NONLINEAR STABILITY OF THE ENSEMBLE KALMAN FILTER WITH ADAPTIVE COVARIANCE INFLATION*

XIN T. TONG†, ANDREW J. MAJDA‡, AND DAVID KELLY§

Abstract. The ensemble Kalman filter and ensemble square root filters are data assimilation methods used to combine high dimensional nonlinear models with observed data. These methods have proved to be indispensable tools in science and engineering as they allow computationally cheap, low dimensional ensemble state approximation for extremely high dimensional turbulent forecast models. From a theoretical perspective, these methods are poorly understood, with the exception of a recently established but still incomplete nonlinear stability theory. Moreover, recent numerical and theoretical studies of catastrophic filter divergence have indicated that stability is a genuine mathematical concern and can not be taken for granted in implementation. In this article we propose a simple modification of ensemble based methods which resolves these stability issues entirely. The method involves a new type of adaptive covariance inflation, which comes with minimal additional cost. We develop a complete nonlinear stability theory for the adaptive method, yielding Lyapunov functions and geometric ergodicity under weak assumptions. We present numerical evidence which suggests the adaptive methods have improved accuracy over standard methods and completely eliminate catastrophic filter divergence. This enhanced stability allows for the use of extremely cheap, unstable forecast integrators, which would otherwise lead to widespread filter malfunction.

Key words. Data assimilation, ensemble Kalman filter, nonlinear stability, ergodicity of Markov chains.

AMS subject classifications. 93E11, 60M20, 62L12, 65C20, 37A25, 60G25.

1. Introduction

With the growing importance of accurate weather forecasting and expanding availability of geophysical measurements, data assimilation for high dimensional dynamical system and data has never been more crucial. The ensemble Kalman filter (EnKF) [1] and ensemble square root filters (ESRF) [2,3] are ensemble based algorithms well designed for this purpose. They quantify the uncertainty of an underlying system using the sample information of a moderate size ensemble $\{V_n^{(k)}\}_{k=1}^K$, thereby significantly reducing the computational cost. The simplicity of these algorithms and their accurate performance has fueled their wide application in various fields of geophysical science [4,5].

Despite their success, there are many unresolved issues with ensemble based methods. First, there is very little theoretical understanding of the methods and notably the stability framework is incomplete. Only very recently has the stability theory been partially understood in the finite ensemble size scenario, with progress on well-posedness in the fully observed case [6] and subsequently in nonlinear stability of the partially observed case, but under the so-called observable energy criterion [7]. The shadowing properties of EnKF have been analyzed in [8], under idealized assumptions on the model.

†Centre for Atmosphere Ocean Science, New York University, New York, NY, 10012 and Courant Institute of Mathematical Sciences, New York University, New York, NY, 10012, USA (tong@cims.nyu.edu).

‡Centre for Atmosphere Ocean Science, New York University, New York, NY, 10012 and Courant Institute of Mathematical Sciences, New York University, New York, NY, 10012, USA (jonjon@cims.nyu.edu).

§Corresponding author, Courant Institute of Mathematical Sciences, New York University, New York, NY, 10012, USA (dtkelly@cims.nyu.edu).

A better understanding of filter stability is sorely lacking, with recent numerical studies [9,10] revealing the mechanistically mysterious phenomenon known as *catstrophic filter divergence*, whereby filter state estimates tend to machine infinity, whilst the underlying signal remains in a bounded set. In [11] it has been established rigorously, through an elementary model, that this divergence is not caused by the instability of the numerical integrator alone, instead the update step in the filter itself plays a crucial role in the genesis of divergence.

In this article we propose a simple modification to EnKF (and ESRF) which resolves these issues completely. The modification is a type of covariance inflation, a widely used strategy for stabilizing and improving the accuracy of filters. Typically the forecast covariance $\widehat{C}_n$ is (in the case of EnKF) additively inflated to $\widehat{C}_n + \lambda_n I$ for some choice of inflation constant λ_n. Since the forecast covariance decides how much uncertainty is held by the forecast prediction, inflation has the affect of pulling the filter back towards the observations, yielding improved stability. Existing methods of covariance inflation, such as constant additive inflation (λ_n is constant) tend to improve accuracy, but are still vulnerable to stability issue like catastrophic filter divergence [9].

The modification we propose selects the inflation strength λ_n *adaptively* and varies according to the distribution of the ensemble. In particular, if the filter is deemed to be performing well, the inflation strength is set to zero and the method is reduced to EnKF (or ESRF, as desired). If the filter is deemed to be 'malfunctioning', then the adaptive inflation is *triggered*. The strength of the inflation becomes larger when the filter strays further into malfunction. To decide when and to what extent the filter is malfunctioning, we employ two simple statistics of the ensemble, Θ_n and Ξ_n, which are based respectively on the ensemble innovation and the cross correlation between observed and unobserved components. The two statistics are so chosen as it is clear from the theoretical framework that these are precisely the two variables which must be controlled to guarantee stability. Nevertheless, there is a quite natural interpretation as to why these two statistics are an effective gauge of filter performance. The full derivation and explanation of the adaptively inflated methods are given in Section 3.

In sections 4 and 5, we develop a complete stability theory for the adaptively inflated methods by extending the stability framework established in [7]. This framework is comprised of two main results: time uniform mean-square estimates on the ensemble via a Lyapunov argument and geometric ergodicity of the signal-ensemble process. We prove that if the underlying model satisfies energy dissipation, then the filter inherits this dissipation and in particular has a Lyapnuov function with compact sub-level sets. This is a vast improvement on the results in [7] for EnKF, since firstly the observable energy criterion is no longer required and secondly the Lyapunov function for the filter is guaranteed to have compact sub-level sets. This latter fact leads to geometric ergodicity of the signal-ensemble process for the adaptively inflated EnKF, which follows as an immediate corollary of the results in [7].

In Section 7, we investigate the performance of the adaptively inflated filter numerically, comparing performance with standard methods, such as non-inflated EnKF and non-adaptively (such as constantly) inflated EnKF. The adaptive method performs at least as well and typically better than standard methods and as expected completely avoids the issue of catastrophic filter divergence , which is quite prevalent in all standard methods. The additional expense in computing the inflation strength is minimal and hence the adaptive methods run at almost identical speed to the standard methods. Most impressively, the adaptive method allows for the use of extremely cheap but unstable integrators, like explicit Euler. We will see that such integrators are useless for

standard methods due to the prevalence of catastrophic filter divergence, but are very successful when used with adaptive methods.

Regarding existing literature, there have recently been many methods proposed which employ a type of adaptive inflation for ensemble based methods [12–16], but our method is truly novel. Moreover, none of the cited methods have an established theoretical understanding, and our framework may serve as a good starting point for developing such an understanding.

The structure of the article is as follows. In Section 2 we define the standard ensemble based methods, EnKF and ESRFs. In Section 3 we define the adaptive inflated modifications of EnKF and ESRFs. In Section 4 we derive time uniform mean-square estimates for the adaptively inflated filters via a Lyapunov function argument. In Section 5 we prove geometric ergodicity of the signal-ensemble process for the adaptively inflated filters. In Section 7 we investigate the performance of these filters numerically. In Section 8 we conclude with a discussion. The Appendix contains technical tools that will be used for the main results.

2. Ensemble based methods and adaptive inflation

2.1. Model setup.   In this paper, we assume the signal (truth) sequence $U_n \in \mathbb{R}^d$ is generated by

$$U_n = \Psi_h(U_{n-1}) + \zeta_n, \tag{2.1}$$

where $\Psi_h : \mathbb{R}^d \to \mathbb{R}^d$ is a deterministic mapping and $\{\zeta_n\}_{n \geq 1}$ is the system noise. The noise ζ_n is assumed to be independent of $\zeta_1, \ldots, \zeta_{n-1}$ when conditioned on the realization of U_{n-1}, the conditional mean is zero $\mathbb{E}(\zeta_n | U_{n-1}) = 0$ and the conditional covariance $\mathbb{E}(\zeta_n \otimes \zeta_n | U_{n-1})$ is denoted by $R_h(U_{n-1})$. Due to the conditional independence of the noise sequence, this gives rise to a Markov chain.

For example, the model is often generated by the solution of a stochastic differential equation (SDE)

$$du_t = \psi(u_t)dt + \Sigma dW_t, \tag{2.2}$$

for a sufficiently regular vector field $\psi : \mathbb{R}^d \to \mathbb{R}^d$, diffusion coefficient $\Sigma \in \mathbb{R}^{d \times e}$ and e-dimensional Wiener process W. We then take $U_n = u_{nh}$ for some fixed $h > 0$. In the notation above, we have $\Psi_h(u_0) = \mathbb{E}(u_h | u_0)$ and $\zeta_n = u_{nh} - \Psi_h(u_{(n-1)h})$. It is easy to see that this satisfies the above conditions. We will refer to this as the *discrete time formulation* of an SDE, not to be confused with a time discretization of an SDE.

2.2. Linear observations.   We assume that the truth U_n is observed linearly with a mean zero Gaussian perturbation

$$Z_n = HU_n + \xi_n.$$

where the observation matrix H is of size $q \times d$ with $H = (H_0, 0_{q \times (d-q)})$ where $H_0 = \mathrm{diag}(h_1, \ldots, h_q)$, $q \leq d$ is the rank of H and $h_1, \ldots, h_q > 0$ are fixed scalars. For the observational noise, we assume that $\mathbb{E}(\xi_n | U_{n-1}) = 0$ and $\mathbb{E}(\xi_n \otimes \xi_n | U_{n-1}) = I_q$.

This seemingly restrictive setting can be assumed without loss of generality. Indeed, any observational noise covariance can be reduced to the identity via a simple rotation on the filtering problem. Suppose that ξ_n has a nonsingular covariance matrix Γ (we do not consider the singular case in this article) and $\Gamma^{-1/2}H$ has an SVD decomposition $\Gamma^{-1/2}H = \Phi \Lambda \Psi^T$, then we rotate the coordinate system and consider

$$\widetilde{U}_n = \Psi^T U_n, \quad \tilde{\xi}_n = \Phi^T \Gamma^{-1/2} \xi_n, \quad \widetilde{Z}_n = \Phi^T \Gamma^{-1/2} Z_n = \Lambda \widetilde{U}_n + \tilde{\xi}_n. \tag{2.3}$$

Hence this change of coordinates reduces the observation matrix and the observational noise covariance to the desired form.

