \in \mathcal{L}(\mathcal{H})$. Recall that $|A^{-1}|_{\mathcal{L}} = \sup_{|u|=1} |A^{-1}u|$. For $u \in \mathcal{H}$ with $|u| = 1$, we have from Lemma 2.7 that

$$1 = |u| = |AA^{-1}u| \geq \sigma_0 |A^{-1}u|.$$

Therefore,

$$|A^{-1}u| \leq \frac{1}{\sigma_0},$$

which implies (2.1). □

2.2.2 Compact operators

A linear operator $K : \mathcal{H} \to \mathcal{G}$ is said to be compact if for any bounded sequence $(u_n)_{n\in\mathbb{N}} \subset \mathcal{H}$, the sequence $(Ku_n)_{n\in\mathbb{N}} \subset \mathcal{G}$ contains a convergent subsequence. We denote the space of compact operators by $\mathcal{K}(\mathcal{H}, \mathcal{G})$ and $\mathcal{K}(\mathcal{H}) = \mathcal{K}(\mathcal{H}, \mathcal{H})$. The following fact is important when considering compact operators on infinite-dimensional spaces.

Proposition 2.8. *Let* $\dim(\mathcal{H}) = \infty$. *If* $K \in \mathcal{K}(\mathcal{H})$, *then* $K^{-1} \notin \mathcal{L}(\mathcal{H})$.

From this proposition, if $A, A^{-1} \in \mathcal{L}(\mathcal{H})$, then $A \notin \mathcal{K}(\mathcal{H})$. For instance, $I_{\mathcal{H}} \notin \mathcal{K}(\mathcal{H})$. A self-adjoint and compact operator is unitarily diagonalizable, which yields the spectral decomposition as follows.

Proposition 2.9 (Spectral theorem)**.** *Let* $K \in \mathcal{L}_{sa}(\mathcal{H}) \cap \mathcal{K}(\mathcal{H})$. *Then, there exist eigenvalues* $(\lambda_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ *and an orthonormal basis* $(\phi_n)_{n\in\mathbb{N}} \subset \mathcal{H}$ *consisting of associated eigenvectors such that*

$$K = \sum_{n\in\mathbb{N}} \lambda_n \phi_n \otimes \phi_n.$$

Corollary 2.10. *Let* $K \in \mathcal{K}(\mathcal{H})$. *Then,* $K^*K \succeq 0$ *and* $K^*K \in \mathcal{L}_{sa}(\mathcal{H}) \cap \mathcal{K}(\mathcal{H})$. *Therefore, there exist* $(s_n)_{n\in\mathbb{N}} \subset [0, \infty)$ *and an orthonormal basis* $(\phi_n)_{n\in\mathbb{N}} \subset \mathcal{H}$ *such that*

$$K^*K = \sum_{n\in\mathbb{N}} s_n^2 \, \phi_n \otimes \phi_n.$$

Here, $(s_n)_{n\in\mathbb{N}}$ *are called singular values of* K.

As an important subclass of compact operators, we introduce trace class operators.

Definition 2.11. *For $T \in \mathcal{K}(\mathcal{H})$, it is said to be trace class if*

$$\operatorname{Tr}|T| \coloneqq \sum_{n\in\mathbb{N}} s_n(T) < \infty, \tag{2.4}$$

where $(s_n(T))_{n\in\mathbb{N}}$ are the singular values of T. We denote the set of trace class operators by $\mathcal{K}_1(\mathcal{H})$. Additionally, if $T \succeq 0$, we define the trace of T by

$$\operatorname{Tr} T \coloneqq \operatorname{Tr}|T|.$$

This definition of the trace of an operator is consistent with that of a matrix. Then, we have the following lemma.

Lemma 2.12 ([25])**.** *For $A \in \mathcal{L}(\mathcal{H})$ and $B \in \mathcal{K}_1(\mathcal{H})$, we have*

$$\operatorname{Tr}|AB| \le |A|_{\mathcal{L}} \operatorname{Tr}|B|.$$

Another important sub class of compact operators is the Hilbert-Schmidt class [25].

Definition 2.13. *An operator $A \in \mathcal{L}(\mathcal{H})$ is called a Hilbert-Schmidt operator if $|A|_{HS} < \infty$, where*

$$|A|_{HS} = \left(\sum_{i\in\mathbb{N}} |A\phi_i|\right)^{\frac{1}{2}}$$

with an orthonormal basis $(\phi_i)_{i\in\mathbb{N}}$ of $\mathcal{H}$.

2.2.3 Perturbation theory of eigenvalue problems

In this thesis, we use the perturbation theory of eigenvalue problems of matrices.

Proposition 2.14 ([52, 77])**.** *Suppose the matrix-valued function $S(t) \in \mathbb{R}^{N\times N}$ is self-adjoint and continuously differentiable in an interval I of t. Then, there exist the eigenvalues $\lambda_n(t),\ n = 1, \dots, N$ of $S(t)$ that are continuously differentiable on I.*

The following lemma is used in the analysis of a filtering algorithm in Section 7.

Lemma 2.15 ([30, 32])**.** *Suppose the same condition as Proposition 2.14, there exists a unitary matrix valued function $U(t)$ on I such that*

$$\frac{d}{dt}\lambda_n(t) = \left[U(t)^* \left(\frac{d}{dt}S(t)\right) U(t)\right]_{nn}.$$

Note that $U(t)$ is not differentiable in general [32].

2.3 Calculations for linear operators

2.3.1 Ensemble of vectors

For $m \in \mathbb{N}$, a set of state vectors $v^{(k)} \in \mathcal{H}$ for $k = 1, \dots, m$ is called an ensemble, and m is called the ensemble size. We use the notation $\boldsymbol{V} = [v^{(k)}]_{k=1}^m \in \mathcal{H}^m$ to denote the ensemble. If $\mathcal{H} = \mathbb{R}^l$ for $l \in \mathbb{N}$, $\boldsymbol{V}$ is equivalent to a matrix in $\mathbb{R}^{l \times m}$. For ensembles $\boldsymbol{U} = [u^{(k)}]_{k=1}^m$ and $\boldsymbol{V} = [v^{(k)}]_{k=1}^m \in \mathcal{H}^m$, the ℓ_2-norm $|\boldsymbol{U}|_2$ is defined by

$$|\boldsymbol{U}|_2 = \left(\frac{1}{m} \sum_{k=1}^m |u^{(k)}|^2 \right)^{\frac{1}{2}}, \tag{2.5}$$

and the products $\boldsymbol{U}\boldsymbol{V}^* \in \mathcal{L}(\boldsymbol{U})$ and $\boldsymbol{U}^*\boldsymbol{V} \in \mathbb{R}^{m \times m}$ are given as

$$\boldsymbol{U}\boldsymbol{V}^* = \sum_{k=1}^m u^{(k)} \otimes v^{(k)}, \quad \boldsymbol{U}^*\boldsymbol{V} = \left[\left\langle u^{(i)}, v^{(j)} \right\rangle \right]_{i,j=1}^m.$$

When we write an ensemble consisting of the same vector $u \in \mathcal{H}$ as $u\mathbf{1} = [u, \dots, u] \in \mathcal{H}^m$ with $\mathbf{1} = (1, \dots, 1) \in (\mathbb{R}^m)^*$, it holds that $|u\mathbf{1}|_2^2 = |u|^2$. Moreover, for $x \in \mathbb{R}^m$, $T \in \mathbb{R}^{m \times m}$, and $A \in \mathcal{L}(\mathcal{H})$, we define

$$\begin{aligned}
\boldsymbol{U}x &= \sum_{k=1}^m x^k u^{(k)} \in \mathcal{H}, \\
u + \boldsymbol{U} &= u\mathbf{1} + \boldsymbol{U} = [u + u^{(k)}]_{k=1}^m \in \mathcal{H}^m, \\
\boldsymbol{U}T &= \left[\sum_{l=1}^m u^{(l)} T_{l,k} \right]_{k=1}^m \in \mathcal{H}^m, \\
A\boldsymbol{U} &= \left[Au^{(k)} \right]_{k=1}^m \in \mathcal{H}^m.
\end{aligned}$$

For an ensemble $\boldsymbol{V} = [v^{(k)}]_{k=1}^m \in \mathcal{H}^m$, $\overline{v} = \frac{1}{m} \sum_{k=1}^m v^{(k)}$ is called the ensemble mean and $d\boldsymbol{V} = [v^{(k)} - \overline{v}]_{k=1}^m \in \mathcal{H}^m$ is called the ensemble perturbation. The ensemble $\boldsymbol{V}$ is then decomposed into the mean and the perturbation, $\boldsymbol{V} = \overline{v}\mathbf{1} + d\boldsymbol{V}$. The (unbiased) ensemble covariance $\mathrm{Cov}_m(\boldsymbol{V}) \in \mathcal{L}_{sa}(\mathcal{H})$ is defined by

$$\mathrm{Cov}_m(\boldsymbol{V}) = \frac{1}{m-1} d\boldsymbol{V} d\boldsymbol{V}^*.$$

It is easy to see $\mathrm{Cov}_m(\boldsymbol{V}) = \mathrm{Cov}_m(d\boldsymbol{V})$ and $\mathrm{Cov}_m(\boldsymbol{V}) \succeq 0$.

The following lemma shows a fundamental property of the ensemble perturbation.

Lemma 2.16. *For any ensemble perturbation $d\boldsymbol{V}$ of $\boldsymbol{V} \in \mathcal{H}^m$, we have*

$$d\boldsymbol{V}\mathbf{1}^* = 0 \in \mathcal{H}. \tag{2.6}$$

Proof. The equality (2.6) is obtained from the definition of the ensemble mean.

$$d\boldsymbol{V}\mathbf{1}^* = \sum_{k=1}^{m}(v^{(k)} - \overline{v}) \cdot 1 = m\left(\frac{1}{m}\sum_{k=1}^{m} v^{(k)} - \frac{1}{m}\sum_{k=1}^{m}\overline{v}\right) = m(\overline{v} - \overline{v}) = 0.$$

□

We also have the following lemma providing equivalent representations of the ℓ_2-norm (2.5) of an ensemble $\boldsymbol{V} \in \mathcal{H}^m$.

Lemma 2.17. *The ℓ_2-norm for $\boldsymbol{V} \in \mathcal{H}^m$ satisfies*

$$|\boldsymbol{V}|_2^2 = \frac{1}{m}\operatorname{Tr}\boldsymbol{V}^*\boldsymbol{V} = \frac{1}{m}\operatorname{Tr}\boldsymbol{V}\boldsymbol{V}^* = |\overline{v}|^2 + |d\boldsymbol{V}|_2^2. \tag{2.7}$$

Proof. The first equality is derived from the definition of $|\boldsymbol{V}|_2$.

$$|\boldsymbol{V}|_2^2 = \frac{1}{m}\sum_{k=1}^{m}|v^{(k)}|^2 = \frac{1}{m}\sum_{k=1}^{m}\left\langle v^{(k)}, v^{(k)}\right\rangle = \frac{1}{m}\operatorname{Tr}\boldsymbol{V}^*\boldsymbol{V}.$$

Let $(\phi_i)_{i\in\mathbb{N}}$ be a complete orthonormal basis of $\mathcal{H}$, we have

$$\begin{aligned}|\boldsymbol{V}|_2^2 &= \frac{1}{m}\sum_{k=1}^{m}|v^{(k)}|^2 = \frac{1}{m}\sum_{k=1}^{m}\sum_{i\in\mathbb{N}}\left\langle v^{(k)}, \phi_i\right\rangle^2 = \frac{1}{m}\sum_{i\in\mathbb{N}}\sum_{k=1}^{m}\left\langle v^{(k)}, \phi_i\right\rangle^2 \\ &= \frac{1}{m}\sum_{i\in\mathbb{N}}\sum_{k=1}^{m}\left\langle \phi_i, (v^{(k)}\otimes v^{(k)})\phi_i\right\rangle = \frac{1}{m}\operatorname{Tr}\boldsymbol{V}\boldsymbol{V}^*.\end{aligned}$$

Owing to the relation $d\boldsymbol{V}\mathbf{1}^* = 0$ in Lemma 2.16, we have $\boldsymbol{V}\boldsymbol{V}^* = \overline{v}\mathbf{1}\mathbf{1}^*\overline{v}^* + d\boldsymbol{V}d\boldsymbol{V}^* = m\overline{v}\,\overline{v}^* + d\boldsymbol{V}d\boldsymbol{V}^*$. Hence, we obtain $\frac{1}{m}\operatorname{Tr}\boldsymbol{V}\boldsymbol{V}^* = |\overline{v}|^2 + |d\boldsymbol{V}|_2^2$. □

2.3.2 Inverse of operators

We use the following technical lemmas to calculate the inverse of an operator. See [44, 75] for other identities.

Lemma 2.18 (Woodbury identity [39, 80])**.** *Let $\mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces and $A : \mathcal{H}_1 \to \mathcal{H}_1, B : \mathcal{H}_1 \to \mathcal{H}_2, C : \mathcal{H}_2 \to \mathcal{H}_2$, and $D : \mathcal{H}_2 \to \mathcal{H}_1$ be linear operators. If A, C, and $A + BCD$ are invertible, then*

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}. \tag{2.8}$$

Here, we denote the identity operator by $I = I_{\mathcal{H}}$. Then, we show the following lemmas.

Lemma 2.19. *Let $A \in \mathcal{L}(\mathcal{H})$. If $I + A$ is invertible, then we have*

$$(I + A)^{-1} = I - (I + A)^{-1}A. \tag{2.9}$$

Moreover, if $A \in \mathcal{L}_{sa}(\mathcal{H})$ and $A \succeq 0$, then

$$0 \preceq A(A + I)^{-1} = (A + I)^{-1}A \preceq I, \quad 0 \preceq (A + I)^{-1} \preceq I. \tag{2.10}$$

Proof. The equality (2.9) is easily confirmed by

$$LHS = (I + A)^{-1}(I + A - A) = I - (I + A)^{-1}A.$$

For (2.10), the equality (2.9) yields $A(A + I)^{-1} = (A + I)^{-1}A$. The inequalities follow from the spectral mapping theorem [25]. □

Lemma 2.20. *Let $\Gamma : \mathcal{H} \to \mathcal{H}$ be invertible and $V \in \mathcal{H}^m$. Then, the operator $VV^* + \Gamma$ is invertible, and*

$$(I + V^*\Gamma^{-1}V)^{-1}V^*\Gamma^{-1} = V^*(VV^* + \Gamma)^{-1}. \tag{2.11}$$

Furthermore,

$$(I + V^*\Gamma^{-1}V)^{-1} = I - V^*(VV^* + \Gamma)^{-1}V. \tag{2.12}$$

Proof. The operator $VV^* + \Gamma$ is invertible owing to $VV^* \succeq 0$ and $\Gamma \succ 0$. Then, we have

$$V^*\Gamma^{-1}(VV^* + \Gamma) = V^*\Gamma^{-1}VV^* + V^* = (I + V^*\Gamma^{-1}V)V^*,$$

which is equivalent to (2.11). For (2.12), using (2.11), we get

$$(I + V^*\Gamma^{-1}V)^{-1}V^*\Gamma^{-1}V = V^*(VV^* + \Gamma)^{-1}V.$$

Therefore,

$$I - V^*(VV^* + \Gamma)^{-1}V = I - (I + V^*\Gamma^{-1}V)^{-1}V^*\Gamma^{-1}V = (I + V^*\Gamma^{-1}V)^{-1},$$

where the last equality follows from (2.9). □

Lemma 2.21. *Let $A \in \mathcal{L}(\mathcal{H})$ be invertible. If $U \in \mathcal{L}(\mathcal{H})$ be unitary, then*

$$UA^{-1}U^* = (UAU^*)^{-1},$$

and for diagonal $\Sigma \succ 0$, we have

$$\Sigma A^{-1}\Sigma = (\Sigma^{-1}A\Sigma^{-1})^{-1}.$$

Proof. Owing to $U^{-1} = U^*$ and Σ is invertible, both of the equalities follow from the fact that $(AB)^{-1} = B^{-1}A^{-1}$ for invertible $A, B \in \mathcal{L}(\mathcal{H})$. □

2.4 Probability theory

We need probability theory to quantify uncertainties emerged in the state estimation problems. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space consisting of a sample space Ω, a σ-algebra $\mathscr{F}$, and a probability measure $\mathbb{P}$. By $\mathbb{E}[\,\cdot\,]$, we express the expectation with respect to this probability space. For a family of subsets $E \subset 2^\Omega$, we denote the smallest σ-algebra containing E by $\sigma(E)$. Let $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$ be a measurable space for a Hilbert space $\mathcal{H}$ and its Borel σ-algebra $\mathcal{B}(\mathcal{H}) = \sigma(\{O \subset \mathcal{H} \mid O : \mathit{open}\})$, and $\mathcal{M}_1(\mathcal{H})$ denote the set of probability measures on this space. For a Banach space $\mathcal{X}$, a measurable function $f : \mathcal{H} \to \mathcal{X}$ and $\mu \in \mathcal{M}_1(\mathcal{H})$, we denote $\mathbb{E}_\mu[f] = \int_{\mathcal{H}} f(u)\, d\mu(u)$ in the meaning of the Pettis integral [3, 80]. For $\mu \in \mathcal{M}_1(\mathcal{H})$, the mean of μ is defined by

$$\varpi_\mu = \mathbb{E}_\mu[x] = \int_{\mathcal{H}} x\, d\mu(x) \in \mathcal{H},$$

and the covariance of μ is defined by

$$C_\mu = \int_{\mathcal{H}} (x - \varpi_\mu) \otimes (x - \varpi_\mu)\, d\mu(x) \in \mathcal{L}_{sa}(\mathcal{H}).$$

A random variable U is a measurable map $U : (\Omega, \mathscr{F}, \mathbb{P}) \to (\mathcal{H}, \mathcal{B}(\mathcal{H}))$. It induces the push-forward measure $\mathbb{P}^U = U_*\mathbb{P} \in \mathcal{M}_1(\mathcal{H})$ called the image measure of U and the σ-algebra associated with U defined by $\sigma(U) = \sigma(\{U^{-1}(E) \mid E \in \mathcal{B}(\mathcal{H})\})$. Similarly, the mean and the covariance of U are given by

$$\varpi_U = \mathbb{E}[U] = \int_{\mathcal{H}} u\, d\mathbb{P}^U(u) = \int_\Omega U(\omega)\, d\mathbb{P}(\omega), \quad C_U = \mathbb{E}[(U - \varpi_U) \otimes (U - \varpi_U)].$$

In addition, for a sub-σ-algebra $\mathscr{G} \subset \mathscr{F}$, the conditional expectation of U with respect to $\mathscr{G}$ is denoted by $\mathbb{E}[U \,|\, \mathscr{G}]$.

Let a time index set $\mathcal{T} = \mathbb{N} \cup \{0\}$ or $[0, \infty) \subset \mathbb{R}$ and a stochastic process $U : \mathcal{T} \times \Omega \to \mathcal{H}$. A family of sub-$\sigma$-algebras $(\mathscr{F}_t)_{t \in \mathcal{T}}$ is called a filtration if $s \leq t \Rightarrow \mathscr{F}_s \subset \mathscr{F}_t$. The filtration associated with the stochastic process U is defined by $\mathscr{F}_t^U = \sigma(\{U_s^{-1}(E) \subset \Omega \mid E \subset \mathcal{B}(\mathcal{H}), s \leq t, s \in \mathcal{T}\})$ for $t \in \mathcal{T}$. The expectation conditioned on $U_0 = u \in \mathcal{H}$ is denoted by $\mathbb{E}^u[f(U_t)] = \mathbb{E}[f(U_t) \,|\, U_0 = u]$ for $t \in \mathcal{T}$ and an integrable function $f : \mathcal{H} \to \mathbb{R}$.

Let μ, ν be measures on a measurable space $(\mathcal{X}, \mathscr{F})$. If $\nu(E) = 0$ for any $E \in \mathscr{F}$ with $\mu(E) = 0$, then ν is said to be absolutely continuous with respect to μ, denoted by $\nu \ll \mu$. A measure space $(\mathcal{X}, \mathscr{F}, \mu)$ is called σ-finite if $\mathcal{X}$ can be written by a countable union of elements of $\mathscr{F}$ with each of μ-finite measure.

Proposition 2.22 (Radon-Nikodým's theorem [80])**.** *Suppose that μ and ν are σ-finite measures on a measurable space $(\mathcal{X}, \mathscr{F})$ and that $\nu \ll \mu$. Then, there exists a measurable function $\rho : \mathcal{X} \to [0, \infty]$ such that, for all measurable function $f : \mathcal{X} \to \mathbb{R}$*

and all $E \in \mathscr{F}$,

$$\int_E f d\nu = \int_E f\rho \, d\mu,$$

whenever either integral exists. Furthermore, any two functions ρ *with this property are equal* μ*-almost everywhere.*

The measurable function ρ in Proposition 2.22 is called the Radon-Nikodým derivative, often denoted by $\rho = \frac{d\nu}{d\mu}$. The Radon-Nikocým derivative plays an important role in the Bayesian inference, which is explained later.

Let us consider a finite dimensional case. For $l \in \mathbb{N}$, if $\mu \in \mathcal{M}_1(\mathbb{R}^l)$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^l$, the Radon-Nikodým derivative is called the probability density function (PDF), and we denote it by p_μ. Similarly, for a random variable U, the PDF of $\mathbb{P}^U$ is denoted by p_U. The Gaussian measure on $\mathbb{R}^l$ is defined by its PDF.

Definition 2.23 (Gaussian measure on $\mathbb{R}^l$)**.** *Let* $\varpi \in \mathbb{R}^l$ *and* $C \in \mathcal{L}_{sa}(\mathbb{R}^l)$ *with* $C \succ 0$*. The Gaussian measure with mean* ϖ *and covariance* C *is defined by its PDF as*

$$p(x) = \frac{1}{\sqrt{\det C(2\pi)^l}} \exp\left(-\frac{1}{2}|x-\varpi|_C^2\right), \tag{2.13}$$

and it is denoted by $\mathcal{N}(\varpi, C) \in \mathcal{M}_1(\mathbb{R}^l)$*.*

We can introduce the following metrics between probability measures on $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$, see [37] for other metrics and relationships among them.

Definition 2.24 ([80])**.** *Let* $\mu, \nu \in \mathcal{M}_1(\mathcal{H})$*. The total variation distance is defined by*

$$d_{TV}(\mu,\nu) = \sup\left\{|\mu(A) - \nu(A)| \mid A \in \mathcal{B}(\mathcal{H})\right\}.$$

The Hellinger distance is given by

$$d_H(\mu,\nu) = \left(\int_{\mathcal{H}} \left|\sqrt{\frac{d\mu}{d\rho}} - \sqrt{\frac{d\nu}{d\rho}}\right|^2 d\rho\right)^{\frac{1}{2}},$$

independent to a reference measure $\rho \in \mathcal{M}_1(\mathcal{H})$*.*

The following inequalities are useful in estimating the differences between measures or moments.

Proposition 2.25 ([37, 80])**.** *For* $\mu, \nu \in \mathcal{M}_1(\mathcal{H})$*,*

$$d_H(\mu,\nu)^2 \le d_{TV}(\mu,\nu) \le 2 d_H(\mu,\nu).$$

Proposition 2.26 ([80]). *Let $\mu, \nu \in \mathcal{M}_1(\mathcal{H})$, $\mathcal{X}$ be a Banach space with a norm $\|\cdot\|$, and $f : \mathcal{H} \to \mathcal{X}$ be a measurable function. Suppose that $\mathbb{E}_\mu[\|f\|^2]$ and $\mathbb{E}_\nu[\|f\|^2] < \infty$. Then, we have*

$$\|\mathbb{E}_\mu[f] - \mathbb{E}_\nu[f]\| \leq 2\sqrt{\mathbb{E}_\mu[\|f\|^2] + \mathbb{E}_\nu[\|f\|^2]}\, d_H(\mu, \nu).$$

We finally review important facts about measures on infinite-dimensional Hilbert spaces. For more details, see Chapter 2 of [80] and references therein.

Proposition 2.27 (Lebesgue measures on Hilbert spaces [80]). *Suppose that a measure μ on an infinite-dimensional Hilbert space $\mathcal{H}$ is invariant under all translations, and is locally finite, i.e., for any $u \in \mathcal{H}$, there exists a measurable O_u such that $u \in O_u$ and $\mu(O_u) < \infty$. Then, μ is the zero measure.*

From this proposition, we cannot define the Lebesgue measure on $\mathcal{H}$. However, the Gaussian measure on $\mathcal{H}$ is well-defined.

Definition 2.28 (Gaussian measure on $\mathcal{H}$). *A measure μ on $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$ is said to be a Gaussian measure if the push-forward measure $\ell_*\mu$ is a (non-degenerate) Gaussian measure on $\mathbb{R}$ for any continuous linear functional $\ell : \mathcal{H} \to \mathbb{R}$.*

Proposition 2.29 (Sazanov's theorem [80]). *Let μ be a Gaussian measure on $\mathcal{H}$ with mean zero. Then, its covariance $C_\mu \in \mathcal{K}_1(\mathcal{H})$ and*

$$\operatorname{Tr} C_\mu = \int_{\mathcal{H}} |x|^2 \, d\mu(x).$$

Conversely, if $C \in \mathcal{L}_{sa} \cap \mathcal{K}_1(\mathcal{H})$ with $\langle Cx, x\rangle > 0$ for any $x \in \mathcal{H}$, then there exists a Gaussian measure μ on $\mathcal{H}$ with covariance $C_\mu = C$.

Proposition 2.29 implies that the covariance of a Gaussian measure should be trace class. We denote $\mu = \mathcal{N}(0, C)$ in Proposition 2.29. Moreover, for the shifted Gaussian random variable $X_\varpi = \varpi + X_0$ with $\varpi \in \mathcal{H}$ and $X_0 \sim \mathcal{N}(0, C)$, we denote $\mathbb{P}^{X_\varpi} = \mathcal{N}(\varpi, C)$. We also use the same notation for a degenerate Gaussian measure for the covariance $C \succeq 0$.

2.5 Bayesian inference

Bayesian inference provides a mathematical framework, to estimate a state u from a noisy observation y and to quantify its uncertainty based on Bayes' theorem. See [26, 79] for more details about the concepts and formulations in infinite-dimensional spaces.

2.5.1 Bayes' theorem

We first introduce Bayes' theorem in the Euclidean spaces $\mathcal{H} = \mathbb{R}^{N_u}$ and $\mathcal{Y} = \mathbb{R}^{N_y}$. We assume that we know the conditional PDF $p_Y(y \,|\, u)$. For instance, if the noisy observation y is generated by $y \sim \mathcal{N}(u, R)$ for $R \in \mathcal{L}_{sa}(\mathbb{R}^{N_y})$ with $R \succ 0$, we have the PDF $p_Y(y \,|\, u)$ as (2.13).

Proposition 2.30 (Bayes' theorem [63, 76])**.** *For an observation $y \in \mathbb{R}^{N_y}$, the conditional PDF of U is given by Bayes' formula,*

$$p_{U|Y}(u \,|\, y) = \frac{p_Y(y \,|\, u) p_U(u)}{p_Y(y)}, \tag{2.14}$$

where $p_U(u)$ is the PDF of U and $p_Y(y) = \int_{\mathbb{R}^{N_y}} p_Y(y \,|\, u) p_U(u)\, du$.

In the context of Bayesian inference, $p_U(u)$ and $p_{U|Y}(u \,|\, y)$ are called by the prior and posterior distributions (densities), respectively. The prior distribution $p_U(u)$ represents the uncertainty in the initial estimate of the state u. For given observation y, the prior distribution $p_U(u)$ is updated into the posterior distribution $p_U(u \,|\, y)$ by multiplying the likelihood $p_Y(y \,|\, u)$ as in Proposition 2.30.

$$p_U(u) \to p_{U|Y}(u \,|\, y) \propto p_Y(y \,|\, u) p_U(u).$$

The posterior distribution $p_{U|Y}(y \,|\, u)$ reflects the uncertainty in the estimate of u after incorporating the information from observation y into the prior knowledge. Figure 1 illustrates Bayes' formula for one-dimensional state and observation space.

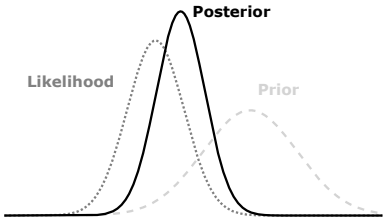

Figure 1: Bayes' formula.

The following lemma provides the explicit formula for the Gaussian posterior distribution.

Lemma 2.31 (Gaussian conditioning [80])**.** *Let $(U, Y) \sim \mathcal{N}(\varpi, C)$ be a joint Gaussian distribution on $\mathbb{R}^{N_u \times N_y}$ with the mean*

$$\varpi = \begin{bmatrix} \varpi_1 \\ \varpi_2 \end{bmatrix} \in \mathbb{R}^{N_u + N_y}$$

and the covariance

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{12}^* & C_{22} \end{bmatrix} \succ 0$$

in block form. The conditional distribution of U provided $Y = y$ is the Gaussian distribution given by

$$\mathbb{P}^U(\,\cdot\,|\,Y = y) \sim \mathcal{N}(\varpi_1 + C_{12}C_{22}^{-1}(y - \varpi_2), C_{11} - C_{12}C_{22}^{-1}C_{12}^*).$$

The Gaussian conditioning is essential in developing the approximated Gaussian algorithms in data assimilation.