If the observation dimension q is larger than the model dimension d, the last $d-q$ diagonal entries of Λ are zero, so the last $d-q$ rows of $\widetilde{Z}_n$ are independent of the signal and play no role in filtering, hence we can ignore and set $d=q$.

Since all the transformations above are linear and bijective, filtering $\widetilde{U}_n$ with $\widetilde{Z}_n$ is equivalent to filtering U_n with Z_n, in the sense that the subsequent assumptions and results hold equally for the original and transformed system. When necessary, this will be clarified within the text.

2.3. Ensemble Kalman filter. In the standard Kalman filtering theory, the conditional distribution of the signal process U_n given the observation sequence $Z_1,\dots,Z_n$ is given by a Gaussian distribution. EnKF inherits this idea by using a group of ensembles $\{V_n^{(k)}\}_{k=1}^{K}$ to represent this Gaussian distribution, as the mean and covariance can be taken as the ensemble mean and covariance. The EnKF operates very much like a Kalman filter, except its forecast step requires a Monte Carlo simulation due to the nonlinearity of the system. In detail, the EnKF is an iteration of following two steps, with (for instance) $\widehat{V}_0^{(k)}$ being sampled from the equilibrium measure of U_n.

- Forecast step: from the posterior ensemble at time $n-1$, $\{V_{n-1}^{(k)}\}_{k=1}^{K}$, a forecast ensemble for time n is generated by

$$\widehat{V}_n^{(k)} = \Psi_h(V_{n-1}^{(k)}) + \zeta_n^{(k)}, \tag{2.4}$$

where $\zeta_n^{(k)}$ are independent samples drawn from the same distribution as ζ_n.

- Analysis step: upon receiving the new observation Z_n, random perturbations of it are generated by adding $\xi_n^{(k)}$:

$$Z_n^{(k)} = Z_n + \xi_n^{(k)},$$

where $\xi_n^{(k)}$ are independent samples drawn from the same distribution as ξ_n. Using the Kalman update rule, each ensemble member is then updated as follow with $\widehat{C}_n$ being the sample covariance of the forecast ensemble:

$$\begin{aligned}
V_n^{(k)} &= \widehat{V}_n^{(k)} - \widehat{C}_n H^T (I + H\widehat{C}_n H^T)^{-1}(H\widehat{V}_n^{(k)} - Z_n^{(k)}) \\
&= (I + \widehat{C}_n H^T H)^{-1}\widehat{V}_n^{(k)} + (I + \widehat{C}_n H^T H)^{-1}\widehat{C}_n H^T Z_n^{(k)},
\end{aligned} \tag{2.5}$$

where

$$\widehat{C}_n = \frac{1}{K-1}\sum_{k=1}^{K}(\widehat{V}_n^{(k)} - \overline{\widehat{V}}_n) \otimes (\widehat{V}_n^{(k)} - \overline{\widehat{V}}_n), \quad \overline{\widehat{V}}_n = \frac{1}{K}\sum_{k=1}^{K}\widehat{V}_n^{(k)}. \tag{2.6}$$

See [6] for a derivation of EnKF as an approximate Monte Carlo method for sampling the true posterior distribution. In the case of linear model dynamics, it has been rigorously shown [17,18] that EnKF is an approximate Monte Carlo method for the true posterior distribution, with convergence in the $K \to \infty$ limit. In the nonlinear case, it has been shown in [19] that EnKF does not recover the true posterior, but rather a specific linear Bayes estimator of the posterior.

Based on our description above, the augmented process $\{U_n, V_n^{(1)}, \dots, V_n^{(K)}\}$ is a Markov chain. As above, we will denote the natural filtration up to time n as $\mathcal{F}_n =$

$\sigma\{U_m, V_m^{(1)}, \ldots, V_m^{(K)}, m \le n\}$, and denote the conditional expectation with respect to $\mathcal{F}_n$ as $\mathbb{E}_n$.

In the sequel it will be convenient to split the ensemble into observed and unobserved components. In particular, for each ensemble member we write $V_n^{(k)} = (X_n^{(k)}, Y_n^{(k)})$ where $X_n^{(k)} \in \mathbb{R}^q$ and $Y_n^{(k)} \in \ker(H)$. Similarly, we write $\widehat{V}_n^{(k)} = (\widehat{X}_n^{(k)}, \widehat{Y}_n^{(k)})$ and for the forecast ensemble covariance

$$\widehat{C}_n = \begin{bmatrix} \widehat{C}_n^X, & \widehat{B}_n \\ \widehat{B}_n^T, & \widehat{C}_n^Y \end{bmatrix},$$

where

$$\widehat{C}_n^X = \frac{1}{K-1} \sum_{k=1}^K (\widehat{X}_n^{(k)} - \overline{\widehat{X}}_n) \otimes (\widehat{X}_n^{(k)} - \overline{\widehat{X}}_n) \quad \overline{\widehat{X}}_n = \frac{1}{K} \sum_{k=1}^K \widehat{X}_n^{(k)}$$

and so forth the other components. In this notation, the update rule (2.5) becomes

$$X_n^{(k)} = \widehat{X}_n^{(k)} - (\widehat{C}_n^X)^T H_0^T (I + H_0 \widehat{C}_n^X H_0^T)^{-1} (H_0 \widehat{X}_n^{(k)} - Z_n^{(k)}) \tag{2.7}$$

for the observed components and

$$Y_n^{(k)} = \widehat{Y}_n^{(k)} - (B_n)^T H_0^T (I + H_0 \widehat{C}_n^X H_0^T)^{-1} (H_0 \widehat{X}_n^{(k)} - Z_n^{(k)}) \tag{2.8}$$

for the unobserved components.

2.4. Ensemble square root filters. One drawback of EnKF comes from its usage of artificial noise $\xi_n^{(k)}$, as this introduces unnecessary sampling errors, particularly when the ensemble size is small [20]. The motivation behind the artificial noise is to make the posterior ensemble covariance

$$C_n := \frac{1}{K-1} \sum_{k=1}^K (V_n^{(k)} - \overline{V}_n) \otimes (V_n^{(k)} - \overline{V}_n), \quad \overline{V}_n := \frac{1}{K} \sum_{k=1}^K V_n^{(k)},$$

satisfy the covariance update of the standard Kalman filter

$$C_n = \widehat{C}_n - \widehat{C}_n H^T (H^T \widehat{C}_n H + I)^{-1} H \widehat{C}_n, \tag{2.9}$$

when the left-hand side is averaged over $\xi_n^{(k)}$ [21–23]. ESRFs, including the ensemble transform Kalman filter (ETKF) and the ensemble adjustment Kalman filter (EAKF), aim to resolve this issue by manipulating the posterior spreads to ensure that (2.9) holds. Both ETKF and EAKF algorithms are described by the following update steps, with the only difference occurring in the assimilation step for the spread. As with EnKF, the initial ensemble $\{V_0^{(k)}\}_{k=1}^K$ is (for instance) sampled from the equilibrium distribution of U_n.

- Forecast step: identical to EnKF, the forecast ensembles at time n is generated from posterior ensembles at time $n-1$:

$$\widehat{V}_n^{(k)} = \Psi_h(V_{n-1}^{(k)}) + \zeta_n^{(k)}.$$

The forecast ensemble covariance $\widehat{C}_n$ is then computed using (2.6).

- Assimilation step for the mean: upon receiving the new observation Z_n, the posterior ensemble mean is updated through

$$\overline{V}_n = \overline{\widehat{V}_n} - \widehat{C}_n H^T (I + H\widehat{C}_n H^T)^{-1}(H\overline{\widehat{V}}_n - Z_n), \quad \overline{\widehat{V}}_n = \frac{1}{K}\sum_{k=1}^{K}\widehat{V}_n^{(k)}. \quad (2.10)$$

- Assimilation step for the spread: The forecast ensemble spread is given by the $d \times K$ matrix

$$\widehat{S}_n = [\widehat{V}_n^{(1)} - \overline{\widehat{V}}_n, \dots, \widehat{V}_n^{(K)} - \overline{\widehat{V}}_n].$$

To update the posterior spread, first find a matrix $T_n \in \mathbb{R}^{d \times d}$ (for ETKF) or $A_n \in \mathbb{R}^{K \times K}$ (for EAKF) such that

$$\frac{1}{K-1}T_n\widehat{S}_n \otimes T_n\widehat{S}_n = \frac{1}{K-1}\widehat{S}_n A_n \otimes \widehat{S}_n A_n = \widehat{C}_n - \widehat{C}_n H^T(H^T\widehat{C}_n H + I)^{-1}H\widehat{C}_n.$$
$$(2.11)$$

The posterior spread is updated to $S_n = T_n\widehat{S}_n$ (for ETKF) or $S_n = \widehat{S}_n A_n$ (EAKF), and the ensemble members are updated to

$$V_n^{(k)} = \overline{V}_n + S_n^{(k)},$$

where $S_n^{(k)}$ denotes the kth column of the updated spread matrix S_n. By construction, the posterior covariance $C_n = (K-1)^{-1}S_n^T S_n$ satisfies (2.9).

At this stage it suffices to know that such A_n and T_n exist, their finer properties play no role in the discussion concerning stability, but we refer the reader to [4, 7].

Based on our description above, the augmented process $\{U_n, V_n^{(1)}, \dots, V_n^{(K)}\}$ is again a Markov chain. As in the previous section, we employ the notation $\mathbb{E}_n$ to denote conditional expectation with respect to $\mathcal{F}_n$.

3. Adaptive inflation methods

In this section we introduce the adaptive inflation modification of EnKF, as well as the square root filters ETKF and EAKF. We start with the modification of EnKF, which we refer to *EnKF-AI* with the suffix standing for 'adaptive inflation'. The nomenclature for the algorithms and their modifications is summarized at the end of the section, in Table 3.1.

The adaptive inflation algorithm is precisely the EnKF algorithm with the forecast covariance $\widehat{C}_n$ replaced by an inflated forecast covariance $\widetilde{C}_n = \widehat{C}_n + \lambda_n I$. More precisely, the filter ensemble $\{V_n^{(k)}\}_{k \le K}$ is governed by

$$V_n^{(k)} = \widehat{V}_n^{(k)} - \widetilde{C}_n H^T(I + H\widetilde{C}_n H^T)^{-1}(H\widehat{V}_n^{(k)} - Z_n^{(k)})$$

$$\widehat{V}_n^{(k)} = \Psi_h(V_{n-1}^{(k)}) + \zeta_n^{(k)}, \quad \overline{\widehat{V}}_n = \frac{1}{K}\sum_{k=1}^{K}\widehat{V}_n^{(k)}, \quad Z_{n+1}^{(k)} = Z_n + \xi_n^{(k)}, \quad (3.1)$$

$$\widetilde{C}_n = \widehat{C}_n + \lambda_n I_d, \quad \widehat{C}_n = \frac{1}{K-1}\sum_{k=1}^{K}(\widehat{V}_n^{(k)} - \overline{\widehat{V}}_n) \otimes (\widehat{V}_n^{(k)} - \overline{\widehat{V}}_n).$$

The inflation parameter λ_n is chosen in such a way that it only plays a role when the filter is malfunctioning. In particular, when the filter is giving accurate predictions the inflation parameter will be zero. The inflation will be 'triggered' when either of two

statistical properties of the filter, Θ_n or Ξ_n, exceeds appropriately chosen thresholds, indicating that filter predictions have begun to stray far from the truth. Hence, we define

$$\lambda_n = \varphi(\Theta_n, \Xi_n), \quad \varphi(x,y) = c_\varphi x(1+y)\mathbb{1}_{\{x>M_1 \text{ or } y>M_2\}} \tag{3.2}$$

where φ plays the role of a cut-off function, c_φ is some fixed positive constant and M_1, M_2 are fixed positive thresholds constant used to decide whether the filter is functioning properly or not.

The first statistical quantity Θ_n, measures how far predicted observations are from actual observations:

$$\Theta_n := \sqrt{\frac{1}{K}\sum_{k=1}^{K}\left|H\widehat{V}_n^{(k)} - Z_n^{(k)}\right|^2}.$$

In a standard Kalman filter, the quantity $H\widehat{V}_n - Z_n$ is called the *innovation process* and is used for to test accuracy of the filter. Hence the quantity Θ_n should be thought of as the norm of the ensemble innovation process.

The second statistical property Ξ_n measures the forecast ensemble covariance between observed components and unobserved components. Using the observed–unobserved notation from Section 2.3, we have $\Xi_n = \|\widehat{B}_n\|$ where $\widehat{B}_n$ is the covariation between observed and unobserved components

$$\widehat{B}_n = \frac{1}{K-1}\sum_{k=1}^{K}(\widehat{X}_n^{(k)} - \overline{\widehat{X}}_n)\otimes(\widehat{Y}_n^{(k)} - \overline{\widehat{Y}}_n).$$

Intuitively, it is important to control this cross-variation, since this avoids the situation whereby the filter magnifies a small error in the observed component and imposes it on the unobserved component. This was a key mechanism leading to the filter instability found in [11].

Although both statistics Θ_n and Ξ_n can be intuitively connected to the performance of the filter, the real reason we choose to control these two statistics is a mathematical one. If one attempts to prove boundedness results of Section 4 for the unmodified EnKF, it becomes clear that the ensembles cannot be stabilized unless one has adequate control on these two quantities. Hence it is a simple observation that an adaptive inflation which guarantees control of these two statistics will therein guarantee boundedness and stability of the corresponding adaptive filter.

3.1. Adaptive inflation for other ensemble based filters. The adaptive inflation approach can be applied to other ensemble based filters for guaranteed stability with minimal changes to the original filtering method.

Two popular modifications to EnKF are constant additive covariance inflation and constant multiplicative covariance inflation. With these two modifications, the analysis step in EnKF operates with the ensemble covariance $\widehat{C}_n$ in (2.5) replaced by

$$\widetilde{C}_n = \widehat{C}_n + \rho I, \quad \widetilde{C}_n = (1+\rho)\widehat{C}_n, \tag{3.3}$$

respectively, and $\rho > 0$ here is a fixed constant. We will refer to EnKF with constant inflation as *EnKF-CI*, with the suffix standing for 'constant inflation'. Typically we will use additive rather than multiplicative inflation, but the theoretical results will hold for both verbatim.

We can combine constant and adaptive inflation by simply adding both simultaneously

$$\widetilde{C}_n = \widehat{C}_n + \rho I + \lambda_n I, \quad \widetilde{C}_n = (1+\rho)\widehat{C}_n + \lambda_n I \tag{3.4}$$

respectively in the analysis step (2.5). We will refer to this as *EnKF-CAI*, with the suffix standing for 'constant+adaptive inflation'.

In the case of the ESRFs introduced in Section 2.4, we follow the similar pattern of simply replacing $\widehat{C}_n$ with some inflated $\widetilde{C}_n$. However, this replacement only occurs when calculating the mean and not when calculating the posterior covariance. This is because C_n is the posterior ensemble covariance, its rank cannot exceed $K-1$, while additive inflation can make the right hand side of (2.9) of rank d, and in most practical cases the model dimension d is larger than the ensemble size K.

Hence the adaptively inflated version of ESRF is given by

$$\begin{aligned}
\overline{V}_n &= \overline{\widehat{V}}_n - \widetilde{C}_n H^T (I + H\widetilde{C}_n H^T)^{-1} (H\overline{\widehat{V}}_n - Z_n), \\
C_n &= \widehat{C}_n - \widehat{C}_n H^T (I + H\widehat{C}_n H^T)^{-1} H\widehat{C}_n,
\end{aligned} \tag{3.5}$$

where

$$\begin{aligned}
\widetilde{C}_n &= \widehat{C}_n + \lambda_n I \quad \text{(ETKF-AI / EAKF-AI)}, \\
\widetilde{C}_n &= \widehat{C}_n + \rho I \quad \text{(ETKF-CI / EAKF-CI)}, \\
\widetilde{C}_n &= \widehat{C}_n + \rho I + \lambda_n I \quad \text{(ETKF-CAI / EAKF-CAI)}.
\end{aligned} \tag{3.6}$$

Notice that we use the same nomenclature from Table 3.1 as in the case of EnKF. The adaptive inflation strength λ_n defined the same as EnKF, but with Θ_n now defined as

$$\Theta_n := \frac{1}{K} \sum_{k=1}^{K} |H\widehat{V}_n^{(k)} - Z_n|^2,$$

since $Z_n^{(k)}$ no longer plays a role in the setting of ESRF.

Acronym	Algorithm	Eq. ref.
EnKF	Ensemble Kalman filter	(2.5)
EnKF-AI	Ensemble Kalman filter w/ adaptive inflation	(3.1)
EnKF-CI	Ensemble Kalman filter w/ constant inflation	(3.3)
EnKF-CAI	Ensemble Kalman filter w/ constant + adaptive inflation	(3.4)

TABLE 3.1. *Acronyms for the ensemble methods.*

4. Energy principles

In this section we show that the adaptive filtering methods introduced in Section 3 inherit an energy principle from the underlying model. Using the notation for the model introduced in Section 2, the energy principle for the model takes the following form:

ASSUMPTION 4.1 (Energy principle). *There exist constants* $\beta_h \in (0,1), K_h > 0$, *such that*

$$\mathbb{E}_{n-1} |U_n|^2 \leq (1-\beta_h)|U_{n-1}|^2 + K_h \quad a.s.,$$

recalling that $\mathbb{E}_{n-1}$ denotes conditional expectation with respect to the σ-algebra $\mathcal{F}_{n-1} := \sigma(U_0, U_1, \ldots, U_{n-1})$. Equivalently,

$$|\Psi_h(u)|^2 + tr(R_h(u)) \le (1-\beta_h)|u|^2 + K_h,$$

for all for all $u \in \mathbb{R}^d$.

REMARK 4.2. There is a slight abuse of notation here, since we also use $\mathbb{E}_{n-1}$ to denote expectation conditioned on the history of the signal-ensemble process. However it is clear that in the above situation the two coincide.

When U_n is given by the discrete time formulation of an SDE (as described in (2.2) and the description below), it follows from a simple Gronwall argument that Assumption 4.1 holds as long as $\langle \psi(u), u \rangle \le -\beta|u|^2 + k$ with constants $\beta, k > 0$. Therefore, the discrete time formulation of truncated Navier–Stokes equation, Lorenz 63 and 96 models (as in (2.2)) all satisfy Assumption 4.1, explicit verifications can be found in Section 2.5 of [7].

It is natural to ask whether a filter inherits an energy principle from its underlying model and hence inherits a type of stability. In [7] it was shown that the non-adaptive ensemble based filters (EnKF, ESRF) inherit a stronger notion of energy principle called the observable energy principle. In particular, if the model satisfies

$$|H\Psi_h(u)|^2 + \text{tr}(HR_h(u)H^T) \le (1-\beta_h)|Hu|^2 + K_h, \qquad (4.1)$$

then the filter satisfies a related energy principle, which a Lyapunov function whose sublevel sets are only compact in the case where H is of full rank (complete observations). This is clearly a much weaker result than one would hope for, since one must assume (4.1) which is not thought to be true for partially observing filters and moreover the result is not as strong as desired since the Lyapunov function cannot be used for ergodicity arguments.

We will now show that the adaptive inflation filters completely avoid these issues. In particular, the adaptively inflated filters inherit the usual energy principle (Assumption 4.1) and moreover, the Lyapnuov function for the ensemble has compact sub-level sets for any choice of observation matrix. The inflation mechanism additionally ensures that the observed directions of the filter HV_n^k never stray from the (perturbed) observations Z_n^k, this is indeed the main purpose of the inflation.