Since the Lebesgue measure does not exist on infinite-dimensional Hilbert spaces according to Proposition 2.27, we cannot consider density functions with respect to it. We thus generalize (2.14) in terms of the Radon-Nikodým derivative of the posterior distribution with respect to the prior distribution. If $\dim(\mathcal{Y}) < \infty$, we can define the posterior distribution as follows.

Proposition 2.32 (Generalized Bayes' formula [26, 29, 80])**.** *Let $\mathcal{Y} = \mathbb{R}^{N_y}$, $h : \mathcal{H} \to \mathcal{Y}$ be continuous, and $\mu \in \mathcal{M}_1(\mathcal{Y})$ with its PDF p_μ be the distribution of observation noises, then the posterior distribution $\mu^y(du) = \mathbb{P}(du\,|\,y)$ is absolutely continuous with respect to the prior distribution $\mu_0 \in \mathcal{M}_1(\mathcal{H})$ and its Radon-Nikodým derivative is given by*

$$\frac{d\mu^y}{d\mu_0}(u) \propto \exp\left(-\Phi(u; y)\right), \tag{2.15}$$

where $\Phi(u; y) = -\log(p_\mu(y - h(u)))$.

2.5.2 Well-posedness of the posterior distribution

The Bayesian inverse problem provides a continuous posterior distribution with respect to the observation data. This is known as the well-posedness of the posterior distribution. Let $\mathcal{H}$ and $\mathcal{Y}$ be two Hilbert spaces with norms $|\cdot|_{\mathcal{H}}$ and $|\cdot|_{\mathcal{Y}}$ respectively. Here, we define the posterior by a potential function and impose assumptions on it.

Assumption 2.33. *Let $\Phi(\,\cdot\,;\,\cdot\,) : \mathcal{H} \times \mathcal{Y} \to \mathbb{R}$.*

(1) For any $\epsilon, r > 0$, there exists $M = M(\epsilon, r) \in \mathbb{R}$ such that

$$\Phi(u; y) \geq M - \epsilon|u|_{\mathcal{H}}^2, \qquad u \in \mathcal{H}, |y|_{\mathcal{Y}} < r.$$

(2) For any $r > 0$, there exists $K = K(r) > 0$ such that

$$\Phi(u; y) \leq K, \qquad |u|_{\mathcal{H}}, |y|_{\mathcal{Y}} < r.$$

(3) For any $r > 0$, there exists $L = L(r) > 0$ such that

$$|\Phi(u_1; y) - \Phi(u_2; y)| \le L|u_1 - u_2|_{\mathcal{H}}, \qquad |u_1|_{\mathcal{H}}, |u_2|_{\mathcal{H}}, |y|_{\mathcal{Y}} < r.$$

(4) For any $\epsilon, r > 0$, there exists $C = C(\epsilon, r) > 0$ such that

$$|\Phi(u; y_1) - \Phi(u; y_2)| \le \exp(\epsilon|u|_{\mathcal{H}}^2 + C)|y_1 - y_2|_{\mathcal{Y}}, \qquad |u|_{\mathcal{H}}, |y_1|_{\mathcal{Y}}, |y_2|_{\mathcal{Y}} < r.$$

Example 2.34. *For $\mathcal{Y} = \mathbb{R}^{N_y}$, $H \in \mathcal{L}(\mathcal{H}, \mathcal{Y})$, and $R \in \mathbb{R}^{N_y \times N_y}$ with $R \succ 0$, we consider a potential function*

$$\Phi(u; y) = |y - Hu|_R^2.$$

Then, $\Phi(u; y)$ satisfies Assumption 2.33. This example is introduced from a Gaussian likelihood. See Chapter 6 of [80] for more general examples.

Proposition 2.35 (Well-posedness of the Bayesian inverse problem [26, 80])**.** *Suppose that Φ satisfies Assumption 2.33 and μ_0 is a Gaussian distribution on $\mathcal{H}$. Then, for any $y \in \mathcal{Y}$, the posterior distribution,*

$$\frac{d\mu^y}{d\mu_0} = Z(y)^{-1} \exp(-\Phi(u; y)), \quad Z(y) = \int_{\mathcal{H}} \exp(-\Phi(u; y))\, d\mu_0(u),$$

is well-defined. Furthermore, for any $r > 0$, there exists $C = C(r) > 0$ such that

$$d_H(\mu^{y_1}, \mu^{y_2}) \le C(r)|y_1 - y_2|_{\mathcal{Y}}, \qquad |y_1|_{\mathcal{Y}}, |y_2|_{\mathcal{Y}} \le r.$$

We note that the condition (1) and (2) in Assumption 2.33 implies the finiteness and positivity of the normalizing constant $Z(y)$ in Proposition 2.35, respectively. Proposition 2.35 implies the local Lipschitz continuity of the posterior distribution with respect to the observation data. In other words, the Bayesian inverse problem provides a robust estimation of the state from uncertain data. From Proposition 2.26 and Proposition 2.35, we have the following corollary.

Corollary 2.36. *Let $\mathcal{X}$ be a Banach space with a norm $\|\cdot\|_{\mathcal{X}}$ and $f : \mathcal{H} \to \mathcal{X}$. Suppose that $\mathbb{E}_{\mu_0}[\|f\|_{\mathcal{X}}^2] < \infty$, then for any $r > 0$, there exists $C = C(r)$ such that*

$$\|\mathbb{E}_{\mu^{y_1}}[f] - \mathbb{E}_{\mu^{y_2}}[f]\|_{\mathcal{X}} \le C|y_1 - y_2|_{\mathcal{Y}}, \qquad |y_1|_{\mathcal{Y}}, |y_2|_{\mathcal{Y}} \le r.$$

By taking $\mathcal{X} = \mathcal{H}$, $\|\cdot\|_{\mathcal{X}} = |\cdot|$, and $f(u) = u$ in Corollary 2.36, the mean of the posterior distribution is continuous with respect to y. On the other hand, the mode of the posterior distribution is not continuous with respect to y in general [63].

2.6 Dynamical systems

We use the theory of dynamical systems to describe mathematical models in data assimilation. In particular, it is essential to consider partial differential equations (PDE) as infinite-dimensional dynamical systems [59, 85]. See also the comprehensive textbook of the theory for finite-dimensional dynamical systems [53]. Let $\mathcal{H}$ be a Hilbert space. A dynamical system is defined by a semigroup on $\mathcal{H}$.

Definition 2.37 (Semigroup)**.** *A semigroup on $\mathcal{H}$ is a continuous family $\{\Psi_t \mid t \geq 0\}$ of mappings from $\mathcal{H}$ to itself satisfying*

(1) $\Psi_0 = id_{\mathcal{H}}$*;*

(2) $\Psi_{t+s} = \Psi_t \circ \Psi_s$ *for all* $t, s \geq 0$*;*

(3) $\Psi_t(u_0)$ *is continuous with respect to* t *and* u_0*.*

A semigroup is often generated by an ordinal differential equation (ODE) or a PDE. For example, let $u(t, x)$ be the solution to a PDE with an initial condition $u(0, x) = u_0(x)$ for $u_0 \in \mathcal{H}$. If $u(t, \cdot\,) \in \mathcal{H}$, we can define $\Psi_t : \mathcal{H} \to \mathcal{H}$ by

$$\Psi_t(u_0)(\cdot\,) = u(t, \cdot\,).$$

Hence, in principle, we assume the well-posedness (i.e., the existence and the uniqueness of the solution and its continuous dependence on the initial condition) of the model equation so that the solution generates a semigroup. For ODEs, there is a sufficient condition for the existence of a semigroup.

Proposition 2.38 (Picard-Lindelöf [59])**.** *Let $\mathcal{F} : \mathbb{R}^{N_u} \to \mathbb{R}^{N_u}$ be a locally Lipschitz, i.e., for any $r > 0$, there exists $L = L(r) > 0$ such that*

$$|\mathcal{F}(u) - \mathcal{F}(v)| \leq L|u - v|$$

for all $|u|, |v| \leq r$. Then, there exists a unique solution to the ODE

$$\frac{du}{dt} = \mathcal{F}(u), \quad u(0) = u_0 \in \mathbb{R}^{N_u},$$

on a time interval $[0, T)$ with $T = T(u_0) > 0$.

The following concepts are fundamental for characterizing the long-term behavior of dynamical systems.

Definition 2.39. *Let Ψ_t be a dynamical system on $\mathcal{H}$.*

(1) For $X \subset \mathcal{H}$, we call X is invariant if $\Psi_t(X) = X$ for all $t \geq 0$.

(2) *For $X, B \subset \mathcal{H}$, we call X attracts B if*

$$\operatorname{dist}(\Psi_t(B), X) \to 0 \quad (t \to \infty),$$

where $\operatorname{dist}(A, B) = \sup_{a \in A} \inf_{b \in B} |a - b|$. Moreover, we call X is attracting if it attracts all bounded subset $B \subset \mathcal{H}$.

(3) *For $\mathscr{A} \subset \mathcal{H}$, we call $\mathscr{A}$ a global attractor if it is compact, invariant, and attracting.*

A global attractor satisfies the following properties.

Proposition 2.40 ([59])**.** *Let Ψ_t be a dynamical system on $\mathcal{H}$.*

(1) *A global attractor $\mathscr{A}$ of Ψ_t is unique.*

(2) *The global attractor $\mathscr{A}$ is the maximal compact invariant set and the minimal attracting set.*

(3) *There exists a global attractor $\mathscr{A}$ if and only if there exists a compact attracting set.*

Instead of the existence of an attracting set, we can show stronger results in many applications.

Definition 2.41 (Absorbing set)**.** *For $X \subset \mathcal{H}$, we call X is absorbing if for any bounded subset $B \subset \mathcal{H}$, there exists $T = T(B) \geq 0$ such that*

$$\Psi_t(B) \subset X$$

for all $t \geq T$.

Remark 2.42. *From Proposition 2.40, the existence of a compact absorbing set implies the existence of a global attractor.*

The following inequality, known as the kinetic energy principle, is useful to show the existence of an absorbing set, i.e., there exists $\lambda, K > 0$ such that

$$\frac{d}{dt}|\Psi_t(u_0)|^2 \leq -\lambda |u_0|^2 + K \tag{2.16}$$

for $t \geq 0$ and $u_0 \in \mathcal{H}$. As a result of the Gronwall lemma, this implies

$$|\Psi_t(u_0)|^2 \leq e^{-\lambda t}|u_0|^2 + \frac{K}{\lambda}(1 - e^{-\lambda t}). \tag{2.17}$$

This is an essential property of dissipative dynamical systems. In general, (2.16) is considered to be the existence of a Lyapunov function $\mathcal{E}(\,\cdot\,) = |\,\cdot\,|^2$ satisfying

$$\frac{d}{dt}\mathcal{E}(u_t) \leq -\lambda \mathcal{E}(u_t) + K \tag{2.18}$$

for $\lambda, K > 0$.

We can also consider these concepts in discrete-time dynamical systems [88]. In Section 5, we discuss the existence of the global attractor and other properties for dissipative dynamical systems with important examples appearing in geophysics.

3. State space model and data assimilation problems

3.1 Finite dimensional problems

3.1.1 State space model

Let us consider the finite-dimensional state space $\mathcal{H} = \mathbb{R}^{N_u}$ and the observation space $\mathcal{Y} = \mathbb{R}^{N_y}$ for $N_u \geq N_y$. We suppose that the time evolution of the true state is modeled by a discrete-time stochastic process $U : \mathbb{N} \times \Omega \to \mathbb{R}^{N_u}$ satisfying

$$U_n = \Psi(U_{n-1}) + \xi_n \tag{3.1}$$

with an uncertain initial state $U_0 \in \mathbb{R}^{N_u}$, where $\Psi : \mathbb{R}^{N_u} \to \mathbb{R}^{N_u}$ is a continuous map. The sequence $(\xi_n)_{n\in\mathbb{N}} \subset \mathbb{R}^{N_u}$ is an i.i.d. stochastic error, which represents modelling and approximation errors. Its mean is zero, and the covariance matrix is represented by a matrix $Q \in \mathbb{R}^{N_u \times N_u}$ with $Q \succeq 0$. The information from the unknown true state is obtained by noisy observations $Y : \mathbb{N} \times \Omega \to \mathbb{R}^{N_y}$,

$$Y_n = h(U_n) + \eta_n, \quad n \in \mathbb{N}, \tag{3.2}$$

where $h : \mathbb{R}^{N_u} \to \mathbb{R}^{N_y}$ is a continuous observation function and $(\eta_n)_{n\in\mathbb{N}} \subset \mathbb{R}^{N_y}$ is an i.i.d. noise sequence with a probability density function (PDF) p_Y. The following assumption is imposed on the observation noise to formulate data assimilation algorithms in Section 4.

Assumption 3.1. *For any $n \in \mathbb{N}$, $\eta_n \sim \mathcal{N}(0, R)$ with $R \in \mathbb{R}^{N_y \times N_y}, R \succ 0$.*

Moreover, the full observation is considered when analyzing data assimilation algorithms in an ideal setting.

Assumption 3.2 (Full observation)**.** *The state is fully observed, i.e., $h = id_{\mathcal{H}}$ and $R = r^2 I_{\mathcal{H}}$ for $r > 0$. If the observation function is a linear operator $H \in \mathcal{L}(\mathcal{H}, \mathcal{Y})$, i.e., $h(u) = Hu$, we suppose that $H = I_{\mathcal{H}}$.*

We also define a finite-dimensional state space model with continuous time, which is convenient for mathematical analysis. In addition, we can consider the following stochastic differential equation (SDE) for a stochastic process $U : [0, \infty) \times \Omega \to \mathbb{R}^{N_u}$,

$$dU_t = \mathcal{F}(U_t)dt + Q^{\frac{1}{2}} dW_t, \tag{3.3}$$

where $\mathcal{F} : \mathbb{R}^{N_u} \to \mathbb{R}^{N_u}$ is continuous, $W = (W_t)_{t\geq 0}$ is the N_u-dimensional Wiener process, and $Q \in \mathbb{R}^{N_u \times N_u}$ with $Q \succeq 0$. For SDEs, see the basic textbook [73], in which

we find the existence and the uniqueness theorem of the solution. In a similar manner, we adopt a continuous-time stochastic observation,

$$dY_t = h(U_t)dt + R^{\frac{1}{2}}dB_t, \tag{3.4}$$

where $h : \mathbb{R}^{N_u} \to \mathbb{R}^{N_y}$ is continuous, and $(B_t)_{t\geq 0}$ is the N_y-dimensional Wiener process independent of W, and $R \in \mathbb{R}^{N_y \times N_y}$ with $R \succeq 0$. We consider the following assumption corresponding to Assumption 3.1.

Assumption 3.3. *The covariance of observation noises is positive definite $R \succ 0$.*

3.1.2 Bayesian data assimilation problems

Let $\mathcal{T} = \mathbb{N} \cup \{0\}$ or $[0, \infty) \subset \mathbb{R}$ be a time index set. Stochastic processes $U : \mathcal{T} \times \Omega \to \mathcal{H}$ and $Y : \mathcal{T} \times \Omega \to \mathcal{Y}$ denote a true state and an observation, respectively. We write the observation up to time t as $\boldsymbol{Y}_t = \{Y_s \mid 0 \leq s \leq t\}$. We first formulate a state (signal) estimation problem, minimizing the L^2-error from the true state using the observations.

Definition 3.4. *For $t \in \mathcal{T}$, a random variable $V_t : \Omega \to \mathcal{H}$ is called an estimator based on the observations $\boldsymbol{Y}_t$ if V_t is $\mathscr{F}_t^Y$-measurable. Furthermore, it is said to be optimal if*

$$\mathbb{E}[|U_t - V_t|^2] = \inf\{\mathbb{E}[|U_t - V|^2] \mid V \in \mathcal{K}_t\},$$

where $\mathcal{K}_t = \{V : \Omega \to \mathcal{H} \mid V \in L^2(\Omega, \mathbb{P})$ is an estimator based on $\boldsymbol{Y}_t\}$. Here, $L^2(\Omega, \mathbb{P})$ is the space of square integrable functions with respect to $\mathbb{P}$ on Ω. The state estimation problem is to construct or approximate the optimal estimator V_t based on the observations $\boldsymbol{Y}_t$.

The following proposition implies that the optimal estimator is obtained by the conditional expectation.

Proposition 3.5 (Optimal estimation [73])**.** *An optimal estimator V_t of the state estimation problem is given by $V_t = \mathbb{E}[U_t \,|\, \mathscr{F}_t^Y]$.*

We then consider a Bayesian formulation of the state estimation problem, in which the estimation is represented by the conditional distribution. Here, we consider the discrete-time system. Let $\boldsymbol{y}_N = \{y_n \mid 0 \leq n \leq N\}$ denote the realizations of observations in a discrete-time interval $0 \leq n \leq N$ for $N \in \mathbb{N}$.

Definition 3.6 (Data assimilation problem)**.** *Let U be the unknown true state and $\boldsymbol{y}_N$ be the given observations up to $N \in \mathbb{N}$. For $n \in \mathbb{N}$, we consider a problem constructing a random variable V_n such that its probability distribution corresponds to the conditional probability distribution of U_n with respect to $\boldsymbol{y}_N$, $\mathbb{P}^{V_n} = \mathbb{P}^{U_n}(\,\cdot \mid \boldsymbol{y}_N)$. It is called a data assimilation problem. Data assimilation problems are classified into the following three types depending on n and N.*

- *Prediction if $n > N$;*
- *Filtering if $n = N$;*
- *Smoothing if $n < N$.*

Since the prediction distribution $\mathbb{P}^{U_n}(\,\cdot\mid \boldsymbol{y}_N)$ $(n > N)$ is obtained as the push-forward of the filtering distribution by the model dynamics, it is sufficient to deal with the filtering and smoothing problems. In Definition 3.6, the distributions are obtained as the posterior distributions using Bayes' formula in Proposition 2.30.

In many real-world applications, observations are often obtained at every discrete time step. In the discrete-time filtering problem, a successive update of the distribution $\mathbb{P}^{V_n}$ is useful. We assume that the model noise ξ_n has a probability density function p_ξ. If the model is deterministic, i.e., the covariance of the noise sequence $(\xi_n)_{n\in\mathbb{N}}$ in (3.1) is O, Dirac's delta function is used instead of p_ξ.

Definition 3.7 (Sequential data assimilation for a finite-dimensional state space)**.** *Suppose U and Y are governed by (3.1) and (3.2) respectively. Let a PDF p_{U_0} represent the initial uncertainty of U_0. The following successive update yields the exact filtering distribution $p_{V_n} = p_{U_n}(\,\cdot\mid \boldsymbol{y}_n)$ for $n \in \mathbb{N}$, starting with $p_{V_0} = p_{U_0}$.*

(I) (Prediction: $p_{V_{n-1}} \to p_{\widehat{V}_n}$) Propagate $p_{V_{n-1}}$ to $p_{\widehat{V}_n}$ using the model dynamics.

$$p_{\widehat{V}_n}(v) = \int_{\mathbb{R}^{N_u}} p_\xi(v - \Psi(v'))p_{V_{n-1}}(v')\,dv'. \tag{3.5}$$

(II) (Analysis: $p_{\widehat{V}_n}, y_n \to p_{V_n}$) Update $p_{\widehat{V}_n}$ to p_{V_n} using Bayes' formula:

$$p_{V_n}(v) = \frac{p_Y(y_n \mid v)p_{\widehat{V}_n}(v)}{\int_{\mathbb{R}^{N_u}} p_Y(y_n \mid v')p_{\widehat{V}_n}(v')\,dv'}, \tag{3.6}$$

where $p_Y(y \mid u)$ is the conditional PDF of Y with respect to $u \in \mathbb{R}^{N_u}$.

The step (I) is known as the prediction (or forecast) step and $\mathbb{P}^{\widehat{V}_n}(dv) = p_{\widehat{V}_n}(v)\,dv$ is called the prediction (or forecast) distribution. Similarly, the step (II) is known as the analysis (or update) step, and $\mathbb{P}^{V_n}(dv) = p_{V_n}(v)\,dv$ is referred to as the analysis (or filtering) distribution. In [63], the two steps (I) and (II) of Definition 3.7 are represented by operator forms on $\mathcal{M}_1(\mathbb{R}^{N_u})$.

$$\mathbb{P}^{\widehat{V}_n} = \mathcal{P}\mathbb{P}^{V_{n-1}}, \quad \mathbb{P}^{V_n} = L_{y_n}\mathbb{P}^{\widehat{V}_n},$$

where $\mathcal{P} : \mathcal{M}_1(\mathbb{R}^{N_u}) \to \mathcal{M}_1(\mathbb{R}^{N_u})$ and $L_{y_n} : \mathcal{M}_1(\mathbb{R}^{N_u}) \to \mathcal{M}_1(\mathbb{R}^{N_u})$ are a Markov transition operator associated with (3.1) and represent Bayes' update using the observation y_n, respectively.

We need to construct numerical algorithms to approximate the exact filtering distribution $\mathbb{P}^{V_n}(dv)$, which is discussed in Section 4. The several layers to estimate the true state in the filtering problem are shown in Figure 2. The layer (a) represents the hidden true states generated by the model dynamics. The noisy observations are obtained in the layer (b) from the true states. The conditional distribution of the true state $\mathbb{P}^{U_{n-1}}(\,\cdot\,|\,\boldsymbol{y}_{n-1})$ is propagated into $\mathbb{P}^{U_n}(\,\cdot\,|\,\boldsymbol{y}_{n-1})$ by the model dynamics and it becomes $\mathbb{P}^{U_n}(\,\cdot\,|\,\boldsymbol{y}_n)$ after conditioned by the observation data as in the layer (c). The layer (d) describes the exact filtering distributions $\mathbb{P}^{\widehat{V}_n}$ and $\mathbb{P}^{V_n}$ defined by the sequential data assimilation process in Definition 3.7. These replicate the conditional distributions in the layer (c). The layer (e) explains a filtering algorithm approximating the sequential data assimilation process. The approximated operations are denoted by $\widetilde{\mathcal{P}}$ and $\widetilde{L}_{y_n}$.

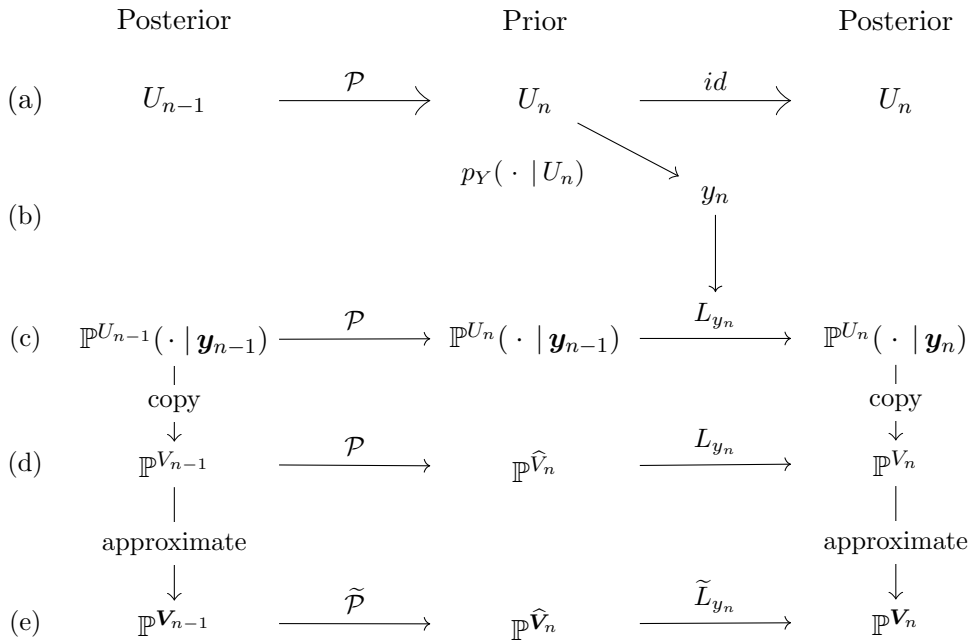

Figure 2: Several layers in the filtering problems. The layers are (a) model dynamics, (b) observation, (c) the conditional distribution, (d) the exact filtering distribution, and (e) an approximated filtering distribution.

The filtering distribution relates to the smoothing distribution as follows. We consider the PDF $p_{\boldsymbol{V}}(\boldsymbol{v}\,|\,\boldsymbol{y}_N)$ of the smoothing distribution for the states $\boldsymbol{v} = (v_0, \dots, v_N) \in \mathbb{R}^{N_u \times (N+1)}$ with respect to observations $\boldsymbol{y}_N$ up to time $n = N$.

Proposition 3.8 ([63]). *For the smoothing distribution $p_{\boldsymbol{V}}(\boldsymbol{v}\,|\,\boldsymbol{y}_N)$ over the discrete time interval $0 \le n \le N$ and the filtering distribution $p_{V_N}(v_N\,|\,y_N)$ at time $n = N$, the marginal of the smoothing distribution with respect to v_N is the filtering distribution.*

$$\int p_{\boldsymbol{V}}(\boldsymbol{v}\,|\,\boldsymbol{y}_N)\,dv_0 \dots dv_{N-1} = p_{V_N}(v_N\,|\,\boldsymbol{y}_N).$$

For deterministic model dynamics, estimating the initial state using all observations automatically yields an estimation of the final state.

Proposition 3.9 ([63])**.** *For the deterministic model dynamics* (3.1) *with* $Q = O$*, the push-forward of the smoothing distribution of* v_0 *is the filtering distribution of* v_N*.*

$$(\Psi^{(N)})_* \, \mathbb{P}^{V_0}(\,\cdot \mid \boldsymbol{y}_N) = \mathbb{P}^{V_N}(\,\cdot \mid \boldsymbol{y}_N),$$

where $\Psi^{(N)}$ *denotes the* N*-fold composition of* Ψ*.*

Next, we review the robustness of Bayesian formulations of data assimilation, which is the dependence of the posterior distributions on the observation data. As a consequence of Proposition 2.35, the well-posedness of the smoothing distribution $\mathbb{P}^{V_0}(\cdot \mid \boldsymbol{y}_N)$ is established for both the deterministic ($Q = O$) and the stochastic ($Q \neq O$) model dynamics (3.1) [63]. From these results and Corollary 2.36, the mean of the estimate of the initial state

$$\varpi_0 = \mathbb{E}[V_0 \mid \boldsymbol{y}_N]$$

is continuous with respect to $\boldsymbol{y}_N$.

3.2 Infinite dimensional problems

3.2.1 Discrete-time state space model

Let $\mathcal{H}$ be an infinite-dimensional Hilbert space. We first consider a discrete-time stochastic process $U : \mathbb{N} \times \Omega \to \mathcal{H}$ satisfying

$$U_n = \Psi(U_{n-1}) + \xi_n \tag{3.7}$$

with an uncertain initial state $U_0 \in \mathcal{H}$, where $\Psi : \mathcal{H} \to \mathcal{H}$ is continuous and $(\xi_n)_{n \in \mathcal{H}} \in \mathcal{H}$ is an i.i.d. and mean zero noise sequence with a covariance $Q \in \mathcal{L}_{sa}(\mathcal{H})$ satisfying $Q \in \mathcal{K}_1(\mathcal{H})$ and $Q \succeq 0$. The observation is a stochastic process $Y : \mathbb{N} \times \Omega \to \mathcal{Y}$, generated by

$$Y_n = h(U_n) + \eta_n, \quad n \in \mathbb{N}, \tag{3.8}$$

where $h : \mathcal{H} \to \mathcal{Y}$ is continuous and $(\eta_n)_{n \in \mathbb{N}} \subset \mathcal{Y}$ is an i.i.d. noise sequence.

When the observation space is finite dimensional, i.e., $\mathcal{Y} = \mathbb{R}^{N_y}$, we can define the posterior distribution $\mathbb{P}^{V_n}$ in terms of its Radon-Nikodým derivative using the generalized Bayes' formula (Proposition 2.32 and Proposition 2.35) when we consider the Gaussian likelihood as in Assumption 3.1. Hence, we consider Bayesian data assimilation problems as in Definition 3.6.

Definition 3.10 (Sequential data assimilation for an infinite-dimensional state space)**.** *Suppose U and Y are governed by (3.7) and (3.8) respectively. Suppose that $\mathcal{Y} = \mathbb{R}^{N_y}$ and the observation noise satisfies Assumption 3.1. The following successive update yields the exact filtering distribution $\mathbb{P}^{V_n} = \mathbb{P}^{U_n}(\cdot \mid \boldsymbol{y}_n)$ for $n \in \mathbb{N}$, starting with $\mathbb{P}^{V_0} = \mathbb{P}^{U_0}$.*

(I) (Prediction: $\mathbb{P}^{V_{n-1}} \to \mathbb{P}^{\widehat{V}_n}$) Propagate $\mathbb{P}^{V_{n-1}}$ to $\mathbb{P}^{\widehat{V}_n}$ using the model dynamics,

$$\mathbb{P}^{\widehat{V}_n}(dv) = \int_{\mathcal{H}} K(v_{n-1}, dv)\mathbb{P}^{V_{n-1}}(dv_{n-1}), \tag{3.9}$$

where $K : \mathcal{H} \times \mathcal{B}(\mathcal{H}) \to [0,1]$ is a transition kernel associated with (3.7).