THEOREM 4.3. *Let $\{V_n^{(k)}\}_{k=1}^K$ denote the EnKF-AI ensemble described by (3.1) with an inflation strength $\lambda_n \ge \varphi(\Theta_n, \Xi_n)$, then we have the following estimates:*

(i) *The ensemble innovation satisfies the almost sure bound*

$$|HV_n^{(k)} - Z_n^{(k)}| \le \sqrt{K} \max\{M_1, \rho_0^{-1} c_\varphi^{-1}\}$$

Here ρ_0 denotes the minimum eigenvalue of $H_0 H_0^T$.

(ii) *Suppose the model U_n additionally satisfies an energy principle (Assumption 4.1) and let*

$$\mathcal{E}_n = 4K\rho_0^{-1}\beta_h^{-1}\|H\|^2|U_n|^2 + \sum_{k=1}^K |V_n^{(k)}|^2.$$

Then there exists constant $\beta_h \in (0,1)$, $D_h > 0$ such that

$$\mathbb{E}_{n-1}\mathcal{E}_n \le (1 - \tfrac{1}{2}\beta_h)\mathcal{E}_{n-1} + D_h.$$

In particular, we have the time uniform bound $\sup_{n\ge 0} \mathbb{E}\mathcal{E}_n < \infty$.

REMARK 4.4.   It will be clear from the proof that the above result also holds verbatim for EnKF-CAI, introduced in (3.4).

A similar energy principle also holds for the adaptive inflation modifications of ESRF. Note that only the second part of the result holds for the ESRF case.

THEOREM 4.5.   Let $\{V_n^{(k)}\}_{k=1}^K$ denote the ETKF-AI / EAKF-AI ensemble described by (3.5) , (3.6) with any inflation strength $\lambda_n \geq \varphi(\Theta_n, \Xi_n)$, then Theorem 4.3 (ii) holds.

REMARK 4.6.   As above, it will be clear from the proof that the above result also holds verbatim for ETKF-CAI / EAKF-CAI introduced in (3.6).

We now provide the proof of Theorem 4.3, the proof Theorem 4.5 is similar and is therefore deferred until the appendix.

Proof.   (Proof of Theorem 4.3.)   Before proceeding, recall the 'observed-unobserved' notation from Section 2.3:

$$H = (H_0, 0), \quad V_n^{(k)} = (X_n^{(k)}, Y_n^{(k)}), \quad \widetilde{C}_n = \begin{bmatrix} \widetilde{C}_n^X, & \widetilde{B}_n \\ \widetilde{B}_n^T, & \widetilde{C}_n^Y \end{bmatrix},$$

$$\widehat{V}_n^{(k)} = (\widehat{X}_n^{(k)}, \widehat{Y}_n^{(k)}), \quad \widehat{C}_n = \begin{bmatrix} \widehat{C}_n^X, & \widehat{B}_n \\ \widehat{B}_n^T, & \widehat{C}_n^Y. \end{bmatrix}$$

We start with part (i). From the assimilation equation (3.1), by applying H to both sides, rearranging and applying the identity $H\widetilde{C}_n H^T = H_0 \widetilde{C}_n^X H_0^T$ we obtain

$$HV_n^{(k)} - Z_n^{(k)} = (I + H_0 \widetilde{C}_n^X H_0^T)^{-1}(H\widehat{V}_n^{(k)} - Z_n^{(k)}). \tag{4.2}$$

Since the minimum eigenvalue of matrix $H_0 H_0^T$ is $\rho_0 > 0$ and $H_0 \widehat{C}_n^X H_0^T$ is positive semi-definite, we have

$$I + H_0 \widetilde{C}_n^X H_0^T = I + H_0 \widehat{C}_n^X H_0^T + \varphi(\Theta_n, \Xi_n) H_0 H_0^T \succeq (1 + \rho_0 \varphi(\Theta_n, \Xi_n)) I_d.$$

Therefore, when $\Theta_n \leq M_1$, we have

$$|(I + H_0 \widetilde{C}_n^X H_0^T)^{-1}(H\widehat{V}_n^{(k)} - Z_n^{(k)})| \leq \frac{|H\widehat{V}_n^{(k)} - Z_n^{(k)}|}{1 + \rho_0 \varphi(\Theta_n, \Xi_n)} \leq |H\widehat{V}_n^{(k)} - Z_n^{(k)}| \leq \sqrt{K} M_1.$$

On the other hand, when $\Theta_n > M_1$, we have

$$|(I + H_0 \widetilde{C}_n^X H_0^T)^{-1}(H\widehat{V}_n^{(k)} - Z_n^{(k)})| \leq \frac{|H\widehat{V}_n^{(k)} - Z_n^{(k)}|}{1 + \rho_0 \varphi(\Theta_n, \Xi_n)} = \frac{|H\widehat{V}_n^{(k)} - Z_n^{(k)}|}{\rho_0 c_\varphi \Theta_n (1 + \Xi_n)}$$

$$\leq \frac{|H\widehat{V}_n^{(k)} - Z_n^{(k)}|}{\rho_0 c_\varphi \Theta_n}$$

$$\leq \sqrt{K} \rho_0^{-1} c_\varphi^{-1}.$$

This bound and (4.2) yield claim (i).

We now move on to part (ii), starting by bounding the observed part. Starting from the expression (2.7) and applying part (i), Young's inequality and the independence of $\xi_n^{(k)}$, we obtain

$$
\begin{aligned}
\mathbb{E}_{n-1}|X_n^{(k)}|^2 &\leq 2\mathbb{E}_{n-1}\left(|X_n^{(k)} - H_0^{-1}Z_n^{(k)}|^2 + |H^{-1}Z_n^{(k)}|^2\right) \\
&\leq 2\rho_0^{-1}\mathbb{E}_{n-1}(|HV_n^{(k)} - Z_n^{(k)}|^2 + |Z_n^{(k)}|^2) \\
&\leq 2\rho_0^{-1}K\max\{M_1^2, \rho_0^{-2}c_\varphi^{-2}\} + 2\rho_0^{-1}\|H\|^2\mathbb{E}_{n-1}|U_n|^2 + 2\rho_0^{-1}\mathbb{E}_{n-1}|\xi_n^{(k)} + \xi_n|^2 \\
&= 2\rho_0^{-1}K\max\{M_1^2, \rho_0^{-2}c_\varphi^{-2}\} + 2\rho_0^{-1}\|H\|^2\mathbb{E}_{n-1}|U_n|^2 + 4\rho_0^{-1}d. \quad (4.3)
\end{aligned}
$$

For the unobserved part, starting from (2.8) and applying Young's inequality (Lemma A.1)

$$
|Y_n^{(k)}|^2 \leq (1 + \tfrac{1}{2}\beta_h)|\widehat{Y}_n^{(k)}|^2 + (1 + 2\beta_h^{-1})|\widehat{B}_n^T H_0^T(I + H_0\widetilde{C}_n^X H_0^T)^{-1}(H_0\widehat{X}_n^{(k)} - Z_n^{(k)})|^2, \quad (4.4)
$$

where $\beta_h \in (0,1)$ is the one appearing in Assumption 4.1. For the first term on the right hand side, using Assumption 4.1 and the elementary inequality $(1 + \tfrac{1}{2}\beta_h)(1 - \beta_h) \leq (1 - \tfrac{1}{2}\beta_h)$ we have

$$
\mathbb{E}_{n-1}(1 + \tfrac{1}{2}\beta_h)|\widehat{Y}_n^{(j)}|^2 \leq (1 + \tfrac{1}{2}\beta_h)\mathbb{E}_{n-1}|\widehat{V}_n^{(j)}|^2 \leq (1 - \tfrac{1}{2}\beta_h)|V_{n-1}^{(j)}|^2 + (1 + \tfrac{1}{2}\beta_h)K_h.
$$

Since $\Xi_n = \|\widehat{B}_n\|$, the second part of (4.4) is bounded by

$$
\begin{aligned}
|\widehat{B}_n^T H_0^T(I + H_0\widetilde{C}_n^X H_0^T)^{-1}(H_0\widehat{X}_n^{(j)} - Z_n^{(j)})|^2 &\leq \Xi_n^2\|H_0\|^2\|(I + H_0\widetilde{C}_n^X H_0^T)^{-1}\|^2|H_0\widehat{X}_n^{(j)} - Z_n^{(j)}|^2 \\
&\leq K\Theta_n^2\Xi_n^2\|H_0\|^2\|(I + H_0\widetilde{C}_n^X H_0^T)^{-1}\|^2
\end{aligned}
$$

where in the last inequality we used the bound $|H_0\widehat{X}_n^{(j)} - Z_n^{(j)}|^2 \leq K\Theta_n^2$. Notice that, as in the proof of part (i)

$$
I + H_0\widetilde{C}_n^X H_0^T \succeq (1 + \lambda_n\rho_0)I_q.
$$

Hence, when $\Theta_n > M_1$ or $\Xi_n > M_2$ we have

$$
\|(I + H_0\widetilde{C}_n^X H_0^T)^{-1}\|\Theta_n\Xi_n \leq \frac{\Theta_n\Xi_n}{1 + \rho_0 c_\varphi\Theta_n(\Xi_n + 1))} \leq \rho_0^{-1}c_\varphi^{-1},
$$

and when $\Theta_n \leq M_1$ and $\Xi_n \leq M_2$,

$$
\|(I + H_0\widetilde{C}_n^X H_0^T)^{-1}\|\Theta_n\Xi_n \leq \Theta_n\Xi_n \leq M_1 M_2.
$$

In conclusion, the following holds almost surely

$$
|\widehat{B}_n^T H_0^T(I + H_0\widetilde{C}_n^X H_0^T)^{-1}(H_0\widehat{X}_n^{(j)} - Z_n^{(j)})|^2 \leq K\|H_0\|^2\max\{M_1^2 M_2^2, \rho_0^{-2}c_\varphi^{-2}\}. \quad (4.5)
$$

By combining the above estimates with (4.4), the sum of (4.3) and (4.4) over all k yields the following inequality with a constant $D_2 > 0$

$$
\mathbb{E}_{n-1}\sum_{k=1}^K |V_n^{(k)}|^2 - 2K\rho_0^{-1}\|H_0\|^2\mathbb{E}_{n-1}|U_n|^2 \leq (1 - \tfrac{1}{2}\beta_h)\sum_{k=1}^K |V_{n-1}^{(k)}|^2 + D_2. \quad (4.6)
$$

Then notice a multiple of Assumption 4.1 is

$$4K\beta_h^{-1}\rho_0^{-1}\|H\|^2\mathbb{E}_{n-1}|U_n|^2 \leq 4K\beta_h^{-1}\rho_0^{-1}\|H\|^2(1-\beta_h)|U_{n-1}|^2 + 4K\beta_h^{-1}\rho_0^{-1}\|H\|^2 K_h.$$

The sum of the two previous two inequalities yields

$$
\begin{aligned}
\mathbb{E}_{n-1}\mathcal{E}_n &= \sum_{k=1}^{K}\mathbb{E}_{n-1}|V_n^{(k)}|^2 + 4K\beta_h^{-1}\rho_0^{-1}(1-\tfrac{1}{2}\beta_h)\|H\|^2\mathbb{E}_{n-1}|U_n|^2 \\
&= \sum_{k=1}^{K}\mathbb{E}_{n-1}|V_n^{(k)}|^2 - 2K\rho_0^{-1}\|H\|^2\mathbb{E}_{n-1}|U_n|^2 + 4K\beta_h^{-1}\rho_0^{-1}\|H\|^2\mathbb{E}_{n-1}|U_n|^2 \\
&\leq (1-\tfrac{1}{2}\beta_h)\sum_{k=1}^{K}|V_{n-1}^{(k)}|^2 + 4K\beta_h^{-1}\rho_0^{-1}\|H\|^2(1-\beta_h)|U_{n-1}|^2 + D_h \\
&\leq (1-\tfrac{1}{2}\beta_h)\sum_{k=1}^{K}|V_{n-1}^{(k)}|^2 + 4K\beta_h^{-1}\rho_0^{-1}\|H\|^2(1-\tfrac{1}{2}\beta_h)^2|U_{n-1}|^2 + D_h \\
&= (1-\tfrac{1}{2}\beta_h)\mathcal{E}_{n_1} + D_h,
\end{aligned}
$$

with $D_h := 4K\beta_h^{-1}\rho_0^{-1}\|H\|^2 K_h + D_2$ and using the fact that $(1-\beta_h) \leq (1-\tfrac{1}{2}\beta_h)^2$.   □

5. Geometric ergodicity

A stochastic process is geometrically ergodic if it converges to a unique invariant measure geometrically fast. In this section we will show that the signal–ensemble process (for each filter introduced in this article) is geometrically ergodic. In particular, if P denotes the Markov transition kernel for the signal ensemble process, then we will show that there exists a constant $\gamma \in (0,1)$ such that

$$\|P^n\mu - P^n\nu\|_{TV} \leq C_{\mu,\nu}\gamma^n, \tag{5.1}$$

where μ,ν are two arbitrary initial probability distributions, $C_{\mu,\nu}$ is a time uniform constant that depends on μ,ν, and $\|\cdot\|_{TV}$ denotes the total variation norm.

Geometric ergodicity is a notion of stability for the filter. In particular it implies that discrepancies in the initialization of the filter will dissipate exponentially quickly. For the linear Kalman filter and optimal filters, it is known that under mild assumptions geometric ergodicity is extended from model to filter [24, 25].

In theorems 5.5, 5.6, and 5.8 of [7], the authors proved that if the system noise ζ_n is non-degenerate and the signal-ensemble process has an energy principle, then EnKF, ETKF and a natural version of EAKF will be geometrically ergodic. In particular, non degeneracy for the system noise ζ_n means the following.

ASSUMPTION 5.1 (Nondegenerate system noise).   *For any constants $R_1, R_2 > 0$, there is a constant $\alpha > 0$ such that*

$$\mathbb{P}(\zeta_n \in \cdot \,|\, U_{n-1} = u) \geq \alpha\lambda_{R_2}(\cdot)$$

for all $|u| \leq R_1$, where $\lambda_{R_2}(dx)$ is the Lebesgue measure of $\mathbb{R}^d$ restricted to $\{u : |u| \leq R_2\}$.

Assumption 5.1 holds for many practical examples. When U_n is produced by time discretization of an SDE (2.2), it suffices to require Σ being nonsingular, see Appendix B in [7] for a detailed discussion.

In contrast with the ergodicity results of [7], we no longer require the assumption of a Lyapunov function with compact sub-level sets, since this is guaranteed by theorems 4.3 and 4.5. This is a vast improvement as in the non-adaptive case such Lyapunov functions are only known to exist in the (essentially) fully observed scenario.

Before stating the result, recall that $M_1, M_2 > 0$ are the thresholds used to decide when to trigger the adaptive inflation mechanism. For EnKF-AI and ETKF-AI, the thresholds have no constraints besides being positive, but for EAKF-AI we require that the constants are sufficiently large.

THEOREM 5.2.   *If the system noise in the signal process is non-degenerate (Assumption 5.1) then the signal-ensemble processes $(U_n, V_n^{(1)}, \dots, V_n^{(K)})$ generated by EnKF-AI (with any $M_1, M_2 > 0$), ETKF-AI (with any $M_1, M_2 > 0$) or a version of EAKF-AI (with M_1, M_2 sufficiently large) are geometrically ergodic.*

REMARK 5.3.   The same result holds verbatim for adaptive filters with additional constant covariance inflation (EnKF-CAI,ETKF-CAI,EAKF-CAI), which are defined in Section 3.1.

As a matter of fact, the above result is an immediate corollary of [7]. The reasoning will become clear after we review the major steps in [7].

5.1. A controllability framework.   The central piece of the arguments is a result of Markov chain theory [26]. Here we use a simple adaptation of the form given in [27, Theorem 2.3].

THEOREM 5.4.   *Let $\mathbf{X}_n$ be a Markov chain in a space E such that*

1. *(Lyapunov function) There is a function $\mathcal{E} : E \mapsto \mathbb{R}^+$ with compact sub-level sets and such that*

$$\mathbb{E}_{n-1} \mathcal{E}(\mathbf{X}_n) \leq (1 - \beta) \mathcal{E}(\mathbf{X}_{n-1}) + K,$$

   for some $\beta \in (0,1)$ and $K > 0$.

2. *(Minorization condition) Let $C_R = \{x : \mathcal{E}(x) \leq R\}$. For all $R > 0$, there exists a probability measure ν on E with $\nu(C_R) = 1$ and a constant $\eta > 0$ such that for any measurable set $A \subset E$*

$$\mathbb{P}(\mathbf{X}_n \in A | \mathbf{X}_{n-1} = x) \geq \eta \nu(A)$$

   for all $x \in C_R$.

Then there is a unique invariant measure π and constants $r \in (0,1), \kappa > 0$ such that

$$\|\mathbb{P}^\mu(\mathbf{X}_n \in \cdot) - \pi\|_{TV} \leq \kappa r^n \left(1 + \int \mathcal{E}(x)\mu(dx)\right).$$

Theorems 4.3 and 4.5 have already provided the Lyapunov function. All that remains is to show the minorization condition.

To obtain minorization, we will exploit the update structure of the ensemble filters, precisely as was done in [7]. Notice that for all ensemble filters, the signal-ensemble process $\mathbf{X}_n := (U_n, V_n^{(1)}, \dots, V_n^{(K)})$ is a Markov chain taking values in $\mathcal{X} := \mathbb{R}^d \times \mathbb{R}^{d \times K}$. The evolution of $\mathbf{X}_n$ is described by the composition of two maps. The first is a random map from $\mathcal{X}$ to a signal-forecast-observation space $\mathcal{Y}$, described by a Markov kernel $\Phi : \mathcal{X} \times \mathcal{B}(\mathcal{Y}) \to [0,1]$. The second is a deterministic map $\Gamma : \mathcal{Y} \to \mathcal{X}$, which combines the forecast with the observed data to produce the updated posterior ensemble. The details

of these maps, as well as the definition of the intermediate space $\mathcal{Y}$, differs between EnKF, ETKF, EAKF, and their adaptive counterparts.

For example, in EnKF, the intermediate space is $\mathcal{Y} := \mathbb{R}^d \times \mathbb{R}^{d \times K} \times \mathbb{R}^{q \times K}$ and the random mapping is

$$(U_{n-1}, V_{n-1}^{(1)}, \ldots, V_{n-1}^{(K)}) \mapsto \mathbf{Y}_n := (U_n, \widehat{V}_n^{(1)}, \ldots, \widehat{V}_n^{(K)}, Z_n^{(1)}, \ldots, Z_n^{(K)}).$$

The deterministic map Γ is given by the Kalman update

$$\Gamma(U_n, \widehat{V}_n^{(1)}, \ldots, \widehat{V}_n^{(K)}, Z_n^{(1)}, \ldots, Z_n^{(K)}) = (U_n, \Gamma^{(1)}, \ldots, \Gamma^{(K)}),$$

where

$$\Gamma^{(k)} = \widehat{V}^{(k)} - \widetilde{C} H^T (I + H \widehat{C} H^T)^{-1} (H \widehat{V}^{(k)} - Z^{(k)}),$$

$$\widetilde{C} = \frac{1}{K-1} \sum_{k=1}^{K} (\widehat{V}^{(k)} - \overline{\widehat{V}}) \otimes (\widehat{V}^{(k)} - \overline{\widehat{V}}). \tag{5.2}$$

For EnKF-AI, we have the same formulas, but with the exception

$$\widetilde{C} = \frac{1}{K-1} \sum_{k=1}^{K} (\widehat{V}^{(k)} - \overline{\widehat{V}}) \otimes (\widehat{V}^{(k)} - \overline{\widehat{V}}) + c_\varphi \Theta_n (1 + \Xi_n) \mathbb{1}_{\Theta_n > M_1 \text{ or } \Xi_n > M_2} I.$$

and

$$\Theta_n = \sqrt{\frac{1}{K} \sum_{k=1}^{K} \left| H \widehat{V}_n^{(k)} - Z_n^{(k)} \right|^2}, \quad \Xi_n = \left\| \frac{1}{K-1} \sum_{k=1}^{K} (\widehat{X}_n^{(k)} - \overline{\widehat{X}}_n) \otimes (\widehat{Y}_n^{(k)} - \overline{\widehat{Y}}_n) \right\|.$$

The corresponding formulas for ETKF, EAKF (and their adaptive counterparts) can be similarly derived and for concreteness are given in Appendix C. Notice that for values of $\mathbf{Y}_n$ close to the origin, the adaptive inflation term will vanish and hence the deterministic map Γ for EnKF-AI coincides with that of EnKF near the origin. Similar statements hold for ETKF-AI and EAKF-AI.