(II) (Analysis: $\mathbb{P}^{\widehat{V}_n}, y_n \to \mathbb{P}^{V_n}$) For $\Phi(u; y) = \frac{1}{2}|y - h(u)|_R^2$, define $\mathbb{P}^{V_n} \in \mathcal{M}_1(\mathcal{H})$ using the generalized Bayes' formula,

$$\frac{d\mathbb{P}^{V_n}}{d\mathbb{P}^{\widehat{V}_n}}(v) \propto \exp(-\Phi(v; y)), \tag{3.10}$$

We have the same relationships as Propositions 3.8 and 3.9 for the infinite-dimensional state spaces [16]. Furthermore, the well-posedness of the smoothing distribution is established in [26].

It is not straightforward to consider a state space model with observations in infinite dimensions. For the case of an infinite-dimensional observation space $\mathcal{Y}$ [58], the positive definiteness of the noise covariance $R \succ 0$ implies $\operatorname{Tr} R = \infty$ from Proposition 2.8. Hence, R cannot be the covariance of any Gaussian distribution on $\mathcal{Y}$ from Proposition 2.29. As a result, Assumption 3.1 is not valid in this context. On the other hand, if $\operatorname{Tr} R < \infty$, then R is not invertible. Thus, the notation $|\cdot|_R = |R^{-\frac{1}{2}} \cdot|$ is only defined on the Cameron-Martin space $\operatorname{Ran}(R^{\frac{1}{2}})$. In this case, the normalizing constant of the posterior distribution $\mathbb{P}^{V_n}$ becomes zero. Therefore, the sequential Bayesian data assimilation can not be considered. See [51] for detailed formulations of data assimilation problems when both $\mathcal{H}$ and $\mathcal{Y}$ are infinite-dimensional. Instead of using the Bayesian formulation, we can consider state (signal) estimation problems as defined in Definition 3.4. Hence, we introduce an alternative assumption to Assumption 3.1 used to define data assimilation algorithms in Section 4.

Assumption 3.11. *For any $n \in \mathbb{N}$, $\eta_n \sim \mathcal{N}(0, \widetilde{R})$ with $\widetilde{R} \in \mathcal{K}_1(\mathcal{Y})$, $\widetilde{R} \succeq 0$, and $\widetilde{R} \preceq R$ for $R \succ 0$.*

3.2.2 Continuous-time state space model

For the continuous-time formulations, we do not consider the stochastic case to avoid dealing with continuous-time stochastic processes in infinite-dimensional spaces such as stochastic partial differential equations. Instead, we consider the infinite-dimensional

dynamical system. To handle this, we first introduce an evolution equation in a Hilbert space $\mathcal{H}$,

$$\frac{du}{dt} = \mathcal{F}(u). \tag{3.11}$$

We then consider noiseless observations in a lower dimensional space $\mathcal{Y}$ with $\dim(\mathcal{Y}) \leq \dim(\mathcal{H})$.

$$y(t) = h(u(t)), \tag{3.12}$$

where $h : \mathcal{H} \to \mathcal{Y}$ is an observation operator. Data assimilation problems in noiseless situations arise from the feedback control of partial differential equations [11]. From the perspective of control theory, it is important to determine whether and how many finite-dimensional control inputs into the simulated state are needed to reconstruct the true state. Such problems have been studied for dissipative dynamical systems, in particular, for the incompressible two-dimensional Navier-Stokes equations [10] and the incompressible three-dimensional Navier-Stokes-alpha equations [2]. This problem is further discussed in Section 5.

We finally remark the relationships between the discrete and continuous-time, finite and infinite-dimensional, deterministic and stochastic settings.

Remark 3.12. *If $Q \neq O$, (3.1) is often used as a discretization of the continuous-time dynamics. The model error ξ_n is interpreted as the cumulative discretization errors over time interval $[t_{n-1}, t_n]$ and in spatial domain [20]. For theoretical simplicity, ξ_n is often assumed to be the Gaussian noise.*

Remark 3.13. *In many applications, the unknown true state is modeled as a continuous-time process. However, the noisy observations are often obtained at discrete time steps with a time interval $\tau > 0$. With Figure 3, we explain the relationships between (3.3) and (3.1), and between (3.11) and (3.7). For the deterministic case (3.11) (resp. (3.3)) with $Q = O$, we suppose that a unique solution exists for any $u_0 \in \mathcal{H}$ and that it generates a one-parameter semigroup $\Psi_t : \mathcal{H} \to \mathcal{H}$ for $t \geq 0$. Then, let $\Psi = \Psi_\tau$ and $U_n = u_{n\tau}$, we obtain (3.7) (resp. (3.1)) with $Q = O$, see [54, 82] for more details. In the case of stochastic dynamics (3.3), let $\widetilde{U}_t$ be the unique solution starting at $U_0 = u$, it suffices to put $\Psi_t(u) = \mathbb{E}^u[\widetilde{U}_t]$, $\Psi = \Psi_\tau$, and $\xi_n = \widetilde{U}_{n\tau} - \Psi\left(\widetilde{U}_{n(\tau-1)}\right)$. See also [88].*

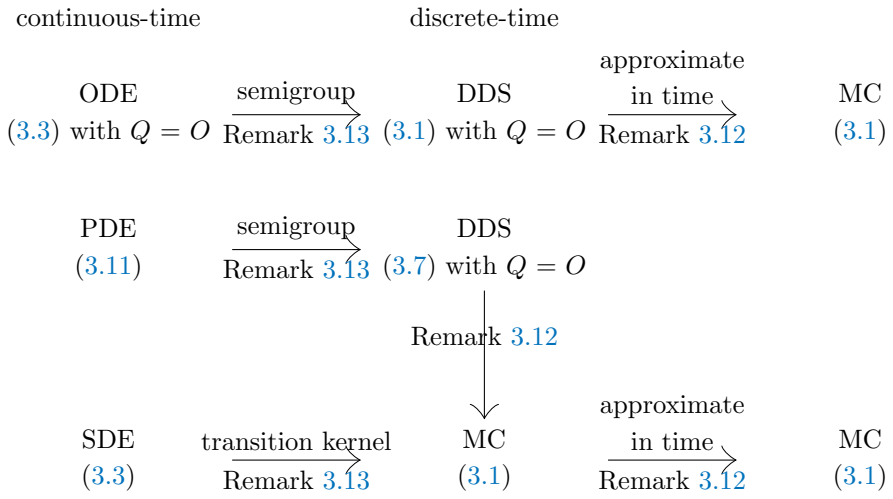

Figure 3: The relationships between various formulations of model dynamics. The abbreviations are as follows: ODE (Ordinary Differential Equation), DDS (Discrete Dynamical System), MC (Markov Chain), PDE (Partial Differential Equation), and SDE (Stochastic Differential Equation).

4. Filtering algorithms

4.1 The Kalman filter

For $N_u, N_y \in \mathbb{N}$, let $F \in \mathbb{R}^{N_u \times N_u}$ and $H \in \mathbb{R}^{N_y \times N_u}$. We then consider a discrete-time, finite-dimensional linear Gaussian system,

$$U_n = FU_{n-1} + \xi_n, \quad Y_n = HU_n + \eta_n, \tag{4.1}$$

where $\xi_n \sim \mathcal{N}(0, Q)$ is Gaussian noise with covariance matrix $Q \succeq 0$ and the Gaussian observation noise η_n satisfies Assumption 3.1. The linear-Gaussian system (4.1) is a special case of (3.1). We also assume that the initial uncertainty follows a Gaussian distribution $U_0 \sim \mathcal{N}(\varpi_0, P_0)$, where $\varpi_0 \in \mathbb{R}^{N_u}$ and $P_0 \in \mathbb{R}^{N_u \times N_u}$ with $P_0 \succ 0$. The Kalman filter (KF), originally proposed by Kalman [49], provides an exact and explicit representation of the filtering distribution (Definition 3.7) for the system (4.1).

Definition 4.1 (KF)**.** *Suppose that the Gaussian distribution $V_{n-1} \sim \mathcal{N}(\varpi_{n-1}, P_{n-1})$. Then, the algorithm of the Kalman filter (KF) is as follows.*

(I) (Prediction: $\varpi_{n-1}, P_{n-1} \to \widehat{\varpi}_n, \widehat{P}_n$) Compute the time evolution of the mean and the covariance:

$$\widehat{\varpi}_n = F\varpi_{n-1}, \tag{4.2}$$
$$\widehat{P}_n = FP_{n-1}F^* + Q. \tag{4.3}$$

(II) (Analysis: $\widehat{\varpi}_n, \widehat{P}_n, y_n \to \varpi_n, P_n$) Using Lemma 2.31, we can compute the mean and the covariance of the posterior distribution:

$$\varpi_n = \widehat{\varpi}_n + K_n(y_n - H\widehat{\varpi}_n), \tag{4.4}$$
$$P_n = (I - K_nH)\widehat{P}_n, \tag{4.5}$$

where K_n is the Kalman gain

$$K_n = \widehat{P}_nH^*(H\widehat{P}_nH^* + R)^{-1}. \tag{4.6}$$

The filtering distribution is represented by $V_n \sim \mathcal{N}(\varpi_n, P_n)$.

We introduce a concept to explain properties of the KF.

Definition 4.2 (Linear estimation)**.** *For $t \in \mathcal{T}$, a random variable $V_n : \Omega \to \mathbb{R}^{N_u}$ is said to be a linear estimator if $\mathbb{E}[|U_n - V_n|^2] = \inf\{|\widetilde{V}_n - U_n|^2 \mid \widetilde{V}_n \in \mathrm{span}(\boldsymbol{Y}_n)\}$.*

Proposition 4.3 (KF [4, 24, 73])**.** *For a linear-Gaussian system (4.1), the followings hold.*

(1) The exact filtering distribution $\mathbb{P}^{U_n}(\,\cdot\,|\,\boldsymbol{Y}_n)$ becomes a Gaussian distribution.

(2) A linear estimator attains the optimal estimator.

(3) The successive updates of the mean and covariance of the Gaussian distribution are given by Definition 4.1.

The equation (4.4) is known as the Kalman update. The following important lemma is a consequence of the Woodbury identity (Lemma 2.18).

Lemma 4.4. *The following identity holds.*

$$I_{N_u} - K_n H = (I_{N_u} + \widehat{P}_n H^* R^{-1} H)^{-1}. \tag{4.7}$$

Proof. It follows from (2.8) in Lemma 2.18 with $A = I_{N_u}$, $B = \widehat{P}_n H^*$, $C = H$, and $D = R^{-1}$ that

$$(I_{N_u} + \widehat{P}_n H^* R^{-1} H)^{-1} = I_{N_u} - \widehat{P}_n H^* (R + H \widehat{P}_n H^*)^{-1} H = I_{N_u} - K_n H.$$

□

Remark 4.5. *For $Q, P_0 \succ 0$, $P_n, \widehat{P}_n \succ 0$ for all $n \in \mathbb{N}$ follow by induction. This is confirmed by the following calculations of the inverse of the covariance matrices. For the (4.3), by taking $A = Q$, $B = F$, $C = P_{n-1}$, and $D = F^*$ in (2.8), we have*

$$\widehat{P}_n^{-1} = (Q + F P_{n-1} F^*)^{-1} = Q^{-1} - Q^{-1} F (P_{n-1}^{-1} + F^* Q^{-1} F)^{-1} F^* Q^{-1}.$$

For the (4.5), Lemma 4.4 yields

$$P_n^{-1} = \widehat{P}_n^{-1} (I_{N_u} - K_n H)^{-1} = \widehat{P}_n^{-1} (I_{N_u} + \widehat{P}_n H^* R^{-1} H) = \widehat{P}_n^{-1} + H^* R^{-1} H.$$

These equalities allow us to compute $\widehat{P}_n^{-1}$ and P_n^{-1} iteratively without directly evaluating $\widehat{P}_n$ and P_n. The inverses of the covariance matrices are known as the precision matrices. See also [63].

Remark 4.6. *For a degenerate matrix Q, the KF works successfully even with a degenerated covariance P_n if $R \succ 0$. Furthermore, Definition 4.1 remains valid for an infinite-dimensional linear-Gaussian system, in which we assume*

- *Linear model: $F \in \mathcal{L}(\mathcal{H})$,*
- *Gaussian model noise: $\xi_n \sim \mathcal{N}(0, Q)$ with $Q \in \mathcal{L}_{sa}(\mathcal{H}), Q \succeq 0, \operatorname{Tr} Q < \infty$,*
- *Linear observation operator: $H \in \mathcal{L}(\mathcal{H}, \mathcal{Y})$ with $\mathcal{Y} \subset \mathcal{H}$,*
- *Gaussian observation noise: η_n satisfies Assumption 3.11,*
- *Gaussian initial uncertainty: $U_0 \sim \mathcal{N}(\varpi_0, P_0)$ with $\varpi_0 \in \mathcal{H}, P_0 \in \mathcal{K}_1(\mathcal{H}), P_0 \succeq 0$.*

Remark 4.7 (Limitations of the KF)**.** *While the KF is theoretically clear, it faces two significant limitations when applied to high-dimensional and complex systems such as atmospheric models.*

(1) The KF assumes that Ψ is linear, whereas atmospheric dynamical models are typically nonlinear.

(2) The dimension of the state space can be extremely large, reaching up to 10^9. Consequently, the covariance matrix becomes $10^9 \times 10^9$, which is too large to store in computer memory.

4.2 Extensions for nonlinear dynamics

In this section, we consider Hilbert spaces $\mathcal{H}$ and $\mathcal{Y}$, the discrete-time nonlinear dynamical system (3.7) and (3.8) with a linear observation $h(u) = Hu$ for $H \in \mathcal{L}(\mathcal{H}, \mathcal{Y})$. The observation noises satisfy Assumption 3.11. We assume that the initial uncertainty $U_0 \sim \mathcal{N}(\varpi_0, P_0)$ where $\varpi_0 \in \mathcal{H}$ and $P_0 \in \mathcal{K}_1(\mathcal{H})$ with $P_0 \succeq 0$. The extended Kalman filter (ExKF) is an extension of the KF using linearization of Ψ to address limitation (1) stated in Remark 4.7.

Definition 4.8 (ExKF)**.** *Assume that Ψ is Fréchet differentiable, and let $D\Psi_u \in \mathcal{L}(\mathcal{H})$ be its derivative at $u \in \mathcal{H}$. For a given prior distribution $\boldsymbol{V}_{n-1}$ with the mean ϖ_{n-1} and the covariance matrix P_{n-1}, the algorithm of the extended Kalman filter (ExKF) is as follows.*

(I) (Prediction: $\varpi_{n-1}, P_{n-1} \to \widehat{\varpi}_n, \widehat{P}_n$)

$$\widehat{\varpi}_n = \Psi(\varpi_{n-1}), \tag{4.8}$$

$$\widehat{P}_n = D\Psi_{\varpi_{n-1}} P_{n-1} D\Psi^*_{\varpi_{n-1}} + Q. \tag{4.9}$$

(II) (Analysis: $\widehat{\varpi}_n, \widehat{P}_n, y_n \to \varpi_n, P_n$)

$$\varpi_n = \widehat{\varpi}_n + K_n(y_n - H\widehat{\varpi}_n), \tag{4.10}$$

$$P_n = (I - K_n H)\widehat{P}_n \tag{4.11}$$

with the Kalman gain $K_n = \widehat{P}_n H^* (H\widehat{P}_n H^* + R)^{-1}$.

In Definition 4.8, the linearizations $D\Psi$ and $D\Psi^*$ are called the tangent linear and adjoint models, respectively. Usually, it is difficult to compute the derivative $D\Psi$ numerically for high-dimensional or complex model dynamics.

The three-dimensional variational method (3DVar) is a simpler algorithm, in which the nonlinear prediction and the Kalman update with a fixed model covariance are repeated.

Definition 4.9 (3DVar)**.** *For the constant model covariance $\widehat{P}_n = P_0$, the algorithm of the 3DVar is as follows.*

(I) (Prediction: $\varpi_{n-1} \to \widehat{\varpi}_n$)

$$\widehat{\varpi}_n = \Psi(\varpi_{n-1}). \tag{4.12}$$

(II) (Analysis: $\widehat{\varpi}_n, y_n \to \varpi_n$)

$$\varpi_n = \widehat{\varpi}_n + K(y_n - H\widehat{\varpi}_n) \tag{4.13}$$

with the Kalman gain $K = P_0 H^*(HP_0H^* + R)^{-1}$.

In the step (II) of Definition 4.9, the Kalman gain K does not depend on time $n \in \mathbb{N}$, and the 3DVar is also known as optimal interpolation. While the 3DVar does not require the computation of the derivative of Ψ, it still faces the issue of large-storage requirements for the full-covariance matrix as pointed out in Remark 4.7.

4.3 Ensemble Kalman filter

We consider the same state space model as in the previous section. The ensemble Kalman filter (EnKF) [34] approximates the filtering distribution by the empirical distribution of an ensemble $\boldsymbol{V}_n = [v_n^{(1)}, \dots, v_n^{(m)}] \in \mathcal{H}^m$ of size $m \in \mathbb{N}$,

$$\mathbb{P}^{V_n}(\,\cdot\,) \approx \frac{1}{m}\sum_{k=1}^{m} \delta_{v_n^{(k)}}(\,\cdot\,),$$

where $\delta_u(\,\cdot\,)$ denotes the Dirac measure at u. Similarly, the distribution $\mathbb{P}^{\widehat{V}_n}$ in the prediction step is approximated by

$$\mathbb{P}^{\widehat{V}_n}(\,\cdot\,) \approx \frac{1}{m}\sum_{k=1}^{m} \delta_{\widehat{v}_n^{(k)}}(\,\cdot\,),$$

where $\widehat{\boldsymbol{V}}_n = [\widehat{v}_n^{(1)}, \dots, \widehat{v}_n^{(m)}] \in \mathcal{H}^m$, and each ensemble member evolves according to the nonlinear dynamical model (3.1). The model covariance is approximated by the ensemble covariance $\widehat{P}_n = \mathrm{Cov}_m \widehat{\boldsymbol{V}}_n \coloneqq \frac{1}{m-1}\sum_{k=1}^{m}(\widehat{v}_n^{(k)} - \overline{\widehat{v}}_n) \otimes (\widehat{v}_n^{(k)} - \overline{\widehat{v}}_n)$ with $\overline{\widehat{v}}_n = \frac{1}{m}\sum_{k=1}^{m} \widehat{v}_n^{(k)}$. Variants of the EnKF use the same prediction step to obtain the prediction ensemble $\widehat{\boldsymbol{V}}_n$ from $\boldsymbol{V}_{n-1}$. However, each variant employs a different approach in the analysis step to generate the analysis ensemble $\boldsymbol{V}_n$ from the prediction ensemble $\widehat{\boldsymbol{V}}_n$ and observation data y_n. A simple and stochastic implementation of the EnKF is known as the perturbed observation (PO) method [19].

Definition 4.10 (PO)**.** *Let $\boldsymbol{V}_0 = [v_0^{(k)}]_{k=1}^m \in \mathcal{H}^m$. The algorithm of the perturbed observation (PO) method consists of the following two steps. For $n \in \mathbb{N}$,*

(I) (Prediction: $\boldsymbol{V}_{n-1} \to \widehat{\boldsymbol{V}}_n$) Compute

$$\widehat{v}_n^{(k)} = \Psi(v_{n-1}^{(k)}) + \xi_n^{(k)}, \quad \xi_n^{(k)} \sim \mathcal{N}(0, Q), \quad k = 1, \dots, m, \tag{4.14}$$

and set $\widehat{\boldsymbol{V}}_n = [\widehat{v}_n^{(k)}]_{k=1}^m \in \mathcal{H}^m$.

(II) (Analysis: $\widehat{\boldsymbol{V}}_n, y_n \to \boldsymbol{V}_n$) Set $\widehat{P}_n = \mathrm{Cov}_m(\widehat{\boldsymbol{V}}_n)$ and replicate observations by adding random perturbations,

$$y_n^{(k)} = y_n + \eta_n^{(k)}, \quad \eta_n^{(k)} \sim \mathcal{N}(0, R), \quad k = 1, \dots, m, \tag{4.15}$$

and update the ensemble,

$$v_n^{(k)} = \widehat{v}_n^{(k)} + K_n(y_n^{(k)} - H\widehat{v}_n^{(k)}), \quad k = 1, \dots, m, \tag{4.16}$$

with the Kalman gain

$$K_n = \widehat{P}_n H^* (H\widehat{P}_n H^* + R)^{-1}. \tag{4.17}$$

Finally, set $\boldsymbol{V}_n = [v_n^{(k)}]_{k=1}^m \in \mathcal{H}^m$. We denote the map from $\widehat{\boldsymbol{V}}_n$ to $\boldsymbol{V}_n$ as $\boldsymbol{V}_n = \boldsymbol{V}_{PO}(\widehat{\boldsymbol{V}}_n; y_n, \widehat{P}_n)$.

Proposition 4.11 (Well-definedness of the PO method [51])**.** *Suppose $R \succ 0$, the PO method is well-defined, i.e., $H\widehat{P}_n H^* + R$ is invertible.*

We can implement the PO method without directly evaluating $\widehat{P}_n$ to avoid successive memory allocations.

Lemma 4.12. *The analysis step (II) of the PO method can be replaced by the following step without evaluating $\widehat{P}_n$.*

(II') (Analysis: $\widehat{\boldsymbol{V}}_n, y_n \to \boldsymbol{V}_n$) Decompose $\widehat{\boldsymbol{V}}_n = \overline{\widehat{v}}_n \mathbf{1} + d\widehat{\boldsymbol{V}}_n$. Set $d\widehat{\boldsymbol{Y}}_n = H d\widehat{\boldsymbol{V}}_n$. Define $y_n^{(k)}$ and $v_n^{(k)}$ for $k = 1, \dots, m$ as in the step (II) with

$$K_n = \frac{1}{m-1} d\widehat{\boldsymbol{V}}_n d\widehat{\boldsymbol{Y}}_n^* \left(\frac{1}{m-1} d\widehat{\boldsymbol{Y}}_n d\widehat{\boldsymbol{Y}}_n^* + R \right)^{-1}. \tag{4.18}$$

We denote this map as $\boldsymbol{V}_n = \boldsymbol{V}_{PO'}(\widehat{\boldsymbol{V}}_n; y_n)$.

Proof. The equality (4.18) follows owing to

$$\frac{1}{m-1} d\widehat{\boldsymbol{V}}_n d\widehat{\boldsymbol{Y}}_n^* = \frac{1}{m-1} d\widehat{\boldsymbol{V}}_n d\widehat{\boldsymbol{V}}_n^* H^* = \widehat{P}_n H^*$$

and

$$\frac{1}{m-1} d\widehat{\boldsymbol{Y}}_n d\widehat{\boldsymbol{Y}}_n^* = \frac{1}{m-1} H d\widehat{\boldsymbol{V}}_n d\widehat{\boldsymbol{V}}_n^* H^* = H\widehat{P}_n H^*.$$

□

As noted in [91], adding artificial noises in the analysis step of the PO method introduces additional errors in approximating the analysis distribution. To avoid this issue, deterministic versions of the EnKF, called the ensemble square root filters (ESRF), have been proposed [5, 14, 91]. These algorithms include the computation of a matrix square root to generate the analysis ensemble deterministically. One implementation of the ESRF algorithm is called the ensemble transform Kalman filter (ETKF) [14].

Definition 4.13 (ETKF)**.** *Let $\boldsymbol{V}_0 = [v_0^{(k)}]_{k=1}^m \in \mathcal{H}^m$. The algorithm of the ensemble transform Kalman filter (ETKF) consists of the following two steps.*

(I) (Prediction: $\boldsymbol{V}_{n-1} \to \widehat{\boldsymbol{V}}_n$) This step is the same as the step (I) of the PO method in Definition 4.10.

(II) (Analysis: $\widehat{\boldsymbol{V}}_n, y_n \to \boldsymbol{V}_n$) Decompose $\widehat{\boldsymbol{V}}_n = \overline{\widehat{v}}_n \mathbf{1} + d\widehat{\boldsymbol{V}}_n$ and set $\widehat{P}_n = \mathrm{Cov}_m(d\widehat{\boldsymbol{V}}_n)$. Update the mean

$$\overline{v}_n = \overline{\widehat{v}}_n + K_n(y_n - H\overline{\widehat{v}}_n) \tag{4.19}$$

with the Kalman gain $K_n = \widehat{P}_n H^*(H\widehat{P}_n H^* + R)^{-1}$. Take a symmetric matrix $T_n \in \mathbb{R}^{m\times m}$ satisfying

$$\frac{1}{m-1} d\widehat{\boldsymbol{V}}_n T_n (d\widehat{\boldsymbol{V}}_n T_n)^* = (I_{\mathcal{H}} - K_n H)\widehat{P}_n, \tag{4.20}$$

and transform the ensemble perturbation $d\boldsymbol{V}_n = d\widehat{\boldsymbol{V}}_n T_n$. The matrix T_n is called a transform matrix. Finally, set the analysis ensemble $\boldsymbol{V}_n = \overline{v}_n \mathbf{1} + d\boldsymbol{V}_n$. We denote the map as $\boldsymbol{V}_n = \boldsymbol{V}_{ETKF}(\widehat{\boldsymbol{V}}_n; y_n, \widehat{P}_n)$.

Theorem 4.1 (Well-definedness of the ETKF [82])**.** *Suppose $R \succ 0$. For any $\widehat{\boldsymbol{V}}_n \in \mathcal{H}^m$, there exists a unique symmetric transform matrix $T_n \in \mathbb{R}^{m\times m}$ satisfying (4.20). It is given by*

$$T_n = \left(I_m + \frac{1}{m-1} d\widehat{\boldsymbol{V}}_n^* H^* R^{-1} H d\widehat{\boldsymbol{V}}_n \right)^{-\frac{1}{2}}. \tag{4.21}$$

To prove Theorem 4.1, we prepare the key property of the Kalman gain K_n.

Lemma 4.14. *Let $\widehat{P}_n \succeq 0$, $H \in \mathcal{L}(\mathcal{H}, \mathcal{Y})$, and $R \in \mathcal{L}(\mathcal{Y})$ with $R \succ 0$. The Kalman gain satisfies*

$$K_n = (I_{\mathcal{H}} - K_n H)\widehat{P}_n H^* R^{-1}, \tag{4.22}$$

and (4.19) is equivalent to

$$(I_{\mathcal{H}} + \widehat{P}_n H^* R^{-1} H)\overline{v}_n = \overline{\widehat{v}}_n + \widehat{P}_n H^* R^{-1} y_n. \tag{4.23}$$

Proof. For simplicity, we omit the time index n in the following proofs since $n \in \mathbb{N}$ is fixed. Owing to $\widehat{P} \succeq 0$ and $R \succ 0$, we have $R + H\widehat{P}H^* \succ 0$ and it is thus invertible. Then, $I_{\mathcal{Y}} + R^{-1}H\widehat{P}H^* = R^{-1}(R + H\widehat{P}H^*)$ is also invertible since a product of two positive definite operators has positive spectrum as in Proposition 2.5. From (2.9) and the fact that $(AB)^{-1} = B^{-1}A^{-1}$ for invertible A, B, we have

$$\begin{aligned}(R + H\widehat{P}H^*)^{-1} &= (I_{\mathcal{Y}} + R^{-1}H\widehat{P}H^*)^{-1}R^{-1} \\ &= [I_{\mathcal{Y}} - (I_{\mathcal{Y}} + R^{-1}H\widehat{P}H^*)^{-1}R^{-1}H\widehat{P}H^*]R^{-1} \\ &= [I_{\mathcal{Y}} - (R + H\widehat{P}H^*)^{-1}H\widehat{P}H^*]R^{-1}.\end{aligned}$$

Hence, we have

$$\begin{aligned}K &= \widehat{P}H^*(R + H\widehat{P}H^*)^{-1} = \widehat{P}H^*[I_{\mathcal{Y}} - (R + H\widehat{P}H^*)^{-1}H\widehat{P}H^*]R^{-1} \\ &= (I_{\mathcal{H}} - KH)\widehat{P}H^*R^{-1},\end{aligned}$$

which is (4.22). On the other hand, it follows from (4.22) that

$$\overline{v} = \overline{\widehat{v}} + K(y - H\overline{\widehat{v}}) = (I_{\mathcal{H}} - KH)\overline{\widehat{v}} + Ky = (I_{\mathcal{H}} - KH)\overline{\widehat{v}} + (I_{\mathcal{H}} - KH)\widehat{P}H^*R^{-1}y.$$

To show (4.23), it is sufficient to verify $(I_{\mathcal{H}} + \widehat{P}H^*R^{-1}H)(I_{\mathcal{H}} - KH) = I_{\mathcal{H}}$ with (4.6). This follows from Lemma 4.4 □

Proof of Theorem 4.1. First, we prove the existence of T_n satisfying (4.20). Let $d\boldsymbol{Y} = Hd\widehat{\boldsymbol{V}}$. Then, the operator $R + \frac{1}{m-1}d\boldsymbol{Y}d\boldsymbol{Y}^* = R + H\widehat{P}H^*$ is invertible, and we consider the symmetric matrix

$$S = I_m - \frac{1}{m-1}d\boldsymbol{Y}^*\left(R + H\widehat{P}H^*\right)^{-1}d\boldsymbol{Y} \in \mathbb{R}^{m\times m}.$$

Then, we have

$$\begin{aligned}\frac{1}{m-1}d\widehat{\boldsymbol{V}}Sd\widehat{\boldsymbol{V}}^* &= \frac{1}{m-1}d\widehat{\boldsymbol{V}}d\widehat{\boldsymbol{V}}^* - \frac{1}{m-1}d\widehat{\boldsymbol{V}}d\boldsymbol{Y}^*\left(R + H\widehat{P}H^*\right)^{-1}\frac{1}{m-1}d\boldsymbol{Y}d\widehat{\boldsymbol{V}}^* \\ &= \widehat{P} - \widehat{P}H^*(R + H\widehat{P}H^*)^{-1}H\widehat{P} = (I - KH)\widehat{P},\end{aligned}$$

in which, we use $\frac{1}{m-1}d\boldsymbol{Y}d\widehat{\boldsymbol{V}}^* = H\widehat{P}$. From (2.12), we have

$$S = \left(I_m + \frac{1}{m-1}d\boldsymbol{Y}^*R^{-1}d\boldsymbol{Y}\right)^{-1} \tag{4.24}$$

and $S \succ 0$. We finally define the transform matrix $T = S^{\frac{1}{2}}$, which is nothing but (4.21). Then T becomes symmetric by definition. □

Theorem 4.1 is an extension of the well-definedness of the ETKF algorithm [60] to Hilbert spaces which can be infinite-dimensional. We also consider a practical implementation of the ETKF as Lemma 4.12.

Lemma 4.15. *The following analysis step is equivalent to the step (II) of the ETKF.*

(II') (Analysis: $\widehat{\boldsymbol{V}}_n, y_n \to \boldsymbol{V}_n$) Decompose $\widehat{\boldsymbol{V}}_n = \overline{\widehat{v}}_n \mathbf{1} + d\widehat{\boldsymbol{V}}_n$. Define the modified transform matrix

$$\widetilde{T}_n = \frac{1}{m-1} T_n^2 d\widehat{\boldsymbol{V}}_n^* H^* R^{-1}(y_n - H\overline{\widehat{v}}_n)\mathbf{1} + T_n, \tag{4.25}$$

and transform

$$\boldsymbol{V}_n = \overline{\widehat{v}}_n \mathbf{1} + d\widehat{\boldsymbol{V}}_n \widetilde{T}_n.$$

We denote the map as $\boldsymbol{V}_n = \boldsymbol{V}_{ETKF'}(\widehat{\boldsymbol{V}}_n; y_n)$.

Note that this alternative step avoids redundant memory allocations in practical numerical computation since we don't need to evaluate the covariance $\widehat{P}_n$ explicitly.

Proof. By multiplying the first term of (4.25) by $d\widehat{\boldsymbol{V}}$, it follows by definition of T_n that

$$\begin{aligned} d\widehat{\boldsymbol{V}}_n \frac{1}{m-1} T_n^2 d\widehat{\boldsymbol{V}}_n^* H^* R^{-1}(y_n - H\overline{\widehat{v}}_n) &= (I_{\mathcal{H}} - K_n H)\widehat{P}_n H^* R^{-1}(y_n - H\overline{\widehat{v}}_n) \\ &= K_n(y_n - H\overline{\widehat{v}}_n), \end{aligned}$$

where the last equality follows from the relation for the Kalman gain (4.22). □

Another implementation of the ESRF is called the ensemble adjustment Kalman filter (EAKF) [5].

Definition 4.16 (EAKF)**.** *Let $\boldsymbol{V}_0 = [v_0^{(k)}]_{k=1}^m \in \mathcal{H}^m$. The algorithm of the ensemble adjustment Kalman filter (EAKF) is as follows.*

(I) (Prediction: $\boldsymbol{V}_{n-1} \to \widehat{\boldsymbol{V}}_n$) This step is the same the step (I) of the PO method Definition 4.10.

(II) (Analysis: $\widehat{\boldsymbol{V}}_n, y_n \to \boldsymbol{V}_n$) Decompose $\widehat{\boldsymbol{V}}_n = \overline{\widehat{v}}_n \mathbf{1} + d\widehat{\boldsymbol{V}}_n$ and set $\widehat{P}_n = \mathrm{Cov}_m(\widehat{\boldsymbol{V}}_n)$. Update the mean

$$\overline{v}_n = \overline{\widehat{v}}_n + K_n(y_n - H\overline{\widehat{v}}_n) \tag{4.26}$$

with the Kalman gain $K_n = \widehat{P}_n H^*(H\widehat{P}_n H^* + R)^{-1}$. Take an appropriate $A_n \in \mathcal{L}(\mathcal{H})$, called an adjustment operator, satisfying

$$\frac{1}{m-1} A_n d\widehat{\boldsymbol{V}}_n (A_n d\widehat{\boldsymbol{V}}_n)^* = (I_{\mathcal{H}} - K_n H)\widehat{P}_n, \tag{4.27}$$

and transform the ensemble perturbation as $d\boldsymbol{V}_n = A_n d\widehat{\boldsymbol{V}}_n$. Finally, set the analysis ensemble $\boldsymbol{V}_n = \overline{v}_n \mathbf{1} + d\boldsymbol{V}_n$. We denote the map as $\boldsymbol{V}_n = \boldsymbol{V}_{EAKF}(\widehat{\boldsymbol{V}}_n; y_n, \widehat{P}_n)$.

The existence of the adjustment operator A_n is guaranteed by the following theorem.

Theorem 4.2 (Well-definedness of the EAKF)**.** *Suppose $R \succ 0$. For any $\widehat{\boldsymbol{V}}_n \in \mathcal{H}^m$, there exists an adjustment operator $A_n \in \mathcal{L}(\mathcal{H})$ satisfying (4.27). It is given by*

$$A_n = \Phi\Sigma E(I_\kappa + \Lambda)^{-\frac{1}{2}}\Sigma^{-1}\Phi^*. \tag{4.28}$$

Here, $\kappa = \operatorname{rank}\widehat{P}_n$, $\Sigma = \operatorname{diag}(s_1, \dots, s_\kappa)$ is the diagonal matrix consisting of the square root of non-zero eigenvalues of $\widehat{P}_n$ with the descending order, $\Phi = [\phi_1, \cdots, \phi_\kappa] \in \mathcal{H}^\kappa$ is the corresponding eigenvectors, i.e.,

$$\widehat{P}_n = \Phi\Sigma^2\Phi^*. \tag{4.29}$$

Moreover, the diagonal matrix $\Lambda \in \mathbb{R}^{\kappa\times\kappa}$ and a matrix $E \in \mathbb{R}^{\kappa\times\kappa}$ consist of the eigenvalues and eigenvectors of the symmetric matrix $\Sigma\Phi^*H^*R^{-1}H\Phi\Sigma \in \mathbb{R}^{\kappa\times\kappa}$ respectively, i.e.,

$$\Sigma\Phi^*H^*R^{-1}H\Phi\Sigma = E\Lambda E^*. \tag{4.30}$$

Proof. The proof is just an extension of that for a rank deficient case in [69, 88] to when the state space is an infinite-dimensional Hilbert space.

$$\begin{aligned}
\frac{1}{m-1}Ad\widehat{\boldsymbol{V}}(Ad\widehat{\boldsymbol{V}})^* &= \frac{1}{m-1}\Phi\Sigma E(I_\kappa+\Lambda)^{-\frac{1}{2}}\Sigma^{-1}\Phi^*d\widehat{\boldsymbol{V}}d\widehat{\boldsymbol{V}}^*\Phi\Sigma^{-1}(I_\kappa+\Lambda)^{-\frac{1}{2}}E^*\Sigma\Phi^* \\
&= \Phi\Sigma E(I_\kappa+\Lambda)^{-\frac{1}{2}}\Sigma^{-1}\Phi^*\widehat{P}_n\Phi\Sigma^{-1}(I_\kappa+\Lambda)^{-\frac{1}{2}}E^*\Sigma\Phi^* \\
&= \Phi\Sigma E(I_\kappa+\Lambda)^{-\frac{1}{2}}I_\kappa(I_\kappa+\Lambda)^{-\frac{1}{2}}E^*\Sigma\Phi^* \\
&= \Phi\Sigma E(I_\kappa+\Lambda)^{-1}E^*\Sigma\Phi^* \\
&= \Phi\Sigma(I_\kappa+E\Lambda E^*)^{-1}\Sigma\Phi^* \\
&= \Phi(\Sigma^{-2}+\Sigma^{-1}E\Lambda E^*\Sigma^{-1})^{-1}\Phi^* \\
&= \Phi(\Sigma^{-2}+\Phi^*H^*R^{-1}H\Phi)^{-1}\Phi^*.
\end{aligned} \tag{4.31}$$

Here, the third equality follows from (4.29), the fifth and sixth equalities follow from Lemma 2.21, and the last equality follow from (4.30). Applying the Woodbury identity (Lemma 2.18) for $A = \Sigma^{-2}$, $B = \Phi^*H^*$, $C = R^{-1}$, and $D = B^*$, we obtain

$$(\Sigma^{-2}+\Phi^*H^*R^{-1}H\Phi)^{-1} = \Sigma^2 - \Sigma^2\Phi^*H^*(R+H\Phi\Sigma^2\Phi^*H^*)^{-1}H\Phi\Sigma^2.$$

Substituting this into (4.31),

$$\begin{aligned}
\frac{1}{m-1}Ad\widehat{\boldsymbol{V}}(Ad\widehat{\boldsymbol{V}})^* &= \Phi(\Sigma^{-2}+\Phi^*H^*R^{-1}H\Phi)^{-1}\Phi^* \\
&= \Phi\Sigma^2\Phi^* - \Phi\Sigma^2\Phi^*H^*(R+H\Phi\Sigma^2\Phi^*H^*)^{-1}H\Phi\Sigma^2\Phi^* \\
&= \widehat{P}_n - \widehat{P}_nH^*(R+H\widehat{P}_nH^*)^{-1}H\widehat{P}_n \\
&= \widehat{P}_n - K_nH\widehat{P}_n = (I_{\mathcal{H}} - K_nH)\widehat{P}_n.
\end{aligned}$$

□

The simplified version of the analysis step (II) of the EAKF is given as follows.

Lemma 4.17. *The following analysis step is equivalent to the step (II) of the EAKF in Definition 4.16.*

(II') (Analysis: $\widehat{\boldsymbol{V}}_n, y_n \to \boldsymbol{V}_n$) Decompose $\widehat{\boldsymbol{V}}_n = \overline{\widehat{v}}_n \mathbf{1} + d\widehat{\boldsymbol{V}}_n$. Update the mean

$$\overline{v}_n = \overline{\widehat{v}}_n + K_n(y_n - H\overline{\widehat{v}}_n)$$

with the Kalman gain given by (4.18). Apply the singular value decomposition

$$\frac{1}{\sqrt{m-1}} d\widehat{\boldsymbol{V}} = \Phi \Sigma \widetilde{E}^*, \tag{4.32}$$

where $\widetilde{E} \in \mathbb{R}^{\kappa \times \kappa}$ is a unitary matrix. Also, apply the eigenvalue decomposition

$$\frac{1}{m-1} d\widehat{\boldsymbol{V}}^* H^* R^{-1} H d\widehat{\boldsymbol{V}} = E' \Lambda' (E')^*, \tag{4.33}$$

where $E' \in \mathbb{R}^{\kappa \times \kappa}$ is a unitary matrix. Define the adjustment operator

$$\widetilde{A}_n = \Phi \Sigma \widetilde{E}^* E' (I_\kappa + \Lambda')^{-\frac{1}{2}} \Sigma^{-1} \Phi^*. \tag{4.34}$$

The rest of the algorithm is the same as (II) in Definition 4.16 with $A_n = \widetilde{A}_n$. We denote the map as $\boldsymbol{V}_n = \boldsymbol{V}_{EAKF'}(\widehat{\boldsymbol{V}}_n; y_n)$.

Proof. By substituting (4.32) into (4.33), we have

$$\widetilde{E} \Sigma \Phi^* H^* R^{-1} H \Phi \Sigma \widetilde{E}^* = E' \Lambda' (E')^*.$$

Hence, it follows from (4.30) that

$$\widetilde{E}^* E' \Lambda' (E')^* \widetilde{E} = \Sigma \Phi^* H^* R^{-1} H \Phi \Sigma = E \Lambda E^*.$$

Since $E, E', \widetilde{E}$ are unitary, we get

$$\Lambda' = \Lambda, \quad \widetilde{E}^* E' = E.$$

This implies $\widetilde{A}_n = A_n$ given by (4.28). □

The EnKF provides a low-rank approximation of the covariance without computing the derivative of Ψ. There still remains the other issue originated from using the finite size ensemble.

Remark 4.18. *Due to the prediction covariance $\widehat{P}_n$ is approximated by an ensemble of m vectors, the rank is bounded by*

$$\operatorname{rank} \widehat{P}_n \leq m - 1.$$

Moreover, the analysis ensemble lies in the subspace spanned by the prediction ensemble. This is so-called the subspace property of the EnKF [78], which is shared with algorithms related to the EnKF [47, 90]. The numerical approaches to this limitation are discussed in Section 4.4.

4.4 Numerical techniques for the EnKF

The choice of the prediction covariance $\widehat{P}_n$ in the analysis step is crucial for the filtering algorithms. In the EnKF, it is approximated by propagating the analysis covariance P_{n-1} from the previous time. In practical numerical applications, $\widehat{P}_n$ is often underestimated in uncertain directions of the high-dimensional state space due to the limited ensemble size $m \ll N_u$, which leads to poor state estimation. To resolve this issue, covariance inflation techniques extend data assimilation algorithms by introducing an additional parameter α. The idea of the covariance inflation is to inflate $\widehat{P}_n$ before the analysis step [8, 69, 91]. The method of introducing inflation depends on the filtering algorithm, as described below.

Definition 4.19 (Additive inflation for the PO method)**.** *Let $\alpha \geq 0$. In the analysis step (II) of Definition 4.10, one defines an inflated covariance $\widehat{P}_n^\alpha = \widehat{P}_n + \alpha^2 I_{\mathcal{H}}$ and computes $\boldsymbol{V}_n = \boldsymbol{V}_{PO}(\widehat{\boldsymbol{V}}_n; y_n, \widehat{P}_n^\alpha)$.*

This approach is called an additive inflation of the covariance, and α is referred to as the inflation parameter.

Definition 4.20 (Multiplicative inflation for the EnKF)**.** *Let $\alpha \geq 1$. In the analysis step of the EnKF algorithms, one introduces the multiplicative inflation as follows.*

(1) In (II) of Definition 4.10, one defines an inflated covariance $\widehat{P}_n^\alpha = \alpha^2 \widehat{P}_n$ and computes $\boldsymbol{V}_n = \boldsymbol{V}_{PO}(\widehat{\boldsymbol{V}}_n; y_n, \widehat{P}_n^\alpha)$.

(1') In (II) of Definition 4.10, one defines an inflated perturbation $d\widehat{\boldsymbol{V}}_n^\alpha = \alpha d\widehat{\boldsymbol{V}}_n$ and ensemble $\widehat{\boldsymbol{V}}^\alpha = \overline{\widehat{v}}_n \mathbf{1} + d\widehat{\boldsymbol{V}}^\alpha$, covariance $\widehat{P}_n^\alpha = \mathrm{Cov}_m(d\widehat{\boldsymbol{V}}_n^\alpha)$, and computes $\boldsymbol{V}_n = \boldsymbol{V}_{PO}(\widehat{\boldsymbol{V}}_n^\alpha; y_n, \widehat{P}_n^\alpha)$. Equivalently, one can compute $\boldsymbol{V}_n = \boldsymbol{V}_{PO'}(\widehat{\boldsymbol{V}}_n^\alpha; y_n)$ in Lemma 4.12.

(2) In (II) of Definition 4.13, one computes $\boldsymbol{V}_n = \boldsymbol{V}_{ETKF}(\widehat{\boldsymbol{V}}_n^\alpha; y_n, \widehat{P}_n^\alpha)$ with $d\widehat{\boldsymbol{V}}_n^\alpha, \widehat{P}_n^\alpha$ as in (1'). Equivalently, one can compute $\boldsymbol{V}_n = \boldsymbol{V}_{ETKF'}(\widehat{\boldsymbol{V}}_n^\alpha; y_n)$ in Lemma 4.15.

(3) In (II) of Definition 4.16, one computes $\boldsymbol{V}_n = \boldsymbol{V}_{EAKF}(\widehat{\boldsymbol{V}}_n^\alpha; y_n, \widehat{P}_n^\alpha)$ with $d\widehat{\boldsymbol{V}}_n^\alpha, \widehat{P}_n^\alpha$ as in (1'). Equivalently, one can compute $\boldsymbol{V}_n = \boldsymbol{V}_{EAKF'}(\widehat{\boldsymbol{V}}_n^\alpha; y_n)$ in Lemma 4.17.

Remark 4.21. *The additive inflation directly improves the rank of $\widehat{P}$, ensuring that $\widehat{P}^\alpha$ is full-rank. However, the multiplicative inflation does not. Instead, the multiplicative inflation for the ensemble (1'), (2), (3) in Definition 4.20 may maintain the rank of the ensemble covariance in a successive data assimilation process since the ensemble $\widehat{\boldsymbol{V}}_n^\alpha$ is inflated before the contraction in the analysis step. This observation plays an important role in establishing the error bound of the ESRF and is discussed in Section 7.*

Remark 4.22. *For the ETKF and the EAKF, the relationships between the prediction and analysis ensembles are summarized as follows. The mean update (4.23) is given by*

$$(I_{\mathcal{H}} + \alpha^2 \widehat{P}_n H^* R^{-1} H)\overline{v}_n = \overline{\widehat{v}}_n + \alpha^2 \widehat{P}_n H^* R^{-1} y_n, \tag{4.35}$$

and the analysis covariance satisfies

$$P_n = \frac{\alpha^2}{m-1} d\widehat{\boldsymbol{V}}_n (I_m + \alpha^2 d\widehat{\boldsymbol{V}}_n^* H^* R^{-1} H d\widehat{\boldsymbol{V}}_n)^{-1} d\widehat{\boldsymbol{V}}_n^*. \tag{4.36}$$

For the PO method, these equalities hold only in the meaning of the conditional expectation provided $\widehat{\boldsymbol{V}}_n$ and y_n.

The covariance inflation techniques practically improve the state estimation error in the ESRF with multiplicative inflation [69, 91], in the PO method with additive inflation [54]. For large-scale atmospheric models, large α is often required, and manual tuning of α is expensive [45]. To avoid this, adaptive tuning algorithms have been developed [6, 7, 72]. Another approach focuses on the residual, which is the difference between the measured observation and the simulated (or predicted) observation, and they derive upper and lower bounds of α for the multiplicative inflation to ensure that the residual remains within a prescribed interval [68].

Related numerical techniques, known as the relaxation-to-prior methods, relax the contraction of the ensemble at the analysis step. For instance, the relaxation to prior perturbation (RTPP) method [94] interpolates the prediction and analysis ensemble perturbations with the ratio of $\alpha \in [0, 1]$ as

$$d\boldsymbol{V}^{\alpha}_{RTPP} = \alpha\, d\widehat{\boldsymbol{V}} + (1-\alpha) d\boldsymbol{V}.$$

Another example is the relaxation to prior spread (RTPS) method [92], which relaxes the contraction of the analysis ensemble spread by multiplying a factor determined for each component in the state space for $\alpha \in [0, 1]$.

$$d\boldsymbol{V}^{\alpha}_{RTPS}[i] = \alpha_i d\boldsymbol{V}[i], \quad \alpha_i = \frac{\alpha |d\widehat{\boldsymbol{V}}[i]|_2 + (1-\alpha)|d\boldsymbol{V}[i]|_2}{|d\boldsymbol{V}[i]|_2},$$

where $\boldsymbol{V}[i] = [(v^{(k)})^i]_{k=1}^m$ denotes an ensemble of the i-th component of each vector for $\boldsymbol{V} \in \mathbb{R}^{N_u \times m}$. Adaptive tuning methods for the RTPP and RTPS have also been proposed [57, 93].

The other technique to avoid covariance underestimation is called the localization [40, 46, 91]. The central idea is to ignore observations in regions far from each state variable during the analysis step. A well-known algorithm using this approach is the local ensemble transform Kalman filter (LETKF) [46]. Although localization techniques are essential technique in the EnKF, mathematical analysis with the localization is not covered in this thesis, see [31, 87, 90].

4.5 Continuous-time algorithms

The continuous version of the KF for (3.3) and (3.4) is called the Kalman Bucy filter (KBF) [73, 80]. Let us consider the continuous-time and finite-dimensional linear system for (3.3) and (3.4) with

$$\mathcal{F}(u) = Fu, \quad h(u) = Hu, \tag{4.37}$$

where $F \in \mathbb{R}^{N_u \times N_u}, H \in \mathbb{R}^{N_y \times N_u}$ for $N_u, N_y \in \mathbb{N}$ and a Gaussian initial uncertainty $U_0 \sim \mathcal{N}(\varpi_0, P_0)$ with $\varpi_0 \in \mathcal{H}, P_0 \in \mathbb{R}^{N_u \times N_u}, P_0 \succeq 0$.

Proposition 4.23 (The Kalman Bucy filter [73, 80])**.** *Suppose the observation noise in* (3.4) *with* (4.37) *satisfies Assumption 3.3. Then,* $V_t = \mathbb{E}[U_t \,|\, \mathscr{F}_t^Y]$ *satisfies the equations in Definition 4.24.*

Definition 4.24 (The Kalman Bucy filter)**.** *The Kalman Bucy filter (KBF) is defined by the following equations.*

$$\begin{aligned} dV_t &= (F - P_t H^* R^{-1} H) V_t \, dt + P_t H^* R^{-1} dY_t \\ &= F V_t \, dt + K_t (dY_t - H V_t \, dt), \\ V_0 &= \varpi_0, \end{aligned}$$

where $K_t = P_t H^* (H P_t H^* + R)^{-1}$ *with the covariance* P_t *of* V_t *satisfying*

$$\frac{dP_t}{dt} = F P_t + P_t F^* + Q - P_t H^* R^{-1} H P_t = F P_t + P_t F^* + Q - K_t R K_t^*.$$

The continuous-time extensions of the EnKF have also been proposed and summarized in [13]. Here, we introduce a deterministic version called the ensemble Kalman Bucy filter (EnKBF), which is available for the nonlinear system (3.3) and (3.4).

Definition 4.25 (The ensemble Kalman Bucy filter [12])**.** *Let the model dynamics follow* (3.3) *and* $\boldsymbol{V}_0 \in \mathcal{H}^m$*. Suppose that the observation noise in* (3.4) *satisfies Assumption 3.3. Then, the EnKBF is given by*

$$dV_t^{(k)} = \mathcal{F}(V_t^{(k)}) \, dt + Q P_t^\dagger (V_t^{(k)} - \overline{v}_t) \, dt - \frac{1}{2} \widetilde{P}_t R^{-1} (h(V_t^{(k)}) \, dt + \overline{h}_t \, dt - 2 \, dY_t)$$

for $k = 1, \dots, m$*, where* $P_t^\dagger$ *is the pseudo inverse of* P_t *and*

$$\begin{aligned} \overline{v}_t &= \frac{1}{m} \sum_{k=1}^m V_t^{(k)}, \quad P_t = \frac{1}{m-1} \sum_{k=1}^m (V_t^{(k)} - \overline{v}_t) \otimes (V_t^{(k)} - \overline{v}_t), \\ \overline{h}_t &= \frac{1}{m} \sum_{k=1}^m h(V_t^{(k)}), \quad \widetilde{P}_t = \frac{1}{m-1} \sum_{k=1}^m (V_t^{(k)} - \overline{v}_t) \otimes (h(V_t^{(k)}) - \overline{h}_t). \end{aligned}$$

5. Dissipative dynamical systems

Let us consider the evolution equation (3.11). We focus on dissipative dynamical systems on Hilbert spaces $\mathcal{H}$, in particular, chaotic systems that appear in atmospheric and oceanic modeling. See the handbook [38] for further examples of dynamical systems in fluid mechanics.

5.1 Dissipativeness and examples

In the mathematical analysis of data assimilation algorithms, it is essential to ensure that the trajectory is bounded and to estimate the time evolution of the error between two trajectories with a small initial perturbation. We consider two assumptions on the deterministic model dynamics (3.11) to characterize dissipative dynamical systems.

Assumption 5.1. *The evolution equation (3.11) has a unique solution for any $u_0 \in \mathcal{H}$, which generates a one-parameter semigroup $\Psi_t : \mathcal{H} \to \mathcal{H}$. In addition, there exists $\rho > 0$ such that Ψ_t has an absorbing ball $B(\rho) = \{v \in \mathcal{H} \mid |v| \le \rho\}$, i.e., $\Psi_t(v) \in B(\rho)$ for any $v \in B(\rho)$ and $t \ge 0$.*

Assumption 5.2. *There exists $\beta \in \mathbb{R}$ such that,*

$$\langle \mathcal{F}(u) - \mathcal{F}(v), u - v \rangle \le \beta |u - v|^2, \tag{5.1}$$

for any $u \in B(\rho)$ and $v \in \mathcal{H}$.

Kelly et al. [54] impose specific assumptions on $\mathcal{F}$ in addition to (5.1), as it is designed for the two-dimensional Navier-Stokes equations with periodic boundary conditions. Assumption 5.2 implies the following lemma to estimate the error growth along the dynamical system.

Lemma 5.3 (The upper bound of the error growth [54, 82])**.** *Suppose that Assumption 5.1 and 5.2 hold. Then,*

$$|\Psi_t(u) - \Psi_t(v)| \le e^{\beta t} |u - v|, \tag{5.2}$$

for any $u \in B(\rho)$, $v \in \mathcal{H}$ and $t > 0$.

From Lemma 5.3, we can interpret β as the (upper bound of) maximum error growth rate in the absorbing ball. If $\beta < 0$, any perturbation contracts to zero exponentially fast, indicating that the dynamical system is not chaotic. It is noteworthy that Lemma 5.3 does not assume the scale of the initial error $|u - v|$.

In the following error analysis, u and v represent the true and analysis states, respectively. Considering the ensemble mean as the analysis state, we require another inequality to estimate the error growth. To this end, we consider the following stronger assumption.

Assumption 5.4. *For $\rho > 0$ in Assumption 5.1 and $\mathcal{F}$ in (3.11), there exists $\beta > 0$ such that*

$$|\mathcal{F}(u) - \mathcal{F}(v)| \leq \beta |u - v|, \quad u, v \in B(\rho).$$

Then, we have the following lemma.

Lemma 5.5 (Upper bound of the error growth for the ensemble mean [82])**.** *Suppose Assumption 5.1 and Assumption 5.4 hold, and that $u_0 \in B(\rho)$ and $v_0^{(n)} \in B(\rho)$ for $n = 1, \ldots, N$. We define $u_t = \Psi_t(u_0)$ and $\overline{v}_t = \frac{1}{N}\sum_{n=1}^{N} \Psi_t(v_0^{(n)})$. Then, for any $\epsilon > 0$, $t > 0$, we have*

$$|\overline{v}_t - u_t|^2 \leq e^{2(\beta+\epsilon)t}(|\overline{v}_0 - u_0|^2 + D) - D, \tag{5.3}$$

where $D = \frac{\beta^2\rho^2}{(\beta+\epsilon)\epsilon}$.

5.1.1 Examples of finite-dimensional dissipative dynamical systems

We introduce two examples of finite-dimensional dissipative dynamical systems.

Example 5.6. *The Lorenz 63 equation (L63) is a three-dimensional nonlinear ordinary differential equation given by*

$$\frac{dx}{dt} = -\sigma x + \sigma y, \tag{5.4a}$$
$$\frac{dy}{dt} = \varrho x - y - xz, \tag{5.4b}$$
$$\frac{dz}{dt} = -bz + xy, \tag{5.4c}$$

where $\sigma > 0$, $b > 1$, and $\varrho > 0$. It was originally proposed by Lorenz [65]. In [85], a shifted version of the L63 equation is considered to analyze its absorbing property. By the changing variables $(x, y, z) \mapsto (x, y, z - \varrho - \sigma)$, (5.4) becomes

$$\frac{dx}{dt} = -\sigma x + \sigma y, \tag{5.5a}$$
$$\frac{dy}{dt} = -\sigma x - y - xz, \tag{5.5b}$$
$$\frac{dz}{dt} = -bz + xy - b(\varrho + \sigma). \tag{5.5c}$$

The L63 equation satisfies the assumptions of dissipative dynamical systems.