Given this formulation, it suffices to show that the push-forward kernel $\Gamma^* \Phi(x, \cdot) = \Phi(x, \Gamma^{-1}(\cdot))$ satisfies the minorization condition. It is easy to see that, given the assumptions on the noise, the kernel $\Phi(x, \cdot)$ has a density with respect to Lebesgue measure, so we simply need to show that the pushforward inherits the density property from Φ. This can be understood as the controllability of the map Γ. To achieve this, we use the following simple fact, which is Lemma 5.4 in [7].

LEMMA 5.5.   *Let Φ be a Markov transition kernel from $\mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^m$ with a Lebesgue density $p(x, y) = p(x, (y_1, y_2))$ and let $\Gamma : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$. Given a compact set C, suppose that there is a point $y^* = (y_1^*, y_2^*) \in \mathbb{R}^n \times \mathbb{R}^m$ and $\beta > 0$ such that*

1. *Reachable from all $x \in C$, that is the density function $p(x, y) > \beta$ for y around y^*,*

2. *Γ is controllable around y^*, that is Γ is C^1 near y^* and $\det(\mathcal{D}_{y_1} \Gamma)|_{y^*} > 0$.*

Then there is a $\delta > 0$ and a neighborhood O_1 of $\Gamma(y^*)$ such that for all $x \in C$

$$\Gamma^* \Phi(x, \cdot) \geq \delta \lambda_{O_1}(\cdot)$$

where λ_{O_1} is the Lebesgue measure restricted to the set O_1. In other words, the minorization condition holds for the transition kernel $\Gamma^* \Phi$.

5.2. Application to ensemble filters with adaptive inflation. To apply Lemma 5.5 to the proof of Theorem 5.2 we will use the variables

$$x = (U_{n-1}, V_{n-1}^{(1)}, \ldots, V_{n-1}^{(K)})$$
$$y_1 = (U_n, \widehat{V}_n^{(1)}, \ldots, \widehat{V}_n^{(K)}) \tag{5.3}$$
$$y_2 = (Z_n, Z_n^{(1)}, \ldots, Z_n^{(K)})$$

The choice of the intermediate point $y^* = (y_1^*, y_2^*)$ can be quite delicate and can simplify the non-degeneracy condition considerably. In particular, we prove the following proposition in [7].

PROPOSITION 5.6. *Let Φ, Γ be the Markov kernel and deterministic map used to define EnKF/ETKF/EAKF, given by the formulas in Appendix C. If the system noise is non-degenerate (Assumption 5.1) then Φ, Γ satisfies both conditions of Lemma 5.5 with the following choice of intermediate point y^*:*

1. EnKF: y^* has all its components being at the origin.

2. ETKF: y^* has all its components being at the origin.

3. EAKF: y_2^* has all its components being at the origin, U^* is at the origin, while $[\widehat{V}^{*(1)}, \ldots, \widehat{V}^{*(K)}]$ is the $d \times K$ matrix given by

$$\begin{cases} \begin{bmatrix} \vec{1} \ \Xi_0 \ 0 \end{bmatrix} & \text{if } d \leq K - 1 \\ \begin{bmatrix} \vec{1} \ \Xi_0 \\ 0 \ 0 \end{bmatrix} & \text{if } d \geq K - 1, \end{cases}$$

where $\vec{1}$ is the $d \times 1$ vector of 1's and Ξ_0 is the $r \times r$ matrix with $r = \min\{K - 1, d\}$:

$$\Xi_0 = \begin{bmatrix} 1 & 1 & \ldots & 1 & -r \\ \vdots & \vdots & \vdots & \cdot^{\cdot^{\cdot}} & 0 \\ 1 & 1 & -3 & \ldots & 0 \\ 1 & -2 & 0 & \ldots & 0 \\ -1 & 0 & 0 & \ldots & 0 \end{bmatrix}.$$

Note in this case, a version of EAKF is picked so Γ is C^1 near y^*.

With this proposition at hand, Theorem 5.2 becomes an immediate corollary.

Proof. (Proof of Theorem 5.2.)   For each of EnKF-AI, ETKF-AI, and EAKF-AI verification of the reachability condition in Lemma 5.5 is identical to that of EnKF, ETKF, and EAKF which is trivial and is done in [7, Proof of theorems 5.5, 5.6, 5.8], respectively. Hence it suffices to check the non-degeneracy of Γ in each of the three cases. For EnKF-AI and ETKF-AI we pick $y^* = (y_1^*, y_2^*)$ to be the origin. Hence for thresholds $M_1, M_2 > 0$, there exist a neighborhood of the origin such that the map Γ coincides with that of EnKF, ETKF respectively, since the inflation mechanism is not triggered within this neighborhood. Hence the non-degeneracy follows immediately from Proposition 5.6. For EAKF-AI, as detailed above, the intermediate point y^* is not chosen at the origin but rather at the point specified in Proposition 5.6. Nevertheless, since this point is fixed, one can pick thresholds M_1, M_2 sufficiently large so that for some neighborhod of y^* the inflation mechanism is not triggered and hence Γ agrees with that of EAKF. Hence the non-degeneracy follows from Proposition 5.6.   □

6. Thresholds from elementary benchmarks

Finally, we discuss the choice of the cutoff thresholds M_1 and M_2 in the formulation (3.2) of φ. While Theorem 4.3 guarantees filter stability for any $M_1, M_2 > 0$, certain values produce better filter performance. Intuitively, these thresholds should be chosen to differentiate malfunctioning forecast ensembles from properly working ones.

Here we devise an elementary benchmark of accuracy that should be surpassed by any properly functioning filter. At each observation step, the benchmark estimate is given by the stationary state U_n conditioned on the observation Z_n. This can be computed using Bayes' formula and only requires access to the invariant measure π of the model. For instance we can compute the conditional expectation via

$$\mathbb{E}(U_n|Z_n) = \frac{\int x\pi(x)\exp(-\frac{1}{2}|Z_n - Hx|^2)dx}{\int \pi(x)\exp(-\frac{1}{2}|Z_n - Hx|^2)dx}. \tag{6.1}$$

The benchmark for accuracy is given by the mean-square error of the above estimator:

$$\text{Error}_A := \mathbb{E}|U_n - \mathbb{E}(U_n|Z_n)|^2 = \mathbb{E}|U_n|^2 - \mathbb{E}|\mathbb{E}(U_n|Z_n)|^2.$$

A properly functioning forecast ensemble should be expected to perform better than this estimator. For instance, one should expect a properly functioning filter to satisfy

$$\frac{1}{K}\sum_{k=1}^{K}\mathbb{E}\left|\widehat{V}_n^{(k)} - U_n\right|^2 \le \text{Error}_A.$$

Consequently, one should expect that

$$\mathbb{E}\Theta_n^2 = \frac{1}{K}\sum_{k=1}^{K}\mathbb{E}|H(\widehat{V}_n^{(k)} - U_n) + \xi_n + \xi_n^{(k)}|^2 \le \sigma_\Theta^2 := \|H\|^2\text{Error}_A + 2d.$$

Moreover using the inequalities $\|a \otimes b\| \le |a||b|$, $|x||y| \le \frac{1}{2}(x^2 + y^2)$ and Lemma A.2 we expect

$$\mathbb{E}\Xi_n \le \frac{1}{K-1}\sum_{k=1}^{K}\mathbb{E}\left|\widehat{X}_n^{(k)} - \overline{\widehat{X}}_n\right|\left|\widehat{Y}_n^{(k)} - \overline{\widehat{Y}}_n\right| \le \frac{1}{2K-2}\sum_{k=1}^{K}\mathbb{E}|\widehat{V}_n^{(k)} - \overline{\widehat{V}}_n|^2$$

$$\le \frac{1}{2K-2}\sum_{k=1}^{K}\mathbb{E}|\widehat{V}_n^{(k)} - U_n|^2 \le M_\Xi := \frac{K}{2K-2}\text{Error}_A.$$

In summary, we expect that a properly functioning filter should satisfy

$$\mathbb{E}\Theta_n^2 \le \sigma_\Theta^2 := \|H\|^2\text{Error}_A + 2d, \quad \mathbb{E}\Xi_n \le M_\Xi := \frac{K}{2K-2}\text{Error}_A. \tag{6.2}$$

Since the integrals appearing in the Bayesian formula (6.1) are often difficult to compute, it is more useful to have an estimator that only relies on low order statistics of π. For this purpose it is natural to introduce a Gaussian approximation of π, replacing π with $\tilde{\pi} = \mathcal{N}(\mathbb{E}U_n, \text{cov}(U_n))$ where the mean and covariance are the true statistics computed from π. The conditional distribution of $\tilde{\pi}$ given the observation Z_n is now a Gaussian measure and can be computed exactly. In particular we have that

$$\widetilde{\mathbb{E}}(U_n|Z_n) = \mathbb{E}U_n - \text{cov}(U_n)H^T(I + H^T\text{cov}(U_n)H)^{-1}(H\mathbb{E}U_n - Z_n).$$

The covariance of the error $r_K = U_n - \widetilde{\mathbb{E}}(U_n|Z_n)$ is given by the Kalman prior–posterior covariance relation

$$\text{cov}(r_K) = \text{cov}(U_n) - \text{cov}(U_n)H^T(I + H\text{cov}(U_n)H^T)^{-1}H\text{cov}(U_n), \qquad (6.3)$$

which can be explicit computed. The mean square error is given by

$$\text{Error}_A = \mathbb{E}|r_K|^2 = \text{tr}(\text{cov}(r_K)).$$

In the case of a linear model, this Gaussian approximation is of course exact.