Proposition 5.7 ([41, 85])**.** *Let $\mathcal{H} = \mathbb{R}^3$ and $u = (x, y, z)^* \in \mathbb{R}^3$. Then, for any $u_0 \in \mathcal{H}$, there exists a unique solution $u(t) \in \mathcal{H}$ to (5.5) with $u(0) = u_0$. The L63 equation satisfies Assumption 5.1 with $\rho = \frac{b(\varrho+\sigma)}{\sqrt{4(b-1)}}$, Assumption 5.2 with $\beta = 2\rho - 1$, and Assumption 5.4 with some $\beta > 0$. Furthermore, there exists a global attractor $\mathscr{A}$.*

Figure 4 shows the projected global attractor of the L63 equation with typical parameters $\sigma = 10, b = \frac{8}{3}, \varrho = 28$. We can enjoy the interactive animation of the trajectories of the L63 equation on on the web site `https://kotatakeda.github.io/lorenz-webgl` [81].

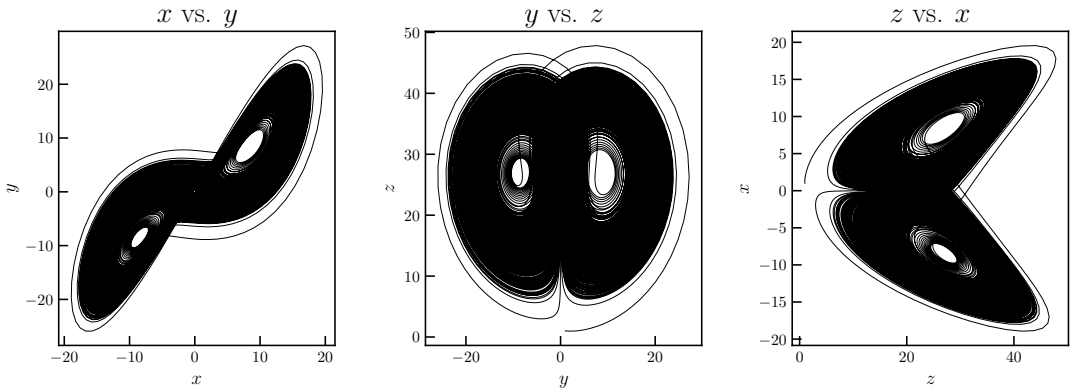

Figure 4: The global attractor of the L63 equation with $\sigma = 10, b = \frac{8}{3}, \varrho = 28$.

Another important example is the Lorenz 96 (L96) equation, which is a spatially extended chaotic phenomenological model in meteorology [66, 67]. The L96 equation also satisfies the assumptions.

Example 5.8. *For a given number of components $J \in \mathbb{N}$, the L96 equation for a state vector $u = (u^1, \dots, u^J)^* \in \mathbb{R}^J$ is given by*

$$\frac{du^i}{dt} = (u^{i+1} - u^{i-2})u^{i-1} - u^i + f, \quad i = 1, \dots, J, \tag{5.6}$$

where $u^{-1} = u^{J-1}$, $u^0 = u^J$, and $u^{J+1} = u^1$ and $f \in \mathbb{R}$ is external forcing.

Proposition 5.9 ([62])**.** *Let $\mathcal{H} = \mathbb{R}^J$. Then, for any $u_0 \in \mathcal{H}$, there exists a unique solution $u(t) \in \mathcal{H}$ to (5.6) with $u(0) = u_0$. The L96 equation satisfies Assumption 5.1 with $\rho = \sqrt{2J}|f|$, Assumption 5.2 with $\beta = 2\rho - 1$, and Assumption 5.4 with $\beta > 0$. Furthermore, there exists the global attractor $\mathscr{A}$.*

5.1.2 Examples of infinite-dimensional dissipative dynamical systems

Let Ω be an open subset of $\mathbb{R}^d$ for $d \in \mathbb{N}$. For $1 \le p \le \infty$, $L^p(\Omega)$ denotes the space of L^p functions on Ω. Similarly, for $n \in \mathbb{N}$ and $1 \le p \le \infty$, $W^{n,p}(\Omega)$ denote the space of L^p functions on Ω whose weak k-th derivatives belong to $L^p(\Omega)$ for $k = 1, \dots, n$. In particular, $H^n(\Omega) = W^{n,2}(\Omega)$. The standard L^2 and H^1 norms are denoted by $|\cdot|$ and $\|\cdot\|$, respectively. For a Banach space $\mathcal{X}$, $1 \le p \le \infty$ (resp. $n \in \mathbb{N}$), $-\infty \le a < b \le \infty$,

$L^p(a,b;\mathcal{X})$ (resp. $H^n(a,b;\mathcal{X})$) denotes the space of L^p (resp. H^n) functions from (a,b) to $\mathcal{X}$. Similarly, for $-\infty < a < b < \infty$, we denote the space of continuous functions from $[a,b]$ to $\mathcal{X}$ by $C([a,b];\mathcal{X})$. The two-dimensional Navier-Stokes equations are given as follows.

Example 5.10 (The two-dimensional Navier-Stokes equations)**.** *For $L > 0$, let $\Omega = [0,L]^2$ and let $\mathscr{V}$ be the set of vector-valued L-periodic trigonometric polynomials $\phi : \Omega \to \mathbb{R}^2$ with $\nabla \cdot \phi = 0$ and $\int_\Omega \phi = 0$. We define its closures with respect to the L^2-norm as $\mathcal{H} = \overline{\mathscr{V}}^{|\cdot|}$ and with respect to the H^1-norm as $\mathcal{V} = \overline{\mathscr{V}}^{\|\cdot\|}$. We consider the Leray-Helmholtz projector $P_\mathcal{H}$, i.e., the L^2-orthogonal projection $P_\mathcal{H} : L^2(\Omega) \to \mathcal{H}$. With the notations in [41, 54, 59, 85], the incompressible two-dimensional Navier-Stokes equations with periodic boundary conditions are given by*

$$\frac{du}{dt} + \mathcal{A}u + \mathcal{B}(u,u) = f, \tag{5.7}$$

where an unbounded linear operator $\mathcal{A} : \mathcal{H} \to \mathcal{V}^*$ is defined as $\mathcal{A} = -\nu\triangle$, a symmetric bilinear operator $\mathcal{B} : \mathcal{V} \times \mathcal{V} \to \mathcal{V}^*$ is defined as $\mathcal{B}(u,v) = \frac{1}{2}[P_\mathcal{H}(u\cdot\nabla)v + P_\mathcal{H}(v\cdot\nabla)u]$, $\nu > 0$ is the kinematic viscosity and $f \in \mathcal{H}$ is a time independent forcing. The domain of $\mathcal{A}$ in $\mathcal{H}$ is denoted by $D(\mathcal{A})$.

We refer to the incompressible two-dimensional Navier-Stokes equations with periodic boundary conditions as the 2D-NSE on a torus.

Proposition 5.11 (The 2D-NSE on a torus [41, 84, 85])**.** *For $u_0, f \in \mathcal{H}$, there exists a unique weak solution to (5.7) with $u(0) = u_0$ satisfying*

$$u \in C([0,T];\mathcal{H}) \cap L^2(0,T;\mathcal{V}), \quad \frac{du}{dt} \in L^2(0,T;\mathcal{V}^*)$$

for any $T > 0$. The semigroup $\Psi_t : \mathcal{H} \ni u_0 \mapsto u_t \in \mathcal{H}$ is well-defined for $t \geq 0$, and it is continuous from $\mathcal{H}$ into itself. A ball $B(\rho) = B_\mathcal{H}(\rho)$ in $\mathcal{H}$ with $\rho = \frac{|f|}{\nu\lambda_1}$ is absorbing where $\lambda_1 > 0$ is the smallest eigenvalue of $\mathcal{A}$. There exists $\beta \in \mathbb{R}$ for Assumption 5.2. Furthermore, if $u_0 \in \mathcal{V}$, there exists a unique strong solution satisfying

$$u \in C([0,T];\mathcal{V}) \cap L^2(0,T;D(\mathcal{A})), \quad \frac{du}{dt} \in L^2(0,T;\mathcal{H})$$

for any $T > 0$. The semigroup $\Psi_t : \mathcal{V} \to \mathcal{V}$ is defined for $t \geq 0$. Assumption 5.1 and 5.2 hold for $\mathcal{V}$ instead of $\mathcal{H}$ with some $\rho > 0$ and $\beta \in \mathbb{R}$. Furthermore, there exists a global attractor $\mathscr{A} \subset \mathcal{V}$.

Similar results to Proposition 5.11 hold when the 2D-NSE is considered under the no-slip Dirichlet boundary condition [84, 85] on a bounded domain. For the three-dimensional Navier-Stokes equations, it is difficult to prove the existence of a global-in-time regular solution. However, for regularized versions like the Camassa-Holm or Navier-Stokes-alpha equations, the global existence and uniqueness have been proved [35, 71]. Similarly, the well-posedness results are obtained for the Leray-alpha [22], and Navier-Stokes-omega equations [64].

5.2 Reconstructing the state from partial observations

As noted in Section 1, one of the essential roles of data assimilation is to reconstruct the true state using only partial observations in a finite-dimensional space. Here, we consider the noiseless observation (3.12) on a Hilbert space to discuss the problem of partial observations in an ideal setting. This problem is related to estimating the degrees of freedom of dynamical systems based on control theory.

5.2.1 Continuous data assimilation for the Navier-Stokes equations

The problem known as "continuous data assimilation" was formulated to obtain an appropriate initial state in the numerical weather prediction using the time series of incomplete observation data, see also [21, 28]. Similar numerical studies [18, 42] investigate whether the small-scale dynamics are subordinated by the large scale dynamics in the atmospheric motions. Let us consider the 2D-NSE (5.7) on a flat torus $\mathbb{T}^2$. Any function $a : \mathbb{T}^2 \to \mathbb{R}^2$ can be represented as

$$a = \sum_{k \in \mathscr{I}} \hat{a}_k \phi_k,$$

where $\mathscr{I} = \{2\pi m \mid m \in \mathbb{Z}^2 \setminus \{0\}\}$ is the index set, $\phi_k(x) = e^{ik\cdot x}$ is the orthogonal basis of $\mathcal{H}$, and $\hat{a}_k = \overline{\hat{a}_{-k}}$ for $k \in \mathscr{I}$. We define the orthogonal projections $\mathcal{P}_\lambda$ for $\lambda > 0$ by

$$\mathcal{P}_\lambda a = \sum_{|k|^2 \le \lambda} \hat{a}_k \phi_k,$$

which are considered as the sparse observation operators with the smallest length scale $\lambda^{-\frac{1}{2}}$ and we write $\mathcal{Q}_\lambda = I - \mathcal{P}_\lambda$. This is a special case for (3.11) and (3.12) by setting $h(u) = \mathcal{P}_\lambda u$. Then, we consider two solutions u_1 and u_2 to (5.7) and decompose them into the large and small scale parts as

$$u_i(t) = p_i(t) + q_i(t), \quad p_i(t) = \mathcal{P}_\lambda u_i(t), \quad q_i = \mathcal{Q}_\lambda u_i(t), \qquad i = 1, 2. \tag{5.8}$$

In the context of the numerical weather prediction, u_1 and u_2 correspond to the true state and its approximation, respectively. Only p_1 is obtained from noiseless observations. The following algorithm, which inserts observed data directly, is called continuous data assimilation (CDA) or the synchronization filter. Let $\varpi \in \mathcal{V}$ be an initial guess for the initial state $u_1(0)$.

Definition 5.12 (CDA for the 2D-NSE on a torus)**.** *To estimate u_1, we copy the large scale part p_1 of u_1 to that of u_2, i.e.,*

$$p_2(t) = p_1(t). \tag{5.9}$$

We approximate the small-scale part q_1 by integrating

$$\frac{dq_2}{dt} + \mathcal{A}q_2 + \mathcal{Q}_\lambda \mathcal{B}((p_1 + q_2), (p_1 + q_2)) = \mathcal{Q}_\lambda f, \quad q_2(0) = \mathcal{Q}_\lambda \varpi. \tag{5.10}$$

The concept of determining modes is important for characterizing the necessary length scale for synchronization [36, 74].

Definition 5.13. *The number of determining modes is the smallest rank of the projection P_λ such that the convergence*

$$\lim_{t\to\infty} |P_\lambda u_1(t) - P_\lambda u_2(t)| = 0$$

implies that

$$\lim_{t\to\infty} |u_1(t) - u_2(t)| = 0$$

for any two solutions u_1 and u_2 to (5.7).

The following result provides the bound for λ.

Proposition 5.14 ([48, 74])**.** *Let u_1 and u_2 be two solutions of (5.7) with corresponding time-independent forcings $f_1, f_2 \in \mathcal{H}$, and initial conditions $u_1(0), u_2(0) \in \mathcal{V}$. Then, there exists a constant $c > 0$ independent of ν, f_i, and any initial conditions such that the convergence*

$$\lim_{t\to\infty} |P_\lambda u_1(t) - P_\lambda u_2(t)| = 0$$

implies that

$$\lim_{t\to\infty} \|u_1(t) - u_2(t)\| = 0$$

for any $\lambda > c\,\mathrm{Gr}(f_1)(2\pi/L)^2$. Here, $\mathrm{Gr}(f) = (L/2\pi\nu)^2 \limsup_{t\to\infty} |f(t)|$ is called the Grashof number and $\|\cdot\|$ is the H^1-norm.

The convergence rate is also estimated in [74]. This analysis provides guidelines on the a necessary dimension of observations needed to estimate the true solution even for data assimilation problems with observation noises. Korn [55, 56] studied continuous data assimilation for the regularized Navier-Stokes equations.

Remark 5.15 (Point wise observations)**.** *In real-world applications, it is often difficult to insert observation data into the model state directly. For example, if the measured data are the values of the exact solution at discrete spatial points, then the exact spatial derivatives cannot be obtained. Azouani et al. [10, 11] proposed a new algorithm for general observation operators based on the control theory, in which the observed information from the true solution is inserted into the vector field of the evolution equation through an interpolant operator.*

5.2.2 Continuous data assimilation for finite-dimensional system

For finite-dimensional dynamical systems such as the L63 and L96 equations, continuous data assimilation can also be well-defined. Let $\mathcal{H} = \mathbb{R}^{N_u}$, and $\mathcal{P} \in \mathbb{R}^{N_u \times N_u}$ be a projection matrix, where each row is the standard basis in $\mathbb{R}^{N_u}$, and set $\mathcal{Q} = I - \mathcal{P}$. Let u and v be the true solution and its approximate solution to (3.11) respectively. The observation function is given by $h(u) = \mathcal{P}u$, and the initial guess is $\varpi \in \mathbb{R}^{N_u}$. We can define the continuous data assimilation for the system as in Definition 5.12.

Definition 5.16 (CDA for ODE)**.** *To estimate u, we replace the large-scale part in v as*

$$v = \mathcal{P}u + q. \tag{5.11}$$

The small-scale part q is defined by integrating

$$\frac{dq}{dt} = \mathcal{Q}\mathcal{F}(\mathcal{P}u + q), \quad q(0) = \mathcal{Q}\varpi. \tag{5.12}$$

For the L96 equation with $J = 3J'$ for $J' \in \mathbb{N}$, we consider the projection matrix,

$$\mathcal{P} = [\phi_1, \phi_2, 0, \phi_4, \phi_5, 0, \dots] \in \mathbb{R}^{J \times J}, \tag{5.13}$$

where $(\phi_j)_{j=1}^J$ is the standard basis of $\mathbb{R}^J$. Note that $\operatorname{rank}(\mathcal{P}) = 2J' = \frac{2}{3}J$.

Proposition 5.17 (CDA for the L96 [62])**.** *Let u be a solution to the L96 equation (5.6) with an initial state $u_0 \in \mathcal{B}(\rho)$ for $\rho > 0$ given by Proposition 5.9, and v is obtained by (5.11) and (5.12) with $\mathcal{P}$ in (5.13) for the L96 equation (5.6). Then, we have*

$$\lim_{t\to\infty} |u(t) - v(t)| = 0.$$

A similar result follows for the L63 with the following projection onto the first component x as in [41]:

$$\mathcal{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{5.14}$$

5.2.3 Discrete data assimilation

Let us return to the case when $\mathcal{H}$ is infinite-dimensional. The results in the previous sections were obtained under the assumption that the observation data is continuous in time. However, in real-world applications, observation data is often obtained at discrete times. We consider a finite-rank orthogonal projection $\mathcal{P}$ on $\mathcal{H}$, and denote $\mathcal{Q} = I - \mathcal{P}$. For a solution u to (3.11), let Ψ_t be the associated semigroup. An increasing sequence $(t_n)_{n\in\mathbb{N}} \subset [0, \infty)$ represents discrete times when observation data are obtained. Let $q_0 \in \mathcal{Q}(\mathcal{H})$ be an initial guess. Hayden et al. [41] proposed discrete data assimilation as an analog of continuous data assimilation.

Definition 5.18 (Discrete data assimilation)**.** *For $n = 0$, let $v_0 = \mathcal{P}u_0 + q_0$. For each $n \in \mathbb{N}$, v_n is defined by*

$$v_n = \mathcal{P}u(t_n) + \mathcal{Q}\Psi_{t_n - t_{n-1}}(v_{n-1}), \tag{5.15}$$

and define the piecewise continuous function in time $v(t)$ by

$$v(t) = \Psi_{t-t_{n-1}}(v_{n-1}), \tag{5.16}$$

for $t \in [t_{n-1}, t_n)$.

Then, the following convergence result is established.

Proposition 5.19 (Discrete data assimilation for the L63 equation [41])**.** *Let $u(t)$ be a solution to (5.4) with $u_0 \in \mathscr{A}$ for the global attractor $\mathscr{A}$. Then, there exists $t^* = t^*(\sigma, b, \varrho) > 0$ such that the approximate solution $v(t)$ defined by (5.15) and (5.16) with $\mathcal{P}$ as in (5.14) and $t_n = \tau n$ converges to $u(t)$ for any $\tau \in (0, t^*]$, i.e.,*

$$\lim_{t\to\infty} |u(t) - v(t)| = 0.$$

Remark 5.20 ([41])**.** *If $\sigma = 10, b = 8/3, \varrho = 28$, then the time t^* is estimated by* 0.000129*.*

We have similar results for the L96 equation and the 2D-NSE on a torus.

Proposition 5.21 (Discrete data assimilation for the L96 equation [62])**.** *Let $u(t)$ be a solution to (5.4) with $u_0 \in \mathcal{B}(\rho)$ for ρ given by Proposition 5.9. Then, there exists $t^* = t^*(F, J) > 0$ such that the approximate solution $v(t)$ defined by (5.15) and (5.16) with $\mathcal{P}$ as in (5.13) and $t_n = \tau n$ converges to $u(t)$ for any $\tau \in (0, t^*]$, i.e.,*

$$\lim_{t\to\infty} |u(t) - v(t)| = 0.$$

Proposition 5.22 (Discrete data assimilation for the 2D-NSE on a torus [41])**.** *Let $u(t)$ be a solution to (5.7) with $u_0 \in \mathscr{A}$ and $q_0 \in \mathcal{V}$. Then, there exists $\lambda^* = \lambda^*(\|q_0\|, |f|, \nu, \Omega) > 0$ such that for any $\lambda > \lambda^*$, there exists $t^* = t^*(\lambda, \|q_0\|, \rho, \nu, \Omega) > 0$ such that the approximate solution $v(t)$ defined by (5.15) and (5.16) with $\mathcal{P} = \mathcal{P}_\lambda$ and $t_n = \tau n$ converges to $u(t)$ for any $\tau \in (0, t^*]$, i.e.,*

$$\lim_{t\to\infty} \|u(t) - v(t)\| = 0.$$

As a corollary, we achieve the convergence for any time interval by taking a sufficiently large $\lambda > 0$.

Corollary 5.23 ([41])**.** *For any $t^* > 0$, there exists $\lambda = \lambda(\rho, \|q_0\|, \nu, \Omega, t^*) > 0$ such that the approximate solution $v(t)$ defined by (5.15) and (5.16) with $t_n = \tau n$ converges to $u(t)$ for any $\tau \in (0, t^*]$, i.e.,*

$$\lim_{t\to\infty} \|u(t) - v(t)\| = 0.$$

5.3 Stochastic dynamical models

While our theory does not currently address stochastic model dynamics, we review existing analyses to guide future extensions of our theory to data assimilation problems with stochastic model dynamics. For the stochastic model dynamics (3.1) and (3.3), the kinetic energy principle (2.17) can be extended as shown in [83, 88], implying the boundedness of the expectation of the solution.

Assumption 5.24 (Kinetic energy principle)**.** *For the discrete-time stochastic dynamics* (3.1)*, there exists* $\lambda \in (0,1), K > 0$ *such that*

$$|\Psi(u)|^2 + \operatorname{Tr} Q \leq (1-\lambda)|u|^2 + K. \tag{5.17}$$

For the continuous-time stochastic model dynamics (3.3), we consider the analog of (2.18) in terms of the infinitesimal generator $\mathscr{L}$ associated with (3.3)

$$\mathscr{L}\mathcal{E}(u) \leq -\lambda'\mathcal{E}(u) + K', \tag{5.18}$$

for $\lambda', K' > 0$ and $\mathcal{E}(\,\cdot\,) = |\,\cdot\,|^2$. By the relations between continuous-time and discrete-time models noted in Remark 3.13, the inequality (5.18) implies that the discrete-time stochastic dynamics satisfies Assumption 5.24. Assumption 5.24 is confirmed for the shifted version L63 equation, the L96 equation, and the 2D-NSE in the same manner as in Section 5.1. These conditions ensure the boundedness of the expectation $\mathbb{E}[|u_t|^2]$. Unlike the deterministic case, however, we cannot guarantee the boundedness of any sample path.

On the other hand, to estimate the maximum error growth rate as in Lemma 5.3, the global Lipschitz condition is imposed, which is a standard assumption for analyzing SDEs.

Assumption 5.25 (Global Lipschitz condition)**.** *For the discrete-time stochastic dynamics* (3.1)*, there exists* $\beta' > 0$ *such that*

$$|\Psi(u) - \Psi(v)| \leq \beta'|u - v|, \tag{5.19}$$

for any $u, v \in \mathbb{R}^{N_u}$*. For the continuous-time stochastic dynamics* (3.3)*, there exists* $\beta > 0$ *such that*

$$|\mathcal{F}(u) - \mathcal{F}(v)| \leq \beta|u - v|, \tag{5.20}$$

for any $u, v \in \mathbb{R}^{N_u}$*.*

As examples for the dissipative stochastic dynamics, the stochastically perturbed L63 and L96 equations are considered in [30, 31], satisfying not Assumption 5.25 but the local Lipschitz condition (Assumption 5.4). In the deterministic case, whether the

condition is local or global is not a crucial since the trajectory is bounded. For the stochastic case, the boundedness of the sample path of the stochastically perturbed L96 equation is only validated numerically as mentioned in [31].

Finally, the non-degeneracy of the model noise plays an essential role in the analysis of the EnKF for stochastic model dynamics [30, 88], which will be discussed in Section 6 later.

Assumption 5.26 (Non-degeneracy of the model noise)**.** *For the state space model* (3.1) *and* (3.3)*, the noise is not degenerate, i.e.,*

$$Q \succ 0. \tag{5.21}$$

6. Mathematical analysis of the EnKF

6.1 Basic properties of the EnKF

We first discuss the EnKF in the infinite ensemble limit $m \to \infty$ since the EnKF is designed to approximate the KF using the Monte Carlo method. The following results ensure the convergence of the EnKF to the KF as $m \to \infty$ for the linear-Gaussian system (4.1). This is referred to as the consistency with the KF.

Proposition 6.1 (Consistency with the KF [58, 70])**.** *For the linear-Gaussian system (4.1), the EnKF converges to the KF. That is to say, let $p \in [1, \infty)$, for any time step $n \in \mathbb{N}$,*

$$\lim_{m\to\infty} \mathbb{E}\left[|\overline{v}_n^m - \varpi_n|^p\right] = 0,$$
$$\lim_{m\to\infty} \mathbb{E}\left[|P_n^m - P_n|_{HS}^p\right] = 0,$$

where $\overline{v}_n^m$ and P_n^m are the ensemble mean and the covariance of the EnKF respectively, ϖ_n and P_n are the mean and the covariance of the KF.

The consistency remains valid when $\mathcal{H}$ and $\mathcal{Y}$ are infinite-dimensional Hilbert spaces [51, 58]. Similarly, in the continuous-time formulation, the EnKBF is consistent with the KBF [13, 30]. It is important to note that the EnKF and EnKBF do not converge to the exact filtering distribution unless the system is linear-Gaussian. In contrast to the consistency with the KF, an analysis with a finite ensemble size m provides a differences between the PO and ESRF. It is shown that the ESRF produces less error than the PO method in approximating the mean and covariance of the analysis distribution with a finite ensemble size m [1].

Secondly, we discuss the EnKF in the infinite time limit $n \to \infty$. The filtering algorithm is said to be stable if two estimates V_n and V_n' generated by the algorithm with the same observations converge to the same solution,

$$\lim_{n\to\infty} |V_n - V_n'| = 0.$$

The stability can also be defined in terms of the filtering distribution as

$$\lim_{n\to\infty} d(\mathbb{P}^{V_n}, \mathbb{P}^{V_n'}) = 0,$$

where $d(\,\cdot\,,\,\cdot\,)$ is a distance between two probability measures. If the convergence rate is exponential, i.e., there exist constants $C > 0$ and $\gamma \in (0, 1)$ such that

$$d(\mathbb{P}^{V_n}, \mathbb{P}^{V_n'}) \leq C\gamma^n, \quad n \in \mathbb{N},$$

then this property is called the exponential stability (or the geometric ergodicity) of the filtering algorithm. In the classical theory of filtering, the controllability and the

observability of the linear-Gaussian system (4.1) ensure the stability of the Kalman filter [4, 24]. For nonlinear problems, the stability is discussed in various contexts [9, 83]. For the 3DVar, the exponential stability has been proven for the two-dimensional Navier-Stokes equations on a torus [17, 15]. The stability of the EnKF has also been established [88]. We introduce the statement of the exponential stability of the EnKF for the stochastic model dynamics (3.1) and the linear observation (3.2) with $h(u) = Hu$.

Proposition 6.2 (Exponential stability of the EnKF [88])**.** *Let U_n be the solution to (3.1) and $V_n^{(1)}, \cdots, V_n^{(m)}$ be the ensemble generated by the EnKF with observations (3.2) for $h(u) = Hu$. We consider the coupled process $\boldsymbol{X}_n = (U_n, V_n^{(1)}, \cdots, V_n^{(m)})$ as a Markov chain on $\mathcal{X} = \mathbb{R}^{N_u} \times \mathbb{R}^{N_u \times m}$, and P denotes the Markov transition kernel of the process $\boldsymbol{X}_n$. Suppose (3.1) satisfies Assumption 5.24 and Assumption 5.26, the observation noise satisfies 3.1, and that there exist constants $\lambda > 0$, $K > 0$, a positive function $\mathcal{E} : \mathcal{X} \to \mathbb{R}$ such that a sublevel set $\{\mathcal{E}(u) \le c\}$ is compact for any $c \in \mathbb{R}$ and*

$$\mathbb{E}[\mathcal{E}(\boldsymbol{X}_n)|\mathscr{F}_{n-1}^{\boldsymbol{X}}] \le (1-\lambda)\mathcal{E}(\boldsymbol{X}_{n-1}) + K, \quad n \in \mathbb{N}.$$

Then, there exists $\gamma \in (0,1)$ such that for any $\mu, \nu \in \mathcal{M}_1(\mathcal{X})$, there exists $C = C(\mu,\nu) > 0$ such that

$$d_{TV}(P^n\mu, P^n\nu) \le C_{\mu,\nu}\gamma^n, \quad n \in \mathbb{N}, \tag{6.1}$$

where $d_{TV}(\,\cdot\,,\,\cdot\,)$ is the total variation distance.

Proposition 6.2 implies that the initial errors of the EnKF will decay exponentially in time. While the filter stability is appropriate for the filtering algorithms, it does not ensure the accurate state estimation by the EnKF.

6.2 Error analysis of the filtering algorithms

We review the error analysis (or accuracy) of the EnKF for the state space model associated with an evolution equation on a Hilbert space $\mathcal{H}$,

$$\frac{du}{dt} = \mathcal{F}(u), \tag{6.2}$$

which is the same as (3.11). Suppose that a unique solution exists for any $u_0 \in \mathcal{H}$ and it generates a one-parameter semigroup $\Psi_t : \mathcal{H} \to \mathcal{H}$ for $t \ge 0$. As described in Remark 3.13, we then consider a discrete dynamical system given by

$$u_n = \Psi(u_{n-1}), \quad n \in \mathbb{N}, \tag{6.3}$$

where $\Psi = \Psi_\tau$ for a time interval $\tau > 0$. The Hilbert space $\mathcal{Y} \subset \mathcal{H}$ is the observation space. The noisy observation $y_n \in \mathcal{Y}$ is obtained by

$$y_n = Hu_n + \eta_n, \tag{6.4}$$

where $H \in \mathcal{L}(\mathcal{H}, \mathcal{Y})$ is a linear observation operator and $(\eta_n)_{n\in\mathbb{N}} \subset \mathcal{Y}$ is i.i.d. noise sequence.