Given the above discussion, an obvious option for the thresholds is

$$M_1 = \sigma_\Theta, \quad M_2 = M_\Xi. \qquad (6.4)$$

we call this *aggressive thresholding*. In this situation the adaptive inflation is triggered as soon as the filter statistics Θ_n, Ξ_n exceed the mean values computed from the crude estimators. A less aggressive strategy is to obtain M_1, M_2 by hypothesis testing, which we do not pursue here.

REMARK 6.1.   *To avoid wasting resources on computing these benchmarks, one could instead use the trivial estimator $U_n = 0$. In this case, one can compute*

$$\mathbb{E}|U_n - \mathbf{0}|^2 = \mathbb{E}|U_n|^2 \leq \beta_h^{-1}K_h.$$

A properly functioning EnKF should have a smaller mean square error in its forecast. And as a consequence, we can replace term Error_A in the bounds above with $\beta_h^{-1}K_h$, which leads to

$$\mathbb{E}\Theta_n^2 \leq \|H\|^2\beta_h^{-1}K_h + 2d, \quad \mathbb{E}\Xi_n \leq \frac{K}{2K-2}\beta_h^{-1}K_h.$$

7. Numerical results

In this section, we will numerically validate our results through the following five mode Lorenz 96 model:

$$\frac{d}{dt}x_i = x_{i-1}(x_{i+1} - x_{i-2}) - x_i + F, \quad i = 1,\dots,5. \qquad (7.1)$$

Here the indices are interpreted in modulo 5 sense. The truth will be given by $U_n = x(nh)$ for some observation time step h. Previous studies in [4,10,28] have shown that different choice of forcing F produces very different strength of turbulence. Here we will pick forcing $F = 4,8,16$, and see that these choices create very different filtering results. In [29], careful numerical studies have suggested that (7.1) does indeed exhibit good mixing properties. However, it is important to note that it is deterministic and therefore does not satisfy the ergodicity assumptions of Theorem 5.2. The justification for using this model for numerical experimentation is twofold: First, it is a realistic prototype for geophysical turbulence. And second, it exhibits that the theoretical results are robust and should hold for much more general assumptions on the mixing properties of the model.

Throughout the section we assume that only one component x_1 is observed, hence we have the observation matrix $H = [1,0,0,0,0]$. The observational noise is chosen to be distributed as $\mathcal{N}(0,0.01)$. We apply a selection of ensemble assimilation methods to this signal-observation system with 6 ensemble members, where the initial positions

are drawn from a Gaussian approximation of the equilibrium measure. For the sake of simplicity, we integrate (7.1) using an explicit Euler scheme with time step $\Delta = 10^{-4}$. For most ensemble based filtering methods, the explicit Euler scheme is a bad choice of integrator for (7.1), since the stiffness of the equation will lead to numerical blow-up for sufficiently large initial conditions. Moreover, the explicit method breaks the dissipative structure of (7.1), hence we are once again stepping outside of the theory developed in the preceding pages. As we shall see, an explicit method leads to a prevalence of catastrophic filter divergence in most ensemble based filtering methods. The adaptive inflation method on the other hand avoids this issue and exhibits high filtering skill. By employing an explicit method, we demonstrate the power of our adaptive inflation method, in that it allows one to use cheap integration schemes without risking catastrophic filter divergence and without sacrificing filtering skill. In Section 7.3, we perform experiments with different choices of integrator (including implicit integrators) and find no improvement in filtering skill over an explicit method with adaptive inflation, despite the vast increase in computational cost.

Here we compare the performance of four ensemble based filters: EnKF, EnKF-AI, EnKF-CI, and EnKF-CAI, which have been defined in sections 2 and 3. The constant inflation is always taken to be additive.

Following the methodology introduced in Section 6, we construct the equilibrium measure of the Lorenz '96 system using data from a long time ($T = 10^4$) simulation, which leads to the performance benchmarks computed by (6.3) and (6.2). The thresholds M_1 and M_2 are obtained through the aggressive strategy (6.4), using the Gaussian approximation of the stationary measure π. In order for the comparison to be unbiased, we use the same realizations of observation noise ξ_n and perturbations $\xi_n^{(k)}$ in all four filters. The statistics are collected for $N = 100$ independent trials, where each runs a total time length of $T = 100$. To eliminate transients, we will only record data from the period $50 \leq T \leq 100$.

The performances of the four filters are measured and compared from three perspectives:

- The most important question here is how often does catastrophic filter divergence appear in standard (non-adaptive) filters, like EnKF and EnKF-CI, and does the adaptive inflation in EnKF-AI and EnKF-CAI prevent it. Catastrophic filter divergence can be identified by checking whether the ensemble at the final time $T = 100$ takes value "NaN", representing machine infinity, in any of its components.

- The overall accuracy of the filters is compared using the average of the root mean square error (RMSE) over $N = 100$ trials among the second half of the time interval, where

$$\text{RMSE} = \sqrt{\frac{2}{T} \sum_{n=T/2h}^{T/h} |\overline{V}_n - U_n|^2}.$$

  Another statistics that is useful in judging the accuracy is the averaged pattern correlation:

$$\text{Cor} = \frac{2}{T} \sum_{n=T/2}^{T} \frac{\langle \overline{V}_n - \overline{U}, U_n - \overline{U} \rangle}{|\overline{V}_n - \overline{U}||U_n - \overline{U}|},$$

where $\overline{U}$ denotes the climatological average of the true model (the mean of the invariant measure) and $\overline{V}_n$ as usual denotes the average of the filter ensemble. We will also plot the of the posterior error, $|\overline{V}_n - U_n|$, for one realization of a typical trajectory for the four filters.

The following experiments have also been carried out with both ESRF methods. All these methods have performance very similar to EnKF based method, as well as the effect of adaptive inflation over them. We do not present these results for brevity.

7.1. Accuracy for different turbulent regimes. In this section we compare the performance of the four filters across the three turbulent regimes. In each experiment we use observation interval $h=0.05$ and (when required) constant additive inflation strength $\rho=0.1$.

7.1.1. Weak turbulence. Here we compare the performance of the filters in a weakly turbulent regime, by setting $F=4$. The performances of the four filters are presented in Table 7.1, where we compare frequency of catastrophic filter divergence, RMSE and pattern correlation. In Figure 7.1, we compare the posterior error for the four filters for one typical trajectory.

We found that, among 100 trials, none of the filters experience catastrophic filter divergence. The adaptive inflation in EnKF-AI has been triggered in 30 trials, and on average 1.96 times in each trial. In EnKF-CAI, the adaptive inflation has been turned on in 9 trials, and only once in each trial. Hence the typical realizations of EnKF and EnKF-AI are identical and similarly for EnKF-CI and EnKF-CAI, since the adaptive inflation is seldom triggered.

In terms of accuracy, all four filters perform quite well. For each filter the RMSE is significantly lower than the benchmark RMSE, which is 3.64 and the pattern correlation is very high. Unsurprisingly, the filters with constant covariance inflation (EnKF-CI and EnKF-CAI) have more skill than EnKF, with lower RMSE and higher pattern correlation. Somewhat more surprising is that EnKF-AI also performs better than EnKF, both in RMSE and pattern correlation, even though the adaptive covariance inflation has only been switched on in 30 of the 100 trials. This indicates that the performance of EnKF must be quite bad in those 30 and the adaptive inflation is needed to ensure the accuracy. Of course, in this regime one could also achieve this with non-adaptive constant inflation.

Filter	EnKF	EnKF-AI	EnKF-CI	EnKF-CAI
Cata. Div.	0%	0%	0%	0%
RMSE	0.89	0.54	0.22	0.22
Pattern Cor.	0.91	0.96	0.98	0.98

TABLE 7.1. *In the weak turbulence regime ($F=4$), for the four algorithms we compare frequency of catastrophic filter divergence (Cata. Div.), RMSE, and pattern correlation (Pattern Cor.) The benchmark RMSE is 3.25.*

We now briefly describe how the thresholds and benchmarks were calculated. The equilibrium mean on each mode is $\bar{x}_i \approx 1.22$ and the variance is $\mathrm{var}(x_i) \approx 3.38$. Using Kalman one time assimilation, the posterior variance of x_1 will be reduced to 0.01, while the benchmark RMSE will be 3.25. From these statistics, we compute $M_1 = \sigma_\Theta = 32.5$ and $M_2 = M_\Xi = 6.2$.
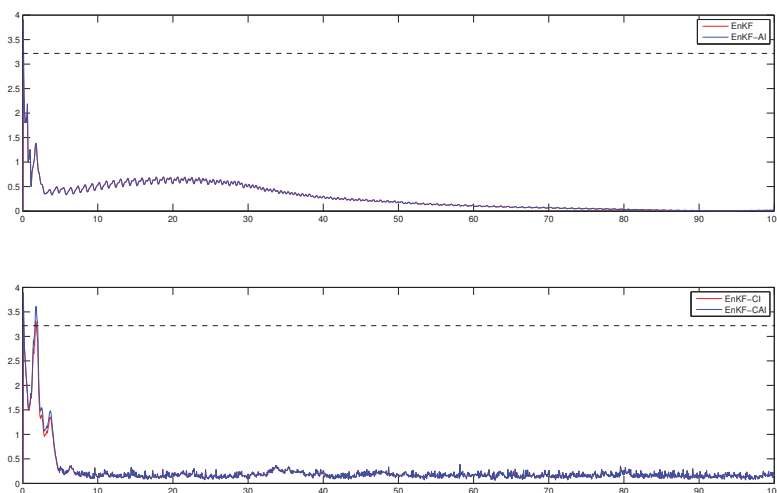
FIG. 7.1. *We compare the posterior error for one realization of the four filters in the weakly turbulent (F = 4) regime. Since the adaptive inflation is not triggered, trajectories collapse into two groups (those with constant inflation and those without). The benchmark RMSE is included (black dash) to emphasize the accuracy of the filters.*

7.1.2. Moderate turbulence.   Here we compare the performance of the filters in a moderately turbulent regime, by setting $F = 8$. The performances of the four filters are presented in Table 7.2, where we compare frequency of catastrophic filter divergence, RMSE, and pattern correlation. In Figure 7.2, we compare the posterior error for the four filters for one typical trajectory.