6.2.1 3DVar

We start with the analysis of the 3DVar algorithm in Definition 4.9.

Proposition 6.3 ([63]). *Let* $\dim(\mathcal{H}) < \infty$ *and* u_n *be a unique solution to* (6.3) *with* $u_0 \in \mathbb{R}^m$. *Suppose that the observation noises in* (6.4) *satisfy*

$$\sup_{n\in\mathbb{N}} |\eta_n| = \epsilon, \tag{6.5}$$

and that the model covariance of the 3DVar $\widehat{P}_0$ *is chosen so that* $(I_{\mathcal{H}} - KH)\Psi : \mathbb{R}^{N_u} \to \mathbb{R}^{N_u}$ *is globally Lipschitz with a constant* $\theta \in (0, 1)$. *Then, for the sequence of the states* $(\varpi_n)_{n\in\mathbb{N}}$ *generated by Definition 4.9 with* $\widehat{P}_0$, *there exists* $c > 0$ *such that*

$$\limsup_{n\to\infty} |\varpi_n - u_n| \le \frac{c}{1-\theta}\epsilon. \tag{6.6}$$

Proof. We review the proof in [63] to explain the essence of the error analysis of data assimilation algorithms for nonlinear dynamical systems. From (4.12) and (4.13), using (6.4) gives

$$\varpi_n = \widehat{\varpi}_n + K(y_n - H\widehat{\varpi}_n) = (I_{\mathcal{H}} - KH)\Psi(\varpi_{n-1}) + KHu_n + K\eta_n.$$

From (6.3), we also have

$$u_n = (I_{\mathcal{H}} - KH)\Psi(u_{n-1}) + KH\Psi(u_{n-1}).$$

By subtracting both sides of these equalities, applying the triangle inequality, and letting $e_n = \varpi_n - u_n$, we obtain

$$|e_n| \le |(I_{\mathcal{H}} - KH)\Psi(\varpi_{n-1}) - (I_{\mathcal{H}} - KH)\Psi(u_{n-1})| + |K\eta_n| \le \theta|e_{n-1}| + c\epsilon,$$

where we use the global Lipschitz constant θ, the bound of the observation noise ϵ, and $c = |K|_{\mathcal{L}}$. When we apply the inequality successively, we have

$$|e_n| = \theta^n|e_{n-1}| + c\epsilon\frac{1-\theta^n}{1-\theta} \to \frac{c\epsilon}{1-\theta} \quad (n \to \infty).$$

□

Remark 6.4. *Similarly, we can obtain the error bound of* $\mathbb{E}[|e_n|^2]$ *if the observation noises* $(\eta_n)_{n\in\mathbb{N}}$ *satisfy Assumption 3.1.*

Proposition 6.3 requires that the global Lipschitz constant of the nonlinear map $(I_{\mathcal{H}} - KH)\Psi$ is less than 1. The combination of the following two conditions is a sufficient condition.

(1) The model dynamics Ψ is global Lipschitz continuous with a constant $\beta' > 0$ as in Assumption 5.25.

(2) The error contraction in the analysis step is estimated by $|I_{\mathcal{H}} - KH|_{\mathcal{L}} < \beta'^{-1}$.

The first condition is common in many analyses. For the second condition, from (4.7), we have

$$|I_{\mathcal{H}} - KH|_{\mathcal{L}} = |(I_{\mathcal{H}} + \widehat{P}_n H^* R^{-1} H)^{-1}|_{\mathcal{L}}.$$

Assuming full observation (Assumption 3.2) and choosing $\widehat{P}_n = P_0 = \alpha^2 I_{\mathcal{H}}$ in the 3DVar, we have

$$|(I_{\mathcal{H}} + \widehat{P}_n H^* R^{-1} H)^{-1}|_{\mathcal{L}} = |(I_{\mathcal{H}} + \alpha^2 r^{-2} I_{\mathcal{H}})^{-1}|_{\mathcal{L}} = \frac{r^2}{r^2 + \alpha^2}. \tag{6.7}$$

Thus, for any $\beta' > 1$, we can achieve $|I_{\mathcal{H}} - KH|_{\mathcal{L}} < \beta'^{-1}$ by taking sufficiently large $\alpha > 0$.

The observation matrix H is not full-rank when we consider partial observations. Hence, we cannot use this approach since it follows that $|I_{\mathcal{H}} - KH|_{\mathcal{L}} \geq 1$ if $\operatorname{Ker} H \neq \{0\}$. In general, we have the following lemma.

Lemma 6.5. *Let $\mathcal{H}$ and $\mathcal{Y}$ be Hilbert spaces. If $\widehat{P}_n H^* R^{-1} H \in \mathcal{L}(\mathcal{H})$ is not full-rank, then*

$$|I_{\mathcal{H}} - KH|_{\mathcal{L}} = |(I_{\mathcal{H}} + \widehat{P}_n H^* R^{-1} H)^{-1}|_{\mathcal{L}} \geq 1.$$

Proof. By assumption, there is $u \in \mathcal{H}$ such that $u \neq 0$ and $\widehat{P}_n H^* R^{-1} H u = 0$. Hence, we get $(I_{\mathcal{H}} + \widehat{P}_n H^* R^{-1} H)^{-1} u = u$. This implies the conclusion. □

The error bounds of the 3DVar are obtained for the dissipative dynamical systems introduced in Section 5 with partial observations. The proofs are similar to that for the discrete and continuous data assimilation in Section 5.2. Moreover, similar results also hold for the continuous-time formulation of the 3DVar. The results are summarized in Table 1.

Algorithm \ Model	L63	L96	2D-NSE
3DVar	[61]	[62]	[17]
Continuous-time 3DVar	[61]	[62]	[15]

Table 1: References of the accuracy results of the 3DVar with partial observations.

6.2.2 PO method

We have two issues when considering the error bound of the EnKF.

(i-1) In the EnKF, we want to estimate the Lipschitz constant $L > 0$ such that

$$|u_n - \overline{\widehat{v}}_n| \leq L |u_{n-1} - \overline{v}_{n-1}|.$$

However, we cannot apply Lemma 5.3 since $\Psi(\overline{v}_{n-1}) \neq \overline{\widehat{v}}_n$ if Ψ is nonlinear.

(i-2) In contrast to the 3DVar, the prediction covariance $\widehat{P}_n$ is not trivial. Therefore, it is difficult to estimate $|I_{\mathcal{H}} + \widehat{P}_n H^* R^{-1} H|_{\mathcal{L}}$ even when the full observation is considered.

The following lemma provides an estimate of the error contraction in the EnKF.

Lemma 6.6. *Suppose Assumption 3.2 holds. Then, the prediction covariance $\widehat{P}_n$ in the EnKF satisfies the following inequality*

$$|I_{\mathcal{H}} - K_n H|_{\mathcal{L}} = |(I_{\mathcal{H}} + r^{-2}\widehat{P}_n)^{-1}|_{\mathcal{L}} \leq 1.$$

Proof. Since $\widehat{P}_n \succeq 0$ by definition, the result follows as a consequence of Lemma 2.3 and Lemma 2.19. □

Kelly et al. [54] consider the error estimate of the state of each ensemble member by defining $e_n^{(k)} = v_n^{(k)} - u_n$ for $k = 1, \dots, m$. They estimate the error growth for each member using Lemma 5.3 in the same manner as the 3DVar to avoid the issue (i-1). They first prove the well-posedness of the PO method based on Lemma 6.6, which implies that the error does not blow up in finite-time.

Proposition 6.7 (Well-posedness of the PO method, modified [54])**.** *Let Assumption 5.1 and Assumption 5.2 for the model dynamics (6.2), and Assumption 3.1 and Assumption 3.2 for the observation (6.4) hold. Let u_n be the solution to (6.3) with $u_0 \in \mathcal{B}(\rho)$, and let $\boldsymbol{V}_n$ be generated by the PO method in Definition 4.10. Then, we have*

$$\mathbb{E}\left[|e_n^{(k)}|^2\right] \leq e^{2\beta\tau n}\mathbb{E}\left[|e_0^{(k)}|^2\right] + 2mr^2\frac{e^{2\beta\tau n} - 1}{e^{2\beta\tau} - 1} \tag{6.8}$$

for $k = 1, \dots, m$ and $n \in \mathbb{N}$.

Proposition 6.7 does not ensure a uniform-in-time error bound when the model dynamics is chaotic, i.e., $\beta > 0$. Hence, they apply additive inflation as described in Definition 4.19. For the full observation with Assumption 3.2, we get the estimate for the additively inflated covariance $\widehat{P}_n^{\alpha}$ as (6.7).

$$\begin{aligned}|(I_{\mathcal{H}} + \widehat{P}_n^{\alpha} H^* R^{-1} H)^{-1}|_{\mathcal{L}} &= |(I_{\mathcal{H}} + r^{-2}(\widehat{P}_n + \alpha^2 I_{\mathcal{H}}))^{-1}|_{\mathcal{L}} \\ &\leq |(I_{\mathcal{H}} + \alpha^2 r^{-2} I_{\mathcal{H}})^{-1}|_{\mathcal{L}} = \frac{r^2}{r^2 + \alpha^2}.\end{aligned}$$

This bound plays a significant role in the error estimate of the PO method with additive inflation, which is stated as follows.

Proposition 6.8 (Accuracy of the PO method with the additive inflation, modified [54])**.** *Let Assumptions 5.1 and 5.2 for the model dynamics (6.2), and Assumptions 3.1 and 3.2 for the observation (6.4) hold. Let u_n be the solution to (6.3) with $u_0 \in \mathcal{B}(\rho)$*

and let $\boldsymbol{V}_n$ *be generated by the PO method in Definition 4.10 with the additive inflation in Definition 4.19 for* $\alpha \geq 0$*. Define* $\theta = (1 + \frac{\alpha^2}{r^2})^{-2} e^{2\beta h} < 1$*. Then, we have*

$$\mathbb{E}\left[|e_n^{(k)}|^2\right] \leq \theta^n \mathbb{E}\left[|e_0^{(k)}|^2\right] + 2mr^2 \frac{1-\theta^n}{1-\theta} \tag{6.9}$$

for $k = 1, \ldots, m$ *and* $n \in \mathbb{N}$*. In particular, if* $\theta < 1$ *then*

$$\lim_{n\to\infty} \mathbb{E}\left[|e_n^{(k)}|^2\right] \leq \frac{2mr^2}{1-\theta}.$$

Remark 6.9. *Proposition 6.7 and Proposition 6.8 hold with Assumption 3.11 instead of Assumption 3.1 owing to the following lemma.*

Lemma 6.10. *Let* $\mathcal{H}$ *be a Hilbert space,* $\eta \sim \mathcal{N}(0, \widetilde{R}), \widetilde{R} \in \mathcal{L}_{sa}(\mathcal{H}), \widetilde{R} \succeq 0, \operatorname{Tr} \widetilde{R} < \infty, \widehat{R} \preceq r^2 I_{\mathcal{H}}$ *for* $r > 0$*. The operator* $\mathcal{P} : \mathcal{H} \to \mathcal{H}$ *is an orthogonal projection with* $\dim(\mathcal{P}(\mathcal{H})) < \infty$*. Then,*

$$\mathbb{E}[|\mathcal{P}\eta|^2] \leq \dim(\mathcal{P}(\mathcal{H}))r^2.$$

Proof. Let $N_{\mathcal{P}} = \dim(\mathcal{P}(\mathcal{H}))$. Then, we have $\mathcal{P}\eta \sim \mathcal{N}(0, \mathcal{P}\widetilde{R}\mathcal{P})$ and

$$\mathbb{E}[|\mathcal{P}\eta|^2] = \operatorname{Tr}(\mathcal{P}\widetilde{R}\mathcal{P}) \leq \operatorname{Tr}(\mathcal{P}(r^2 I_{\mathcal{H}})\mathcal{P}) = r^2 \operatorname{Tr}(\mathcal{P}) = N_{\mathcal{P}} r^2.$$

□

Next, we consider the continuous-time formulation of the EnKF, the EnKBF defined in Definition 4.25, in a finite-dimensional state space $\mathcal{H} = \mathbb{R}^{N_u}$ and a finite-dimensional observation space $\mathcal{Y} = \mathbb{R}^{N_y}$. The uniform-in-time error bound is also obtained for the EnKBF with full observations and small noises.

Proposition 6.11 (Accuracy of the EnKBF [30])**.** *Let Assumptions 5.25, 5.26 for the stochastic model dynamics (3.3), and Assumption 3.2 for the observation (3.4) hold. Suppose that the observation noise variance* r^2 *is chosen to be sufficiently small. The initial ensemble* $\boldsymbol{V}_0 \in \mathbb{R}^{N_u \times m}$ *is chosen so that the initial ensemble covariance* $P_0 \in \mathbb{R}^{N_u \times N_u}$ *is invertible, and the bounds* $\lambda_{max}(P_0) \leq C_1 r$ *and* $\lambda_{min}(P_0) \geq C_2 r$ *are satisfied for* $C_1, C_2 > 0$*. Let* U_t *be the solution to (3.3) and* $\boldsymbol{V}_t \in \mathbb{R}^{N_u \times m}$ *be the ensemble generated by Definition 4.25. Then, we have*

$$\lim_{t\to\infty} \mathbb{E}[|U_t - \overline{v}_t|^2] = O(r). \tag{6.10}$$

Without any inflation technique, this result holds owing to the non-degeneracy of the model noise (Assumption 5.26).

7. Error analysis of the ESRF

7.1 Well-posedness of the ESRF

We consider the state space model given by (6.2)-(6.4) in Section 6.2. Let the ensemble of state estimation errors be defined as $\boldsymbol{E}_n = [e_n^{(k)}]_{k=1}^m \in \mathcal{H}^m$ and let $\mathscr{F}_n$ be the σ-algebra generated by initial uncertainties in $\boldsymbol{V}_0$ and the observation noise sequence $(\eta_k)_{k=1}^n$ for $n \in \mathbb{N}$. We establish the well-posedness of the ETKF.

Theorem 7.1 (Well-posedness of the ETKF [82])**.** *Suppose that Assumption 5.1 and Assumption 5.4 are satisfied by the model dynamics (6.2), and Assumption 3.2 and Assumption 3.11 are satisfied by the observation (6.4). Let u_n be the solution to (6.3) with $u_0 \in B(\rho)$, and let $\boldsymbol{V}_n$ be generated by the ETKF (Definition 4.13). Then, we have the following upper bound.*

$$\mathbb{E}\left[|\boldsymbol{E}_n|_2^2\right] \leq e^{2\beta h n}\mathbb{E}\left[|\boldsymbol{E}_0|_2^2\right] + (m-1)r^2\frac{e^{2\beta h n}-1}{e^{2\beta h}-1}, \quad n \in \mathbb{N}. \tag{7.1}$$

We need the following lemma for the analysis step of the ETKF in Definition 4.13, and it is a variant of Proposition 3.2 in [54] for the PO method.

Lemma 7.1. *The following holds for the transform matrix*

$$d\boldsymbol{V}_n\mathbf{1}^* = d\widehat{\boldsymbol{V}}_n\mathbf{1}^* = 0 \in \mathcal{H}, \tag{7.2}$$

$$T_n\mathbf{1}^* = \mathbf{1}^*. \tag{7.3}$$

Moreover, the ensembles satisfy the relation

$$(I_{\mathcal{H}} + \widehat{P}_n H^* R^{-1} H)\boldsymbol{V}_n = \widehat{\boldsymbol{V}}_n T_n^{-1} + \widehat{P}_n H^* R^{-1} y_n \mathbf{1}. \tag{7.4}$$

Proof. We omit the time index n in the following proofs for simplicity since $n \in \mathbb{N}$ is fixed. The equality (2.6) in Lemma 2.16 yields

$$S^{-1}\mathbf{1}^* = \left(I_m + \frac{1}{m-1}d\boldsymbol{V}^* H^* R^{-1} H d\boldsymbol{V}\right)\mathbf{1}^* = \mathbf{1}^*,$$

where $S = T^2$ defined by (4.24). Hence, we have

$$S\mathbf{1}^* = \mathbf{1}^*. \tag{7.5}$$

Then, we prove that $\mathbf{1}^*$ is also an eigenvector of $T = S^{\frac{1}{2}}$ with an eigenvalue 1. Since S is symmetric, it is diagonalized as $S = UDU^*$ with a unitary matrix $U \in \mathbb{R}^{m\times m}$ and a diagonal matrix $D \in \mathbb{R}^{m\times m}$. Then, (7.5) is equivalent to

$$S\mathbf{1}^* = \mathbf{1}^* \Leftrightarrow UDU^*\mathbf{1}^* = \mathbf{1}^* \Leftrightarrow DU^*\mathbf{1}^* = U^*\mathbf{1}^*.$$

Putting $u = U^*\mathbf{1}^* = (u^1, \cdots, u^m)^* \in \mathbb{R}^m$ and $d^j > 0$ as j th diagonal element of D for $j = 1, \ldots, m$, we rewrite the last equality for each component as

$$d^j u^j = u^j, \quad j = 1, \ldots, m.$$

This implies that $d^j = 1$ or $u^j = 0$ for each $j = 1, \ldots, m$. Hence, we have

$$(d^j)^{\frac{1}{2}} u^j = u^j, \quad j = 1, \ldots, m,$$

and $D^{\frac{1}{2}} U^* \mathbf{1}^* = U^* \mathbf{1}^*$. By definition, T is written as $T = UD^{\frac{1}{2}}U^*$, and this yields $T\mathbf{1}^* = \mathbf{1}^*$, which is (7.3).

The last equality (7.4) is shown as follows. From (4.23), we have

$$(I_{\mathcal{H}} + \widehat{P}H^*R^{-1}H)\overline{v} = \overline{\widehat{v}} + \widehat{P}H^*R^{-1}y \in \mathcal{H}.$$

By using $\widehat{P} = \mathrm{Cov}_m(\widehat{\boldsymbol{V}})$, (4.24) and $S = T^2$, we obtain

$$\begin{aligned}(I_{\mathcal{H}} + \widehat{P}H^*R^{-1}H)d\boldsymbol{V} &= (I_{\mathcal{H}} + \widehat{P}H^*R^{-1}H)d\widehat{\boldsymbol{V}}T = d\widehat{\boldsymbol{V}}\left[I_m + \frac{1}{m-1}d\widehat{\boldsymbol{V}}^*H^*R^{-1}Hd\widehat{\boldsymbol{V}}\right]T \\ &= d\widehat{\boldsymbol{V}}S^{-1}T = d\widehat{\boldsymbol{V}}T^{-1} \in \mathcal{H}^m.\end{aligned}$$

Finally, owing to (4.23) and $\overline{\widehat{v}}\mathbf{1}T^{-1} = \overline{\widehat{v}}\mathbf{1}$,

$$\begin{aligned}(I_{\mathcal{H}} + \widehat{P}H^*R^{-1}H)V &= (I_{\mathcal{H}} + \widehat{P}H^*R^{-1}H)(\overline{v}\mathbf{1} + d\boldsymbol{V}) = \overline{\widehat{v}}\mathbf{1} + \widehat{P}H^*R^{-1}y\mathbf{1} + d\widehat{\boldsymbol{V}}T^{-1} \\ &= \overline{\widehat{v}}\mathbf{1}T^{-1} + d\widehat{\boldsymbol{V}}T^{-1} + \widehat{P}H^*R^{-1}y\mathbf{1} = \widehat{\boldsymbol{V}}T^{-1} + \widehat{P}H^*R^{-1}y\mathbf{1}.\end{aligned}$$

This finishes the proof. □

(Proof of Theorem 7.1). From Assumption 3.2, the relation (7.4) is equivalent to

$$(I_{\mathcal{H}} + r^{-2}\widehat{P}_n)\boldsymbol{V}_n = \widehat{\boldsymbol{V}}_n T_n^{-1} + r^{-2}\widehat{P}_n y_n \mathbf{1}. \tag{7.6}$$

Let $\boldsymbol{U}_n = u_n\mathbf{1} \in \mathcal{H}^m$. From (7.3), we have $\boldsymbol{U}_n = \boldsymbol{U}_n T_n^{-1}$. Hence,

$$(I_{\mathcal{H}} + r^{-2}\widehat{P}_n)\boldsymbol{U}_n = \boldsymbol{U}_n + r^{-2}\widehat{P}_n\boldsymbol{U}_n = \boldsymbol{U}_n T_n^{-1} + r^{-2}\widehat{P}_n\boldsymbol{U}_n. \tag{7.7}$$

Setting $\widehat{\boldsymbol{E}}_n = \widehat{\boldsymbol{V}}_n - \boldsymbol{U}_n$ and subtracting (7.7) from (7.6) yields

$$(I_{\mathcal{H}} + r^{-2}\widehat{P}_n)\boldsymbol{E}_n = \widehat{\boldsymbol{E}}_n T_n^{-1} + r^{-2}\widehat{P}_n(y_n - u_n)\mathbf{1} = \widehat{\boldsymbol{E}}_n T_n^{-1} + r^{-2}\widehat{P}_n\eta_n\mathbf{1}.$$

Owing to $r^{-2}\widehat{P}_n \succeq 0$, $I_{\mathcal{H}} + r^{-2}\widehat{P}_n$ is invertible. Hence, multiplying $(I_{\mathcal{H}} + r^{-2}\widehat{P}_n)^{-1}$, we obtain

$$\boldsymbol{E}_n = (I_{\mathcal{H}} + r^{-2}\widehat{P}_n)^{-1}\widehat{\boldsymbol{E}}_n T_n^{-1} + (I_{\mathcal{H}} + r^{-2}\widehat{P}_n)^{-1} r^{-2}\widehat{P}_n\eta_n\mathbf{1}.$$

Let us divide $\boldsymbol{E}_n$ into the following two terms and evaluate them separately.

$$R_1 = (I_{\mathcal{H}} + r^{-2}\widehat{P}_n)^{-1}\widehat{\boldsymbol{E}}_n T_n^{-1}, \tag{7.8}$$

$$R_2 = (I_{\mathcal{H}} + r^{-2}\widehat{P}_n)^{-1} r^{-2}\widehat{P}_n \eta_n \mathbf{1}. \tag{7.9}$$

Here, the dimension of $\mathrm{Ran}(P_n)$ is $m-1$ at most since $\widehat{P}_n$ consists of m vectors with one constraint. Let Π_n be the projection from $\mathcal{H}$ to $\mathrm{Ran}(P_n)$, and we have $R_2 = (I_{\mathcal{H}} + r^{-2}\widehat{P}_n)^{-1} r^{-2}\widehat{P}_n \Pi_n \eta_n \mathbf{1}$. From (2.10), we have $(I_{\mathcal{H}} + r^{-2}\widehat{P}_n)^{-1} r^{-2}\widehat{P}_n \preceq I$ owing to $r^{-2}\widehat{P}_n \succeq 0$. This leads to

$$|R_2|_2^2 \le |\Pi_n \eta_n \mathbf{1}|_2^2 = |\Pi_n \eta_n|^2. \tag{7.10}$$

Let $J = I_{\mathcal{H}} + r^{-2}\widehat{P}_n$. Then we have $J, J^{-1} \in \mathcal{L}_{sa}(\mathcal{H})$ and $|J^{-1}|_{\mathcal{L}} \le 1$. We obtain

$$R_1 R_1^* = J^{-1}\widehat{\boldsymbol{E}}_n T_n^{-2} \widehat{\boldsymbol{E}}_n^* J^{-1}.$$

Considering the relations $d\widehat{\boldsymbol{E}}_n = d\widehat{\boldsymbol{V}}_n$ and $\widehat{\boldsymbol{E}}_n d\widehat{\boldsymbol{V}}_n^* = d\widehat{\boldsymbol{V}}_n \widehat{\boldsymbol{E}}_n^* = d\widehat{\boldsymbol{V}}_n d\widehat{\boldsymbol{V}}_n^*$, we have

$$\begin{aligned}
\widehat{\boldsymbol{E}}_n T_n^{-2} \widehat{\boldsymbol{E}}_n^* &= \widehat{\boldsymbol{E}}_n \left[I_m + \frac{r^{-2}}{m-1} d\widehat{\boldsymbol{V}}_n^* d\widehat{\boldsymbol{V}}_n \right] \widehat{\boldsymbol{E}}_n^* \\
&= \widehat{e}_n \mathbf{1} (\widehat{e}_n \mathbf{1})^* + d\widehat{\boldsymbol{V}}_n d\widehat{\boldsymbol{V}}_n^* + d\widehat{\boldsymbol{V}}_n d\widehat{\boldsymbol{V}}_n^* r^{-2}\widehat{P}_n \\
&= \widehat{e}_n \mathbf{1} (\widehat{e}_n \mathbf{1})^* + d\widehat{\boldsymbol{V}}_n d\widehat{\boldsymbol{V}}_n^* J,
\end{aligned}$$

where $\widehat{e}_n = \overline{\widehat{v}}_n - u_n$. Since J^{-1} is self-adjoint, we have

$$R_1 R_1^* = J^{-1}\widehat{e}_n \mathbf{1} (\widehat{e}_n \mathbf{1})^* J^{-1} + J^{-1} d\widehat{\boldsymbol{V}}_n d\widehat{\boldsymbol{V}}_n^* = J^{-1}\widehat{e}_n \mathbf{1} (J^{-1}\widehat{e}_n \mathbf{1})^* + J^{-1} d\widehat{\boldsymbol{V}}_n d\widehat{\boldsymbol{V}}_n^*.$$

Then, $|R_1|_2^2$ is bounded by

$$\begin{aligned}
|R_1|_2^2 &= |J^{-1}\widehat{e}_n|^2 + \frac{1}{m}\mathrm{Tr}(J^{-1} d\widehat{\boldsymbol{V}}_n d\widehat{\boldsymbol{V}}_n^*) \le |J^{-1}|_{\mathcal{L}}^2 |\widehat{e}_n|^2 + |J^{-1}|_{\mathcal{L}} \frac{1}{m}\mathrm{Tr}(d\widehat{\boldsymbol{V}}_n d\widehat{\boldsymbol{V}}_n^*) \\
&\le |\widehat{e}_n|^2 + |d\widehat{\boldsymbol{V}}_n|_2^2 = |\widehat{\boldsymbol{E}}_n|_2^2.
\end{aligned}$$

The first inequality follows from Lemma 2.12, and the second inequality holds owing to $|J^{-1}|_{\mathcal{L}} \le 1$. From this with Lemma 5.3, we obtain the upper bound of $|R_1|_2$ as follows.

$$|R_1|_2^2 \le |\widehat{\boldsymbol{E}}_n|_2^2 \le e^{2\beta h} |\boldsymbol{E}_{n-1}|_2^2. \tag{7.11}$$

Since R_1 and R_2 are conditionally independent under $\mathscr{F}_{n-1}$, it follows from (7.10) and (7.11) that

$$\begin{aligned}
\mathbb{E}_{n-1}[|\boldsymbol{E}_n|_2^2] &= \mathbb{E}_{n-1}[|R_1|_2^2] + \mathbb{E}_{n-1}[|R_2|_2^2] \le e^{2\beta h}\mathbb{E}_{n-1}[|\boldsymbol{E}_{n-1}|_2^2] + \mathbb{E}_{n-1}[|\Pi_n \eta_n|^2] \\
&= e^{2\beta h}\mathbb{E}_{n-1}[|\boldsymbol{E}_{n-1}|_2^2] + (m-1)r^2,
\end{aligned}$$

where the conditional expectation is denoted by $\mathbb{E}_{n-1}[\,\cdot\,] \coloneqq \mathbb{E}[\,\cdot \mid \mathscr{F}_{n-1}]$. Taking the expectation yields

$$\mathbb{E}[|\boldsymbol{E}_n|_2^2] \le e^{2\beta h}\mathbb{E}[|\boldsymbol{E}_{n-1}|_2^2] + (m-1)r^2$$

since the conditional expectation satisfies $\mathbb{E}[\mathbb{E}_{n-1}[\,\cdot\,]] = \mathbb{E}[\,\cdot\,]$ in general. Applying this inequality repeatedly, we obtain (7.1). □

We also have the well-posedness of the EAKF.

Theorem 7.2 (Well-posedness of the EAKF)**.** *Under the same conditions as in Theorem 7.1, for $\boldsymbol{V}_n$ generated by the EAKF (Definition 4.16), the same error bound as in (7.1) holds.*

Proof. The proof is the same as that for the ETKF except for the estimate of R_1. As before, we omit the time index n in the following proofs for simplicity since $n \in \mathbb{N}$ is fixed. It is clear that

$$u = (I_{\mathcal{H}} - K)u + Ku. \tag{7.12}$$

In the EAKF, the analysis ensemble is given by

$$\boldsymbol{V} = \overline{v}\mathbf{1} + d\boldsymbol{V} = \left[(I_{\mathcal{H}} - K)\overline{\widehat{v}} + Ky\right]\mathbf{1} + Ad\widehat{\boldsymbol{V}}. \tag{7.13}$$

By subtracting (7.13) from (7.12), we have

$$\boldsymbol{E} = (I_{\mathcal{H}} - K)\widehat{e}\mathbf{1} + Ad\widehat{\boldsymbol{V}} + K(y - u)\mathbf{1}. \tag{7.14}$$

Recalling Lemma 4.4, we have $(I_{\mathcal{H}} - K)^{-1} = I_{\mathcal{H}} + r^{-2}\widehat{P} = J$. In addition, Lemma 4.14 yields $K = (I_{\mathcal{H}} - K)r^{-2}\widehat{P} = (I_{\mathcal{H}} + r^{-2}\widehat{P})^{-1}r^{-2}\widehat{P}$. Therefore, the equality (7.14) becomes

$$\boldsymbol{E} = J^{-1}\widehat{e}\mathbf{1} + Ad\widehat{\boldsymbol{V}} + (I_{\mathcal{H}} + r^{-2}\widehat{P})^{-1}r^{-2}\widehat{P}\eta\mathbf{1}.$$

We divide the ensemble of the error $\boldsymbol{E}$ as $\boldsymbol{E} = R_1 + R_2$ with

$$\begin{aligned} R_1 &= J^{-1}\widehat{e}\mathbf{1} + Ad\widehat{\boldsymbol{V}}, \\ R_2 &= (I_{\mathcal{H}} + r^{-2}\widehat{P})^{-1}r^{-2}\widehat{P}\eta\mathbf{1}. \end{aligned}$$

Note that R_2 is the same as in the ETKF. Since $J = (I_{\mathcal{H}} - K)^{-1}$, it follows from the definition of the adjustment operator (4.27) that

$$Ad\widehat{\boldsymbol{V}}(Ad\widehat{\boldsymbol{V}})^* = (m-1)(I_{\mathcal{H}} - K)\widehat{P} = J^{-1}d\widehat{\boldsymbol{V}}d\widehat{\boldsymbol{V}}^*.$$

Considering $d\widehat{\boldsymbol{V}}\mathbf{1}^* = 0$, we have

$$R_1R_1^* = J^{-1}\widehat{e}\mathbf{1}(J^{-1}\widehat{e}\mathbf{1})^* + Ad\widehat{\boldsymbol{V}}d\widehat{\boldsymbol{V}}^*A^* = J^{-1}\widehat{e}\mathbf{1}(J^{-1}\widehat{e}\mathbf{1})^* + J^{-1}d\widehat{\boldsymbol{V}}d\widehat{\boldsymbol{V}}^*.$$

This equal to $R_1R_1^*$ in the proof for the ETKF. □

Remark 7.2. *Theorem 7.1 and Theorem 7.2 ensure the accuracy of the ESRF in a short time interval. Additionally, these results hold even when $\mathcal{H}$ is infinite-dimensional and they are analogous to Proposition 6.7 for the PO method. Compared to Proposition 6.7, a constant coefficient 2 does not appear in the second term of the upper bound (7.1) in Theorem 7.1 and Theorem 7.2, indicating that the PO method is more influenced by the variance of observation noises due to its stochastic implementation.*

7.2 Uniform-in-time error bound of the ESRF

We estimate the uniform-in-time error bound of the analysis error $e_n = \overline{v}_n - u_n$. Due to the issue (i-1) in Section 6.2.2, we need to use Lemma 5.5 instead of Lemma 5.3 to estimate the error growth for the ensemble mean.

Theorem 7.3 (Accuracy of the ETKF with the multiplicative inflation [82])**.** *Suppose* $\dim(\mathcal{H}) = N_u < \infty$. *Assumption 5.1 and Assumption 5.4 are satisfied by the model dynamics (6.2), and Assumption 3.11 and Assumption 3.2 are satisfied by the observation (6.4). Let* u_n *be the solution to (6.3) with* $u_0 \in B(\rho)$*, and let* $\boldsymbol{V}_n$ *be generated by the ETKF in Definition 4.13 with the multiplicative inflation in Definition 4.20 for* $\alpha \geq 1$*. In addition, suppose that the ensemble size* $m \in \mathbb{N}$ *is large enough to satisfy* $\lambda_{min}(P_0) \geq \lambda_0$ *with* $\lambda_0 > 0$*, and that* $v_n^{(k)} \in B(\rho)$ *for* $k = 1, \dots, m$ *and* $n \in \mathbb{N}$*. Then, for any* $\epsilon > 0$*, there exists* $\alpha_0 = \alpha_0(\rho, \beta, m, \lambda_0, \tau, r, \epsilon) \geq 1$ *such that the following hold for any* $\alpha \geq \alpha_0$*.*

(i) There exists $\lambda_* = \lambda_*(\rho, \beta, m, \lambda_0, \tau, r, \alpha) > 0$ *such that* $\lambda_{min}(\widehat{P}_n) > \lambda_*$ *for all* $n \in \mathbb{N}$*.*

(ii) For $n \in \mathbb{N}$ *and* $\theta = (1 + \frac{\alpha^2}{r^2}\lambda_*)^{-2} e^{2(\beta+\epsilon)\tau}$*, we have*

$$\mathbb{E}[|e_n|^2] \leq \theta^n (\mathbb{E}[|e_0|^2] + D) + N_u r^2 \frac{1-\theta^n}{1-\theta} + \left(\frac{(1-\theta^n)(1-\Theta)}{1-\theta} - 1 \right) D, \quad (7.15)$$

where $D = \frac{\beta^2 \rho^2}{(\beta+\epsilon)\epsilon}$ *and* $\Theta = (1 + \frac{\alpha_0^2}{r^2}\lambda_*)^{-2}$*. Moreover, if* $\theta < 1$*, we have*

$$\lim_{n\to\infty} \mathbb{E}[|e_n|^2] \leq \frac{N_u r^2}{1-\theta} + \left(\frac{1-\Theta}{1-\theta} - 1 \right) D. \quad (7.16)$$

Proof. For simplicity, we write $\widehat{\lambda}_n^{min} = \lambda_{min}(\widehat{P}_n)$ and $\lambda_n^{min} = \lambda_{min}(P_n)$. We first estimate the change of the eigenvalue from λ_{n-1}^{min} to $\widehat{\lambda}_n^{min}$ in the prediction step. To this end, we interpolate the prediction step as $\widehat{v}_t^{(k)} = \Psi_t(v_{n-1}^{(k)})$, $\widehat{P}_t = \frac{1}{m-1}\sum_{k=1}^m (\widehat{v}_t^{(k)} - \overline{\widehat{v}}_t) \otimes (\widehat{v}_t^{(k)} - \overline{\widehat{v}}_t)$, $\lambda_t = \lambda_{min}(\widehat{P}_t)$ for $t \in [0, \tau]$. Note that $\widehat{v}_t^{(k)} \in \mathcal{B}(\rho)$ from Assumption 5.1. The differentiation of $\widehat{P}_t$ with respect to t yields

$$\frac{d}{dt}\widehat{P}_t = \frac{1}{m-1}\sum_{k=1}^m (F(\widehat{v}_t^{(k)}) - \overline{F}_t) \otimes (\widehat{v}_t^{(k)} - \overline{\widehat{v}}_t) + (\widehat{v}_t^{(k)} - \overline{\widehat{v}}_t) \otimes (F(\widehat{v}_t^{(k)}) - \overline{F}_t), \quad (7.17)$$

where $\overline{F}_t = \frac{1}{m}\sum_{k=1}^m F(\widehat{v}_t^{(k)})$. Owing to $\frac{1}{m}\sum_{k=1}^m \widehat{v}_t^{(k)} - \overline{\widehat{v}}_t = 0$, we have

$$\frac{d}{dt}\widehat{P}_t = \frac{1}{m-1}\sum_{k=1}^m (F(\widehat{v}_t^{(k)}) - F(\overline{\widehat{v}}_t)) \otimes (\widehat{v}_t^{(k)} - \overline{\widehat{v}}_t) + (\widehat{v}_t^{(k)} - \overline{\widehat{v}}_t) \otimes (F(\widehat{v}_t^{(k)}) - F(\overline{\widehat{v}}_t)). \quad (7.18)$$

It follows from Lemma 2.15 that, for $t \in (0, \tau)$, there exists $w = w_t \in \mathcal{H}$ with $|w| = 1$ such that

$$\frac{d}{dt}\lambda_t = \left\langle w, \frac{d}{dt}\widehat{P}_t w \right\rangle.$$

To derive the lower bound of $\frac{d}{dt}\lambda_t$, we consider the absolute value of the right-hand side of (7.18). Owing to $|w| = 1$ and Assumption 5.4, we have

$$\begin{aligned}
\left|\left\langle w, \frac{d}{dt}\widehat{P}_t w \right\rangle\right| &\le \left| \frac{2}{m-1} \sum_{k=1}^{m} \left\langle F(\widehat{v}_t^{(k)}) - F(\overline{\widehat{v}}_t), w \right\rangle \left\langle \widehat{v}_t^{(k)} - \overline{\widehat{v}}_t, w \right\rangle \right| \\
&\le 2 \left(\frac{1}{m-1} \sum_{k=1}^{m} \left\langle F(\widehat{v}_t^{(k)}) - F(\overline{\widehat{v}}_t), w \right\rangle^2 \right)^{\frac{1}{2}} \left(\frac{1}{m-1} \sum_{k=1}^{m} \left\langle \widehat{v}_t^{(k)} - \overline{\widehat{v}}_t, w \right\rangle^2 \right)^{\frac{1}{2}} \\
&\le \beta \frac{1}{m-1} \sum_{k=1}^{m} |\widehat{v}_t^{(k)} - \overline{\widehat{v}}_t|^2 \le 8 \frac{m}{m-1} \beta \rho^2.
\end{aligned}$$

The last inequality holds owing to $\widehat{v}_t^{(k)} \in B(\rho)$ and the assumption of Theorem 7.3. Hence, it follows that

$$\frac{d}{dt}\lambda_t \ge -a$$

with $a = 8\frac{m}{m-1}\beta\rho^2 > 0$. Integrating it from $t = 0$ to τ, we have

$$\widehat{\lambda}_n^{min} = \lambda_\tau \ge e^{-a\tau}\lambda_0 = e^{-a\tau}\lambda_{n-1}^{min}. \tag{7.19}$$

The next step is to address the change of the eigenvalue in the analysis step. From Assumption 3.2 and (4.36), we have

$$P_{n-1} = \frac{\alpha^2}{m-1} d\widehat{\boldsymbol{V}}_{n-1}(I_m + \alpha^2\gamma^{-2}\widetilde{P}_{n-1})^{-1} d\widehat{\boldsymbol{V}}_{n-1}^*,$$

where $\widetilde{P}_{n-1} = \frac{1}{m-1} d\widehat{\boldsymbol{V}}_{n-1}^* d\widehat{\boldsymbol{V}}_{n-1} \in \mathbb{R}^{m\times m}$. Next, for fixed $n \in \mathbb{N}$, we show that the eigenvectors of $\widehat{P}_{n-1}$ are also the eigenvectors of P_{n-1}. Indeed, if $\phi \in \mathcal{H}$ satisfies $\widehat{P}_{n-1}\phi = \lambda\phi$ with an eigenvalue $\lambda \ge 0$, we have

$$\widetilde{P}_{n-1} d\widehat{\boldsymbol{V}}_{n-1}^*\phi = \frac{1}{m-1} d\widehat{\boldsymbol{V}}_{n-1}^* d\widehat{\boldsymbol{V}}_{n-1} d\widehat{\boldsymbol{V}}_{n-1}^* = d\widehat{\boldsymbol{V}}_{n-1}^* \widehat{P}_{n-1}\phi = \lambda d\widehat{\boldsymbol{V}}_{n-1}^*\phi.$$

Hence, it follows that

$$\begin{aligned}
P_{n-1}\phi &= \frac{\alpha^2}{m-1} d\widehat{\boldsymbol{V}}_{n-1}(I_m + \alpha^2\gamma^{-2}\widetilde{P}_{n-1})^{-1} d\widehat{\boldsymbol{V}}_{n-1}^*\phi \\
&= \frac{\alpha^2}{m-1} d\widehat{\boldsymbol{V}}_{n-1} \frac{1}{1+\alpha^2\gamma^{-2}\lambda} d\widehat{\boldsymbol{V}}_{n-1}^*\phi \\
&= \frac{\alpha^2}{1+\alpha^2\gamma^{-2}\lambda}\widehat{P}_{n-1}\phi = \frac{\alpha^2\lambda}{1+\alpha^2\gamma^{-2}\lambda}\phi.
\end{aligned}$$

Since the map $\lambda \mapsto \frac{\alpha^2\lambda}{1+\alpha^2\gamma^{-2}\lambda}$ is monotonically increasing, we obtain the relation between the minimum eigenvalues

$$\lambda_{n-1}^{min} = \frac{\alpha^2\widehat{\lambda}_{n-1}^{min}}{1+\frac{\alpha^2}{\gamma^2}\widehat{\lambda}_{n-1}^{min}}. \tag{7.20}$$

Combining (7.19) and (7.20), we obtain the inequality for $\widehat{\lambda}_n^{min}$.

$$\widehat{\lambda}_n^{min} \geq \frac{e^{-a\tau}\alpha^2\widehat{\lambda}_{n-1}^{min}}{1+\frac{\alpha^2}{\gamma^2}\widehat{\lambda}_{n-1}^{min}}.$$

We now consider the following discrete dynamical system of the eigenvalue

$$\lambda_{n+1} = g(\lambda_n), \quad \lambda_0 > 0,$$

where $g(\lambda) = \frac{e^{-a\tau}\alpha^2\lambda}{1+\frac{\alpha^2}{\gamma^2}\lambda}$. Note that $\lambda_n > 0$ for $n \in \mathbb{N}\cup\{0\}$. Let $\lambda_\infty = \frac{\gamma^2}{\alpha^2}(e^{-a\tau}\alpha^2 - 1)$ be a fixed point of the dynamical system, i.e., $\lambda_\infty = g(\lambda_\infty)$. Then, if $e^{-a\tau}\alpha^2 > 1$, the ratio $\frac{g(\lambda)}{\lambda}$ satisfies $\frac{g(\lambda)}{\lambda} \geq 1$ (resp. < 1) for $\lambda \leq \lambda_\infty$ (resp. $\lambda > \lambda_\infty$). Hence, we have $\lim_{n\to\infty}\lambda_n = \lambda_\infty$. On the other hand, $\lim_{n\to\infty}\lambda_n = 0$ if $e^{-a\tau}\alpha^2 \leq 1$. Therefore, we obtain the lower bound

$$\widehat{\lambda}_n^{min} \geq \min\left\{\widehat{\lambda}_0^{min}, \frac{\gamma^2}{\alpha^2}(e^{-a\tau}\alpha^2-1)\right\} = \min\left\{e^{-a\tau}\lambda_0, \frac{\gamma^2}{\alpha^2}(e^{-a\tau}\alpha^2-1)\right\} = \lambda_* > 0 \tag{7.21}$$

if and only if $e^{-a\tau}\alpha^2 > 1$.

Finally, we establish the recurrence inequality for $\mathbb{E}[|e_n|^2]$. Owing to Assumption 3.2, the equality (4.35) is reduced to

$$(I + \alpha^2\gamma^{-2}\widehat{P}_n)\overline{v}_n = \overline{\widehat{v}}_n + \alpha^2\gamma^{-2}\widehat{P}_n y_n.$$

As in the proof of Theorem 7.1, the error is divided into the two parts as $e_n = r_1 + r_2$, where

$$r_1 = (I_{\mathcal{H}} + \alpha^2\gamma^{-2}\widehat{P}_n)^{-1}\widehat{e}_n, \tag{7.22}$$
$$r_2 = (I_{\mathcal{H}} + \alpha^2\gamma^{-2}\widehat{P}_n)^{-1}\alpha^2\gamma^{-2}\widehat{P}_n(y_n - u_n), \tag{7.23}$$

and $\widehat{e}_n = \overline{\widehat{v}}_n - u_n$. From (2.10), we have $|(I_{\mathcal{H}} + \alpha^2\gamma^{-2}\widehat{P}_n)^{-1}\alpha^2\gamma^{-2}\widehat{P}_n|_{\mathcal{L}} \leq 1$. We thus obtain

$$|r_2| \leq |y_n - u_n| = |\eta_n|. \tag{7.24}$$

From this lower bound of the minimum eigenvalue (7.21), we have $|(I_{\mathcal{H}}+\alpha^2\gamma^{-2}\widehat{P}_n)^{-1}|_{\mathcal{L}} \leq (1+\frac{\alpha^2}{\gamma^2}\lambda_*)^{-1}$. Hence,

$$|r_1|^2 \leq \frac{1}{(1+\frac{\alpha^2}{\gamma^2}\lambda_*)^2}|\widehat{e}_n|^2 \leq \frac{e^{2(\beta+\epsilon)h}}{(1+\frac{\alpha^2}{\gamma^2}\lambda_*)^2}(|e_{n-1}|^2 + D) - \frac{D}{(1+\frac{\alpha^2}{\gamma^2}\lambda_*)^2}, \tag{7.25}$$

where Lemma 5.5 is used for $\epsilon > 0$.

As in the proof of Theorem 7.1, we compute the expectations of r_1 and r_2 separately

$$\begin{aligned}
\mathbb{E}[|e_n|^2] &= \mathbb{E}[|r_1|^2] + \mathbb{E}[|r_2|^2] \le \mathbb{E}\left[|r_1|^2\right] + \mathbb{E}[|\eta_n|^2] = \mathbb{E}\left[|r_1|^2\right] + N_u\gamma^2 \\
&\le \frac{e^{2(\beta+\epsilon)h}}{(1+\frac{\alpha^2}{\gamma^2}\lambda_*)^2}(\mathbb{E}[|e_{n-1}|^2] + D) - \frac{D}{(1+\frac{\alpha^2}{\gamma^2}\lambda_*)^2} + N_u\gamma^2 \\
&\le \theta(\mathbb{E}[|e_{n-1}|^2] + D) - \frac{D}{(1+\frac{\alpha^2}{\gamma^2}\lambda_*)^2} + N_u\gamma^2 = \theta(\mathbb{E}[|e_{n-1}|^2] + D) + \delta,
\end{aligned}$$

where $\delta = N_u\gamma^2 - \Theta D$. This leads to

$$\mathbb{E}[|e_n|^2] + D \le \theta(\mathbb{E}[|e_{n-1}|^2] + D) + \delta + D.$$

Applying this inequality repeatedly, we finally have

$$\begin{aligned}
\mathbb{E}[|e_n|^2] &\le \theta^n(\mathbb{E}[|e_0|^2] + D) + (\delta + D)\frac{1-\theta^n}{1-\theta} - D \\
&= \theta^n(\mathbb{E}[|e_0|^2] + D) + \left(N_u\gamma^2 + (1-\Theta)D\right)\frac{1-\theta^n}{1-\theta} - D \\
&= \theta^n(\mathbb{E}[|e_0|^2] + D) + N_u\gamma^2\frac{1-\theta^n}{1-\theta} + \left(\frac{(1-\theta^n)(1-\Theta)}{1-\theta} - 1\right)D.
\end{aligned}$$

Moreover, if $\theta < 1$, (7.16) holds in the limit of $n \to \infty$. □

Similarly, we obtain the error bound of the EAKF.

Theorem 7.4 (Accuracy of the EAKF with multiplicative inflation)**.** *Suppose that the same assumption as Theorem 7.3 for $\boldsymbol{V}_n$ generated by the EAKF in Definition 4.16 with the multiplicative inflation in Definition 4.20 for $\alpha \ge 1$. Then, the same error bound as (7.15) and (7.16) hold.*

Proof. The proof is same as that of Theorem 7.3 since it only uses the relationships in Remark 4.22:

$$\begin{aligned}
(I_{\mathcal{H}} + \alpha^2\widehat{P}_nH^*R^{-1}H)\overline{v}_n &= \overline{\widehat{v}}_n + \alpha^2\widehat{P}_nH^*R^{-1}y_n, \\
P_n &= \frac{\alpha^2}{m-1}d\widehat{\boldsymbol{V}}_n(I_m + \alpha^2 d\widehat{\boldsymbol{V}}_n^*H^*R^{-1}Hd\widehat{\boldsymbol{V}}_n)^{-1}d\widehat{\boldsymbol{V}}_n^*.
\end{aligned}$$

These are shared with the ETKF and EAKF. □

Luo and Hoteit [68] derive both the upper and lower bounds for the multiplicative inflation parameter α, and they ensure the residual error $e_n^r = y_n - H\overline{v}_n$ remains within a prescribed interval. The bound for α is adaptively computed using the prediction residual error $\widehat{e}_n^r = y_n - H\overline{\widehat{v}}_n$ before the analysis step at each time. They do not impose any assumptions on the model dynamics. However, their theory only guarantees the

bound for the residual error e_n^r, not the actual error $e_n = u_n - \overline{v}_n$. In contrast, the present results of the accuracy ensure the bound for the actual error e_n, provided that the constants ρ and β for the model dynamics are estimated.

Compared to the additive inflation in Proposition 6.8, the multiplicative inflation in Theorem 7.3 and Theorem 7.4 inflates the spread of the ensemble. As a result, we cannot obtain the error bound for all members in the analysis ensemble $\boldsymbol{V}_n$. Instead, we establish the error bound for the analysis mean $\overline{v}_n$ using Lemma 5.5, which introduces the additional constant D in (7.16). However, the additional term becomes negligible and the filtering error tend to be the order of the observation noise in a certain limit.

Corollary 7.3. *Under the same assumptions of Theorem 7.3, in the accurate observation limit (i.e., $r \to 0$), the filtering error (7.16) satisfies*

$$\lim_{n\to\infty} \mathbb{E}[|e_n|^2] = O(r^2).$$

Proof. It is clear that $\Theta = \left(\frac{r^2}{r^2+\alpha^2\lambda_*}\right)^2 = O(r^4)$ and $\theta = O(r^4)$. Then, we have

$$\begin{aligned}\frac{1-\Theta}{1-\theta} - 1 &= (1-\Theta)(1+\theta+O(\Theta^2)) - 1 = 1 - \Theta + \theta + O(\Theta^2) - 1 \\ &= \Theta(e^{2(\beta+\epsilon)h} - 1) + O(\Theta^2) = O(r^4).\end{aligned}$$

Therefore, it follows that

$$\lim_{n\to\infty} \mathbb{E}[|e_n|^2] = \frac{mr^2}{1-\theta} + \left(\frac{1-\Theta}{1-\theta} - 1\right) D = mr^2(1+O(r^4)) + O(r^4) = O(r^2).$$

□

7.3 Numerical examples

The Observing System Simulation Experiment (OSSE) or twin experiment is a standard process used to validate data assimilation algorithms numerically. It uses only synthetic data generated directly from the numerical model rather than real world observations to avoid issues related to the imperfect models and unknown measurement noises.

Definition 7.4 (OSSE)**.** *The standard process of the OSSE is as follows.*

(1) Compute the true solution $(u_n)_{n=1}^N$ to (6.2) numerically.

(2) Generate random observations $(y_n)_{n=1}^N$ as in (6.4).

(3) Assimilate the observed data $(y_n)_{n=1}^N$ and obtain the analysis states $(\varpi_n)_{n=1}^N$ using a data assimilation algorithm.

(4) Compute the error between u_n and ϖ_n for $n = 1, \dots, N$.

In the EnKF, the analysis state is the ensemble mean, i.e., $\varpi_n = \overline{v}_n$. We use the square error (SE) at time n,

$$\mathrm{SE}_n = |\varpi_n - u_n|^2. \tag{7.26}$$

This is used in the error bounds (7.15) and (7.16) of Theorem 7.3. The root mean square error (RMSE) at time n is also used

$$\mathrm{RMSE}_n = \frac{|\varpi_n - u_n|}{\sqrt{N_u}} = \sqrt{\frac{\mathrm{SE}_n}{N_u}}, \tag{7.27}$$

which is normalized by the dimension of the model dynamics. These errors contain the numerical error as well as the state estimation error since we use a numerical approximation of the true solution in the step (1) of Definition 7.4.

Let us consider the L96 equation (5.6) as an example of the OSSE. We set $J = 40$ and $f = 8$. With these parameters, the L96 equation can exhibit chaotic behavior since we have $\beta = 2\rho - 1 = 2\sqrt{2J}f - 1 > 0$ in (5.1) of Assumption 5.2 as stated in Proposition 5.9. We first compute the solution $u(t)$ up to $T = N\Delta t$ by using the fourth-order Runge-Kutta method with a time step size $\Delta t = 0.01$, and $N = 14400$ time steps. The initial condition is set as

$$\begin{aligned} u_0 &= (f * 1.001, f, \cdots, f)^* \in \mathbb{R}^J \\ &= (8.008, 8.0, \cdots, 8.0)^* \in \mathbb{R}^{40}. \end{aligned}$$

The evolution of the first component $u^1(t)$ of the solution $u(t)$ for $0 \le t \le 144$ is shown in Figure 5(a). The projection onto the first two components $(u^1(t), u^2(t)) \in \mathbb{R}^2$ is shown in Figure 5(b).

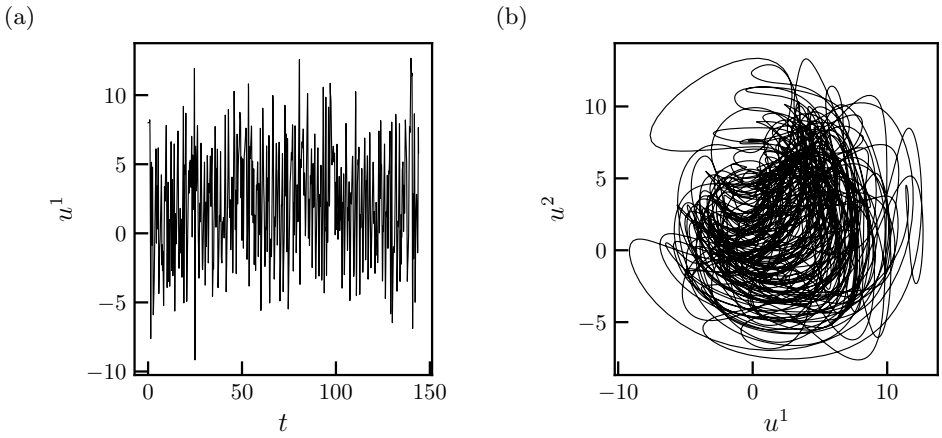

Figure 5: The solution of the L96 equation (5.6) for $(J, f) = (40, 8)$.

We show the time evolution of the norm $|u(t)|/\sqrt{J}$ in Figure 6 to confirm the boundedness of the solution. The norm quickly decays in the initial stage, which demonstrates the dissipativeness of the L96 equation. After that, the norm remains around the value of 4.0 and fluctuates aperiodically, indicating that the trajectory is bounded.

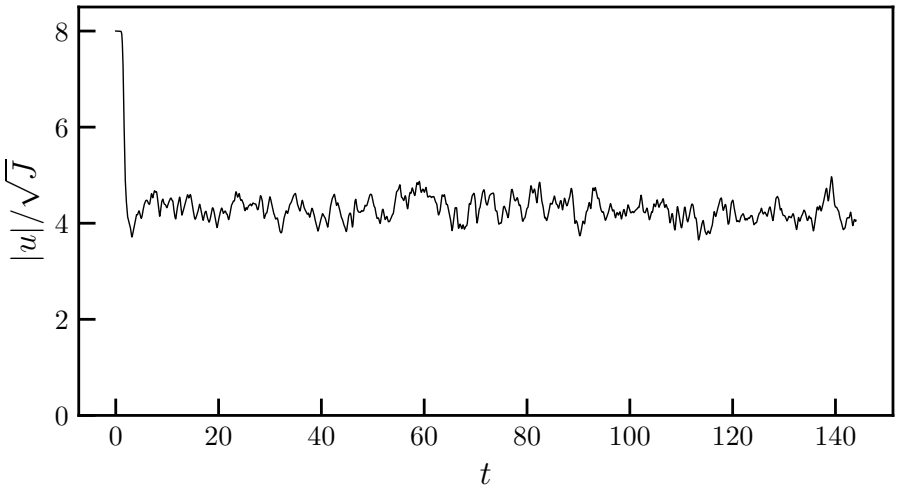

Figure 6: The time evolution of the norm $|u(t)|/\sqrt{J}$.

Now, let $u(t)$ and $u^\epsilon(t)$ be two solutions to the L96 equation starting from the initial states $u(0)$ and $u^\epsilon(0)$ with $|u^\epsilon(0) - u(0)|/\sqrt{J} \approx \epsilon$ for a small scale $\epsilon > 0$. We show in Figure 7 the evolution of the normalized error, $|\delta u(t)|/\sqrt{J}$, where $\delta u(t) = u(t) - u^\epsilon(t)$ and $\epsilon = 10^{-4}$. This indicates that the solution of the L96 equation exhibits sensitivity to the initial perturbation, showing a chaotic behavior.

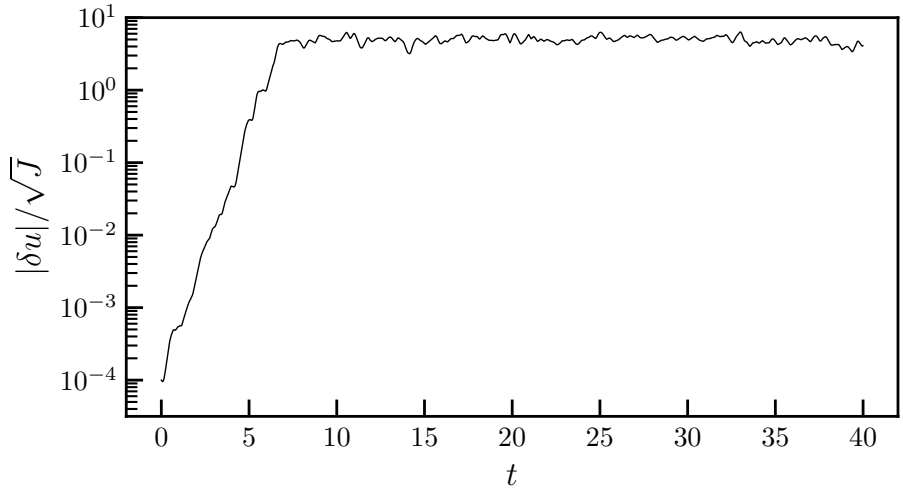

Figure 7: Log plot of the time evolution of the normalized error $|\delta u(t)|/\sqrt{J}$ for $\epsilon = 10^{-4}$.

We apply the ETKF to the L96 equation to validate the analysis in the previous section. For a time interval $\tau = 5\Delta t = 0.05$, we approximate the forward map Ψ_τ by the numerical solution $u(t)$ computed above. First, we discard the solution up to $t = N_0\tau$ with $N_0 = 1440$ so that it falls into an absorbing ball of the L96 equation. Then, we sample the hidden true states $(u_n)_{n=1}^{N_t}$ with $N_t = 480$ as follows:

$$u_n = u(n\tau + N_0\tau), \quad n = 1, \dots, N_t.$$

This process corresponds to the step (1) of the OSSE in Definition 7.4. Second, under Assumption 3.2 for the observation system (6.4), we generate a time series of observations $(y_n)_{n=1}^{N_t}$ as follows.

$$y_n = u_n + \eta_n, \ \eta_n \sim N(0, r^2 I), \quad n = 1, \dots, N_t,$$

where the variance of the observation noises is $r^2 = 0.1$. This process corresponds to the step (2) of the OSSE in Definition 7.4.

In step (3) of the OSSE in Definition 7.4, we apply the ETKF with an ensemble size $m = J + 1$. The initial ensemble is given by $\boldsymbol{V}_0 = [\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_J, -\sum_{i=1}^J \boldsymbol{e}_i] \in \mathbb{R}^{J\times m}$ for the standard basis $\boldsymbol{e}_i$ $(i = 1, \dots, J)$ of $\mathbb{R}^J$ so that the assumption $\lambda_{min}(C_0) > 0$ in Theorem 7.3 holds. From $\boldsymbol{V}_0$ and $(y_n)_{n=1}^{N_t}$, we compute the prediction ensemble $(\widehat{\boldsymbol{V}}_n)_{n=1}^{N_t}$ and the analysis ensemble $(\boldsymbol{V}_n)_{n=1}^{N_t}$ using the ETKF with the multiplicative covariance inflation in Definition 4.20 for $\alpha = 1.0, 1.1$, and 5.0.

The expectation $\mathbb{E}[\mathrm{SE}_n]$ is approximated by averaging over 20 different random seeds to generate observation noises. Figure 8(a) shows $\mathbb{E}[\mathrm{SE}_n]$ for $\alpha = 1.0, 1.1$, and 5.0. We also show the error bound (7.16) for the case when the error growth is sufficiently

suppressed by the inflation, i.e., $\theta \ll 1$. In this case, the bound is approximated by Jr^2. In addition, Figure 8(b) shows the minimum eigenvalue of the prediction covariance $\hat{\lambda}_{min} = \lambda_{min}(\widehat{P}_n)$ for $\alpha = 1.0, 1.1$, and 5.0. These figures indicate that without the inflation ($\alpha = 1.0$), $\mathbb{E}[\mathrm{SE}_n]$ is larger than the theoretical bound (Figure 8(a)), and the $\hat{\lambda}_{min} \approx 10^{-10}$ is negligibly small compared to the variance of observation noise $r^2 = 0.1$ (Figure 8(b)). On the other hand, a large inflation parameter ($\alpha = 5.0$) leads to that $\hat{\lambda}_{min}$ is bounded below by a value $\approx 10^{-2}$, which is about 8 digits larger in magnitude than that with $\alpha = 1.0$ (Figure 8(b)). Then, $\mathbb{E}[\mathrm{SE}_n]$ is smaller than or has the same order with the theoretical bound except at the initial time (Figure 8(a)). This result supports Theorem 7.3, considering the approximation error of the theoretical bound and numerical errors. With a relatively small inflation parameter ($\alpha = 1.1$), $\mathbb{E}[\mathrm{SE}_n]$ is even smaller than when $\alpha = 5.0$ although $\hat{\lambda}_{min}$ still takes a small value. Therefore, our estimate does not characterize the optimal α. Nevertheless, this numerical result is consistent with our theory since a large inflation parameter and a lower bound of the minimum eigenvalues are sufficient conditions to obtain the error bound (7.16).

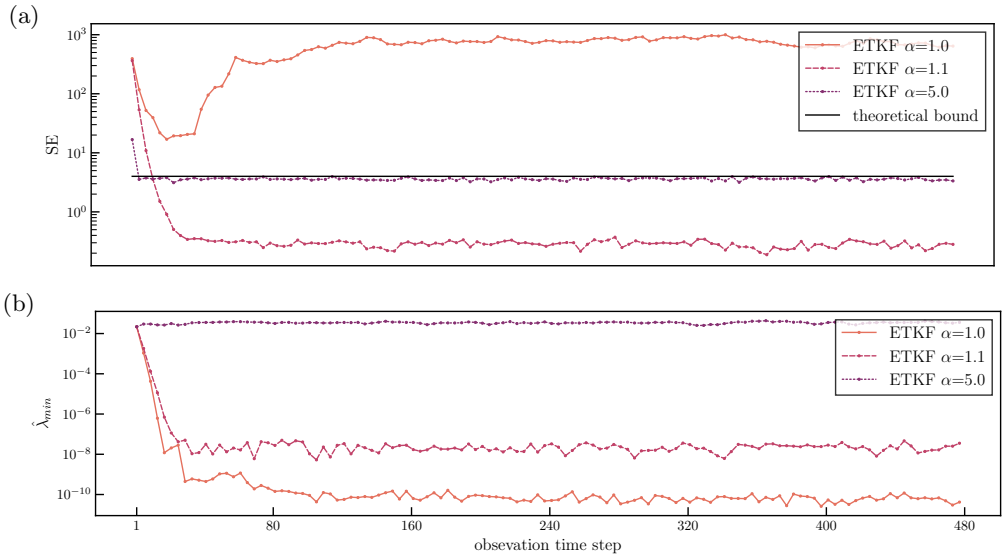

Figure 8: Plot of (a) the evolution $\mathbb{E}[\mathrm{SE}_n]$ and (b) Plot of the minimum eigenvalue $\hat{\lambda}_{min}$ vs. the observation time step n for the ETKF with the multiplicative inflation parameter $\alpha = 1.0, 1.1, 5.0$.

We finally investigate the dependence of the time-averaged SE on the variance of observation noise r^2. Figure 9 shows the log-log plot of the time averaged SE for $\alpha = 1.1, 5.0$ vs. r^2 with $r = 10^{-5}, \ldots, 10^{-1}$. Here, SE is computed using one seed for random observations for each parameter pair (r, α). For $\alpha = 5.0$, the theoretical estimate $\mathrm{SE} = O(r^2)$ in Corollary 7.3 is validated.

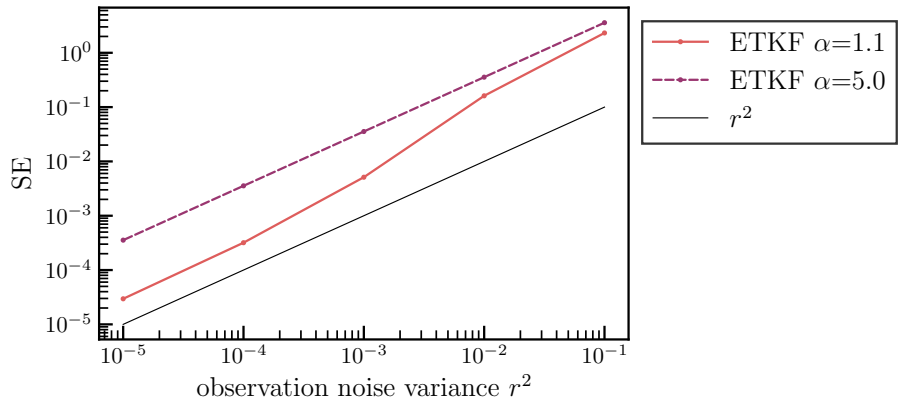

Figure 9: The log-log plot of the time averaged SE vs. r^2 for the ETKF with multiplicative inflation for $\alpha = 1.1, 5.0$ and $r = 10^{-5}, \ldots, 10^{-1}$.

8. Summary and future directions

The theoretical aspects of the EnKF have been investigated for dissipative dynamical systems on Hilbert spaces. Regarding the EnKF, the consistency with the KF, the exponential stability, and the error bound under ideal conditions have been established. In particular, the main contribution of this thesis is summarized in the following three points [82]. First, we prove that the filtering error of the ESRF is bounded for any finite time, even in an infinite-dimensional state space. Second, with the multiplicative inflation, we determine the minimum value of the inflation parameter α sufficient to obtain the uniform-in-time error bound when the state space has a finite dimension. These are the theoretical extensions of the analysis for the PO method to the case of the ESRF. Last, the numerical example validates the error bound and demonstrates that the bound for α may be improved.

We show future directions. First, the accuracy results for the ESRF with multiplicative inflation, Theorem 7.3 and Theorem 7.4 can be extended to the case for the PO method with the multiplicative ensemble inflation (1') in Definition 4.20. To this end, we need to estimate the eigenvalues of the ensemble covariance P_n randomly generated in the analysis step of the PO method. Second, the error analysis of the EnKF are limited in the case of full observations, which is an unrealistic setting in applications. We may extend the results for the 3DVar with partial observations in Table 1 to that for the EnKF. For instance, considering the L63 equation, we conjecture that the error bound might be obtained for the PO method with the additive inflation using the partial observation operator (5.14).

Third, we assume that the ensemble size is larger than the state space dimension, which is another theoretical limitation. In applications of the EnKF, it is usually unexpected due to the limited computational resources. The dimension reduction in dynamical systems is necessary to obtain the error bound with a small ensemble. For instance, the determining modes for the 2D-NSE on a torus define a finite-dimensional subspace, and partial observations from this subspace can reconstruct the state in the whole space. Therefore, we could obtain the error bound if an ensemble size is larger than the dimension of the subspace. Last, this thesis assumes that the exact time evolution of dynamical models can be approximated without numerical errors, as discussed in Section 6.2. Therefore, in the future, we should include numerical errors in the error analysis of data assimilation algorithms by utilizing the theory of numerical analysis. Relevant topics are addressed in [27] for the discretization of the Bayesian inverse problem formulated in infinite-dimensional spaces and [76] for the treatment of imperfect models.

Bibliography

[1] O. Al-Ghattas and D. Sanz-Alonso, *Non-asymptotic analysis of ensemble Kalman updates: Effective dimension and localization*, Information and Inference: A Journal of the IMA, 13 (2024), p. iaad043.

[2] D. A. F. Albanez, H. J. N. Lopes, and E. S. Titi, *Continuous data assimilation for the three-dimensional Navier-Stokes-alpha model*, Asymptot. Anal., 97 (2016), pp. 139–164.

[3] C. D. Aliprantis and K. C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer Berlin, Heidelberg, 3 ed., 2006. https://link.springer.com/book/10.1007/3-540-29587-9.

[4] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Courier Corporation, May 2012.

[5] J. L. Anderson, *An Ensemble Adjustment Kalman Filter for Data Assimilation*, Mon. Weather Rev., 129 (2001), pp. 2884–2903.

[6] ——, *An adaptive covariance inflation error correction algorithm for ensemble filters*, Tellus Dyn. Meteorol. Oceanogr., 59 (2007), pp. 210–224.

[7] ——, *Spatially and temporally varying adaptive covariance inflation for ensemble filters*, Tellus Dyn. Meteorol. Oceanogr., 61 (2009), pp. 72–83.

[8] J. L. Anderson and S. L. Anderson, *A Monte Carlo Implementation of the Nonlinear Filtering Problem to Produce Ensemble Assimilations and Forecasts*, Mon. Weather Rev., 127 (1999), pp. 2741–2758.

[9] R. Atar and O. Zeitouni, *Exponential stability for nonlinear filtering*, Annales de l'Institut Henri Poincare (B) Probability and Statistics, 33 (1997), pp. 697–725.

[10] A. Azouani, E. Olson, and E. S. Titi, *Continuous Data Assimilation Using General Interpolant Observables*, J Nonlinear Sci, 24 (2014), pp. 277–304.

[11] A. Azouani and E. S. Titi, *Feedback control of nonlinear dissipative systemsby finite determining parameters- A reaction-diffusion paradigm*, EECT, 3 (Wed Oct 01 00:00:00 UTC 2014), pp. 579–594.

[12] K. Bergemann and S. Reich, *An ensemble Kalman-Bucy filter for continuous data assimilation*, Meteorol. Z., 21 (2012), pp. 213–219.

[13] A. N. Bishop and P. Del Moral, *On the mathematical theory of ensemble (linear-Gaussian) Kalman–Bucy filtering*, Math. Control Signals Syst., 35 (2023), pp. 835–903.

[14] C. H. Bishop, B. J. Etherton, and S. J. Majumdar, *Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects*, Mon. Weather Rev., 129 (2001), pp. 420–436.

[15] D. Blömker, K. Law, A. M. Stuart, and K. C. Zygalakis, *Accuracy and stability of the continuous-time 3DVAR filter for the Navier–Stokes equation*, Nonlinearity, 26 (2013), p. 2193.

[16] C. E. A. Brett, K. F. Lam, K. J. H. Law, D. S. McCormick, M. R. Scott, and A. M. Stuart, *Stability of Filters for the Navier-Stokes Equation*, Oct. 2011.

[17] ——, *Accuracy and stability of filters for dissipative PDEs*, Physica D: Nonlinear Phenomena, 245 (2013), pp. 34–45.

[18] G. L. Browning, W. D. Henshaw, and H.-O. Kreiss, *A numerical investigation of the interaction between the large and small scales of the two-dimensional incompressible Navier–Stokes equations*, Res. Rep. -UR-98-1712 Los Alamos Natl. Lab., (1998).

[19] G. Burgers, P. J. van Leeuwen, and G. Evensen, *Analysis Scheme in the Ensemble Kalman Filter*, Mon. Weather Rev., 126 (1998), pp. 1719–1724.

[20] A. Carrassi, M. Bocquet, L. Bertino, and G. Evensen, *Data assimilation in the geosciences: An overview of methods, issues, and perspectives*, WIREs Clim. Change, 9 (2018), p. e535.

[21] J. Charney, M. Halem, and R. Jastrow, *Use of Incomplete Historical Data to Infer the Present State of the Atmosphere*, J. Atmospheric Sci., 26 (1969). https://journals.ametsoc.org/view/journals/atsc/26/5/1520-0469_1969_026_1160_uoihdt_2_0_co_2.xml.

[22] A. Cheskidov, D. D. Holm, E. Olson, and E. S. Titi, *On a Leray–α model of turbulence*, Proc. R. Soc. Math. Phys. Eng. Sci., 461 (2005), pp. 629–649.

[23] A. J. Chorin and J. E. Marsden, *A Mathematical Introduction to Fluid Mechanics*, vol. 4 of Texts in Applied Mathematics, Springer, New York, NY, 1993.

[24] C. K. Chui and G. Chen, *Kalman Filtering*, Springer International Publishing, Cham, 2017.

[25] J. B. Conway, *A Course in Functional Analysis*, vol. 96 of Graduate Texts in Mathematics, Springer, New York, NY, 2007.

[26] S. L. Cotter, M. Dashti, J. C. Robinson, and A. M. Stuart, *Bayesian inverse problems for functions and applications to fluid mechanics*, Inverse Problems, 25 (2009), p. 115008.

[27] S. L. Cotter, M. Dashti, and A. M. Stuart, *Approximation of Bayesian Inverse Problems for PDEs*, SIAM J. Numer. Anal., 48 (2010), pp. 322–345.

[28] R. Daley, *Atmospheric Data Analysis*, Cambridge University Press, Cambridge, UK, 1991.

[29] M. Dashti and A. M. Stuart, *The Bayesian Approach to Inverse Problems*, in Handbook of Uncertainty Quantification, R. Ghanem, D. Higdon, and H. Owhadi, eds., Springer International Publishing, Cham, 2017, pp. 311–428.

[30] J. de Wiljes, S. Reich, and W. Stannat, *Long-Time Stability and Accuracy of the Ensemble Kalman-Bucy Filter for Fully Observed Processes and Small Measurement Noise*, Siam J. Appl. Dyn. Syst., 17 (2018), pp. 1152–1181.

[31] J. de Wiljes and X. T. Tong, *Analysis of a localised nonlinear ensemble Kalman Bucy filter with complete and accurate observations*, Nonlinearity, 33 (2020), pp. 4752–4782.

[32] L. Dieci and T. Eirola, *On smooth decompositions of matrices*, Siam J. Matrix Anal. Appl., 20 (1999), pp. 800–819.

[33] G. Evensen, *Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics*, J. Geophys. Res. Oceans, 99 (1994), pp. 10143–10162.

[34] ——, *Data Assimilation: The Ensemble Kalman Filter*, Springer, Berlin, Heidelberg, 2009.

[35] C. Foias, D. D. Holm, and E. S. Titi, *The three dimensional viscous camassa-holm equations, and their relation to the navier-stokes equations and turbulence theory*, J. Dyn. Differ. Equ., 14 (2001).

[36] C. Foias, O. Manley, R. Rosa, and R. Temam, *Navier-Stokes Equations and Turbulence*, Encyclopedia of Mathematics and Its Applications, Cambridge University Press, Cambridge, 2001.

[37] A. L. Gibbs and F. E. Su, *On Choosing and Bounding Probability Metrics*, Int. Stat. Rev. Rev. Int. Stat., 70 (2002), pp. 419–435.

[38] Y. Giga and A. Novotný, *Handbook of Mathematical Analysis in Mechanics of Viscous Fluids*, Springer Cham, 1 ed., 2018.

[39] G. H. Golub and C. F. V. Loan, *Matrix Computations*, Johns Hopkins University Press, 2013.

[40] T. M. Hamill, J. S. Whitaker, and C. Snyder, *Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter*, Mon. Weather Rev., 129 (2001), pp. 2776–2790.

[41] K. Hayden, E. Olson, and E. S. Titi, *Discrete data assimilation in the Lorenz and 2D Navier-Stokes equations*, Phys. -Nonlinear Phenom., 240 (2011), pp. 1416–1425.

[42] W. D. Henshaw, H.-O. Kreiss, and J. Ystrom, *Numerical Experiments on the Interaction Between the Large- and Small-Scale Motions of the Navier-Stokes Equations*, Multiscale Model. Simul., 1 (2003), pp. 119–149.

[43] M. Hladnik and MATJA. OMLADIC, *Spectrum of the Product of Operators*, Proc. Amer. Math. Soc., 102 (1988), pp. 300–302.

[44] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.

[45] P. L. Houtekamer and F. Zhang, *Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation*, Mon. Weather Rev., 144 (2016), pp. 4489–4532.

[46] B. R. Hunt, E. J. Kostelich, and I. Szunyogh, *Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter*, Physica D: Nonlinear Phenomena, 230 (2007), pp. 112–126.

[47] M. A. Iglesias, K. J. H. Law, and A. M. Stuart, *Ensemble Kalman methods for inverse problems*, Inverse Problems, 29 (2013), p. 045001.

[48] D. A. Jones and E. S. Titi, *Upper Bounds on the Number of Determining Modes, Nodes, and Volume Elements for the Navier-Stokes Equations*, Indiana Univ. Math. J., 42 (1993), pp. 875–887. https://www.jstor.org/stable/24897124.

[49] R. E. Kalman, *A New Approach to Linear Filtering and Prediction Problems*, J. Basic Eng, 82 (1960), pp. 35–45.

[50] E. Kalnay, *Atmospheric Modeling, Data Assimilation and Predictability*, Nov. 2002.

[51] I. Kasanicky, *Ensemble Kalman Filter on High and Infinite Dimensional Spaces*, PhD thesis, Charles University, Prague, Dec. 2016.

[52] T. Kato, *Perturbation Theory for Linear Operators*, vol. 132 of Classics in Mathematics, Springer, Berlin, Heidelberg, 1995.

[53] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Encyclopedia of Mathematics and Its Applications, Cambridge University Press, Cambridge, 1995.

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

[55] P. Korn, *Data assimilation for the Navier–Stokes-{α} equations*, Physica D: Nonlinear Phenomena, 238 (2009), pp. 1957–1974.

[56] ——, *On degrees of freedom of certain conservative turbulence models for the Navier–Stokes equations*, Journal of Mathematical Analysis and Applications, 378 (2011), pp. 49–63.

[57] S. Kotsuki, Y. Ota, and T. Miyoshi, *Adaptive covariance relaxation methods for ensemble data assimilation: Experiments in the real atmosphere*, Q. J. R. Meteorol. Soc., 143 (2017), pp. 2001–2015.

[58] E. Kwiatkowski and J. Mandel, *Convergence of the Square Root Ensemble Kalman Filter in the Large Ensemble Limit*, Siam-Asa J. Uncertain. Quantif., 3 (2015), pp. 1–17.

[59] W. Lamb, *Robinson, J. C.Infinite-dimensional dynamical systems (Cambridge University Press, 2001) 461pp., 0 521 63564 0 (paperback), £24.95, 0 521 63204 8 (hardback), £70*, Proc. Edinb. Math. Soc., 46 (2003), pp. 252–253.

[60] T. Lange and W. Stannat, *Mean field limit of Ensemble Square Root filters - discrete and continuous time*, FoDS, 3 (Thu Jan 21 19:00:00 EST 2021), pp. 563–588.

[61] K. Law, A. Shukla, and A. Stuart, *ANALYSIS OF THE 3DVAR FILTER FOR THE PARTIALLY OBSERVED LORENZ'63 MODEL*, Discrete Contin. Dyn. Syst., 34 (2014), pp. 1061–1078.

[62] K. J. H. Law, D. Sanz-Alonso, A. Shukla, and A. M. Stuart, *Filter accuracy for the Lorenz 96 model: Fixed versus adaptive observation operators*, Phys. -Nonlinear Phenom., 325 (2016), pp. 1–13.

[63] K. J. H. Law, A. M. Stuart, and K. C. Zygalakis, *Data Assimilation: A Mathematical Introduction*, Springer, 2015.

[64] W. Layton, C. C. Manica, M. Neda, and L. G. Rebholz, *Numerical analysis and computational comparisons of the NS-alpha and NS-omega regularizations*, Computer Methods in Applied Mechanics and Engineering, 199 (2010), pp. 916–931.

[65] E. N. Lorenz, *Deterministic Nonperiodic Flow*, J. Atmospheric Sci., 20 (1963), pp. 130–141. https://journals.ametsoc.org/view/journals/atsc/20/2/1520-0469_1963_020_0130_dnf_2_0_co_2.xml.

[66] ——, *Predictability: A problem partly solved*, Proc Semin. Predict. Read. UK ECMWF 1996, 1 (1996), pp. 1–18.

[67] E. N. Lorenz and K. A. Emanuel, *Optimal Sites for Supplementary Weather Observations: Simulation with a Small Model*, J. Atmospheric Sci., 55 (1998), pp. 399–414.

[68] X. Luo and I. Hoteit, *Covariance Inflation in the Ensemble Kalman Filter: A Residual Nudging Perspective and Some Implications*, Mon. Weather Rev., 141 (2013), pp. 3360–3368.

[69] A. J. Majda and J. Harlim, *Filtering Complex Turbulent Systems*, Filter. Complex Turbul. Syst., (2012), pp. 1–357.

[70] J. Mandel, L. Cobb, and J. D. Beezley, *On the convergence of the ensemble Kalman filter*, Appl. Math., 56 (2011), pp. 533–541.

[71] J. E. Marsden and S. Shkoller, *Global well–posedness for the Lagrangian averaged Navier–Stokes (LANS–α) equations on bounded domains*, Philos. Trans. R. Soc. Lond. Ser. Math. Phys. Eng. Sci., 359 (2001), pp. 1449–1468.

[72] T. Miyoshi, *The Gaussian Approach to Adaptive Covariance Inflation and Its Implementation with the Local Ensemble Transform Kalman Filter*, Mon. Weather Rev., 139 (2011), pp. 1519–1535.

[73] B. Øksendal, *Stochastic Differential Equations*, Universitext, Springer, Berlin, Heidelberg, 2003.

[74] E. Olson and E. S. Titi, *Determining Modes for Continuous Data Assimilation in 2D Turbulence*, Journal of Statistical Physics, 113 (2003), pp. 799–840.

[75] K. B. Petersen and M. S. Pedersen, *The Matrix Cookbook*. http://www2.compute.dtu.dk/pubdb/pubs/3274-full.html, Nov. 2012.

[76] S. Reich and C. Cotter, *Probabilistic Forecasting and Bayesian Data Assimilation*, Cambridge University Press, Cambridge, 2015.

[77] F. Rellich, *Perturbation theory of eigenvalue problems*. https://archive.org/details/perturbationtheo00rell.

[78] D. Sanz-Alonso, A. Stuart, and A. Taeb, *Inverse Problems and Data Assimilation*, London Mathematical Society Student Texts, Cambridge University Press, Cambridge, 2023.

[79] A. M. Stuart, *Inverse problems: A Bayesian perspective*, Acta Numer., 19 (2010), pp. 451–559.

[80] T. Sullivan, *Introduction to Uncertainty Quantification*, vol. 63 of Texts in Applied Mathematics, Springer International Publishing, Cham, 2015.

[81] K. Takeda, *Lorenz System*. https://kotatakeda.github.io/lorenz-webgl/, 2024.

[82] K. Takeda and T. Sakajo, *Uniform error bounds of the ensemble transform Kalman filter for chaotic dynamics with multiplicative covariance inflation*, SIAMASA J. Uncertain. Quantif., (2024, in press).

[83] T.-J. Tarn and Y. Rasis, *Observers for nonlinear stochastic systems*, IEEE Trans. Autom. Control, 21 (1976), pp. 441–448.

[84] R. Temam, *Navier-Stokes Equations and Nonlinear Functional Analysis*, CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics, Jan. 1995.

[85] ——, *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, Springer New York, NY, 1997.

[86] M. K. Tippett, J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, *Ensemble Square Root Filters*, Mon. Weather Rev., 131 (2003). https://journals.ametsoc.org/view/journals/mwre/131/7/1520-0493_2003_131_1485_esrf_2.0.co_2.xml.

[87] X. T. Tong, *Performance Analysis of Local Ensemble Kalman Filter*, J Nonlinear Sci, 28 (2018), pp. 1397–1442.

[88] X. T. Tong, A. J. Majda, and D. Kelly, *Nonlinear stability and ergodicity of ensemble based Kalman filters*, Nonlinearity, 29 (2016), pp. 657–691.

[89] X. T. Tong, A. J. Majda, and D. Kelly, *Nonlinear stability of the ensemble Kalman filter with adaptive covariance inflation*, Comm. Math. Sci., 14 (2016), pp. 1283–1313.

[90] X. T. Tong and M. Morzfeld, *Localized ensemble Kalman inversion*, Inverse Problems, 39 (2023), p. 064002.

[91] J. S. Whitaker and T. M. Hamill, *Ensemble Data Assimilation without Perturbed Observations*, Mon. Weather Rev., 130 (2002), pp. 1913–1924.

[92] ——, *Evaluating Methods to Account for System Errors in Ensemble Data Assimilation*, Mon. Weather Rev., 140 (2012), pp. 3078–3089.

[93] Y. Ying and F. Zhang, *An adaptive covariance relaxation method for ensemble data assimilation*, Q. J. R. Meteorol. Soc., 141 (2015), pp. 2898–2906.

[94] F. Zhang, C. Snyder, and J. Sun, *Impacts of Initial Estimate and Observation Availability on Convective-Scale Data Assimilation with an Ensemble Kalman Filter*, Mon. Weather Rev., 132 (2004), pp. 1238–1253.