Among 100 trials, catastrophic filter divergence takes place 12 times for EnKF, but never takes place for the other three. The adaptive inflation in EnKF-AI has been turned on in 96 trials, while on average 17.5 times in each of these trials. As for EnKF-CAI, the adaptive inflation has been turned on in 20 trials, while on average 6.05 times each of these trials.

In terms of accuracy, EnKF will not be analyzed due to the frequent occurrence of catastrophic filter divergence. The EnKF-AI avoids the issue of catastrophic filter divergence, however it has very little skill, since the RMSE is higher than the benchmark RMSE 7.02 and the pattern correlation is quite low. However, the filters with constant inflation (EnKF-CI and EnKF-CAI), behave almost identically, display significantly better skill, beating the benchmark RMSE, and having high pattern correlation.

Filter	EnKF	EnKF-AI	EnKF-CI	EnKF-CAI
Cata. Div.	12%	0%	0%	0%
RMSE	NaN	8.6	3.61	3.57
Pattern Cor.	NaN	0.55	0.89	0.89

TABLE 7.2. *In the moderate turbulence regime (F = 8), for the four algorithms we compare frequency of catastrophic filter divergence (Cata. Div.), RMSE, and pattern correlation (Pattern Cor.) The benchmark RMSE is 7.02.*
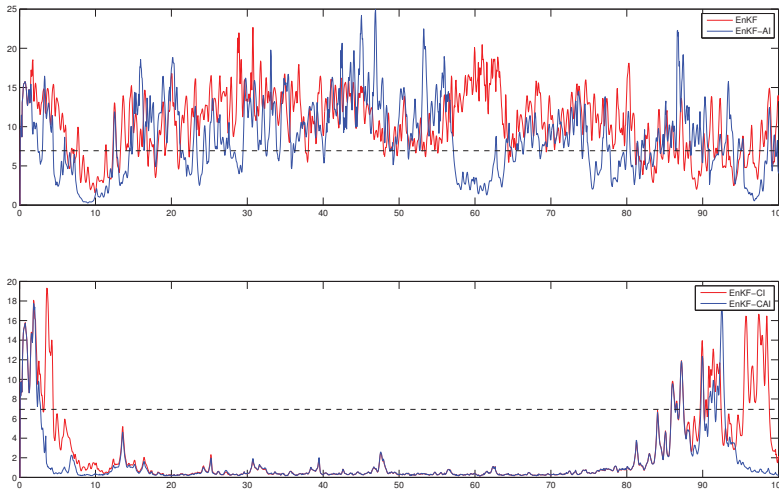
FIG. 7.2. *We compare the posterior error for one realization of the four filters in the moderately turbulent ($F = 8$) regime. The EnKF and EnKF-AI do not perform well, as they frequently exceed the benchmark (black dashed). Adding constant covariance inflation significantly improves the performance.*

We now briefly describe how the thresholds and benchmarks were calculated. The equilibrium mean on each mode is $\bar{x}_i \approx 2.28$ and the variance is $\mathrm{var}(x_i) \approx 12.6$. Using Kalman one time assimilation, the posterior variance of x_1 will be reduced to 0.01, while the benchmark RMSE will be 7.02. From these statistics, we compute $M_1 = \sigma_\Theta = 69.56$ and $M_2 = M_\Xi = 28.8$.

7.1.3. Strong turbulence. Here we compare the performance of the filters in a highly turbulent regime, by setting $F = 16$. The performances of the four filters are presented in Table 7.3, where we compare frequency of catastrophic filter divergence, RMSE and pattern correlation. In Figure 7.3, we compare the posterior error for the four filters for one typical trajectory.

In this highly turbulent regime, both non-adaptive filters, EnKF and EnKF-CI display frequent catastrophic filter divergence, making them practically useless. Among 100 trials, catastrophic filter divergence occurs in every trial of EnKF and in 18 trials of EnKF-CI. Both EnKF-AI and EnKF-CAI experience no catastrophic filter divergence. The adaptive inflation has been turned on in all trials of EnKF-AI, occurring 97.53 times on average per trial. The adaptive inflation has been turned on for EnKF-CAI in 80 trials, occurring 18.7 times on average per trial.

In terms of accuracy, the EnKF-AI exhibits very little skill, with RMSE significantly higher than the benchmark RMSE 12.93 and with low pattern correlation. The EnKF-CAI however does slightly beat the benchmark RMSE and displays moderate pattern correlation. In Section 7.2, we show that the performance of EnKF-CAI is sharply improved by a simple tuning of parameters.

We now briefly describe how the thresholds and benchmarks were calculated. The equilibrium mean on each mode is $\bar{x}_i \approx 3.1$ and the variance is $\mathrm{var}(x_i) \approx 40.6$. Using Kalman one time assimilation, the posterior variance of x_1 will be reduced to 0.01, while

Filter	EnKF	EnKF-AI	EnKF-CI	EnKF-CAI
Cata. Div.	100%	0%	18%	0%
RMSE	NaN	24.48	NaN	11.91
Pattern Cor.	NaN	0.23	NaN	0.69

TABLE 7.3. *In the highly turbulence regime (F = 16), for the four algorithms we compare frequency of catastrophic filter divergence (Cata. Div.), RMSE, and pattern correlation (Pattern Cor.) The benchmark RMSE is 12.93.*
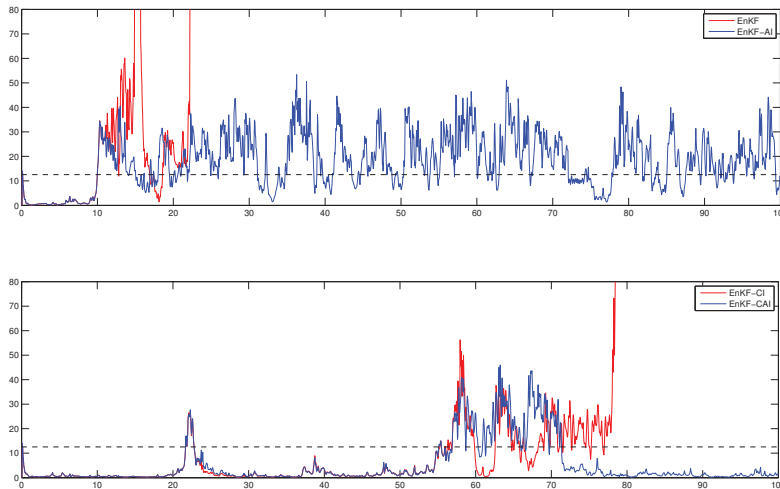

FIG. 7.3. *We compare the posterior error for one realization of the four filters in the strongly turbulent (F = 16) regime. The filter is very unstable, EnKF (red, upper panel) explodes to machine infinity, and EnKF-CI (red, lower panel) does the same with significant probability. The adaptive inflation mechanism, when applied to these two filters, is triggered frequently, preventing the ensemble from exploding. The performance of EnKF-AI (blue, upper panel) is worse than the benchmark (black dash). The performance of EnKF-CAI (blue, lower panel), is much better and outperforms the benchmark.*

the benchmark RMSE will be 12.93. From these statistics, we compute $M_1 = \sigma_\Theta = 127.6$ and $M_2 = M_\Xi = 81.4$.

7.2. Dependence on constant inflation strength and observation time.
Based on previous discussions, the performance of EnKF-CI and EnKF-CAI are satisfactory in the weakly and moderately turbulent regimes, but the strongly turbulent regime poses a significant challenge. In this section, we study how constant covariance inflation strength ρ and observation time h affect the performance of EnKF-CI and EnKF-CAI in the $F = 16$ regime.

In Table 7.4 we compare the performance of EnKF-CAI for different choices of constant inflation strength ρ, including RMSE and pattern correlation. We also include the frequency of catastrophic filter divergence in EnKF-CI to emphasize the impact of adaptive inflation.

As constant inflation strength is decreased, the frequency of catastrophic filter divergence increases. This agrees with the notion that removing inflation simply results

in EnKF, which in Table 7.3 was shown to experience catastophic filter divergence in 100% of trials.

In terms of accuracy, the EnKF-CAI can be optimized by choosing a moderately small constant inflation strength. In particular, when ρ is in the 0.01–0.02 range the RMSE is significantly smaller than the benchmark RMSE 12.93 and the pattern correlation is relatively high. Surprisingly, the optimal choice of ρ for EnKF-CAI is very sub-optimal for EnKF-CI, with around half of all trials experiencing catastrophic filter divergence. This emphasizes the importance of creating the right balance between adaptive inflation and constant inflation.

ρ	1	0.5	0.2	0.1	0.05	0.02	0.01	0.005
EnKF-CI Cata. Div.	3%	8%	18%	18%	28%	42%	57%	75%
EnKF-CAI RMSE	13.05	13.62	13.43	11.91	8.82	8.51	9.3	10.51
EnKF-CAI Cor.	0.64	0.65	0.66	0.69	0.70	0.70	0.75	0.70

TABLE 7.4. *In the $F = 16$ regime we display for different choice of constant inflation strength ρ, the frequency of catastrophic filter divergence for EnKF-CI (EnKF-CI Cata. Div.), the RMSE for EnKF-CAI and the pattern correlation for EnKF-CAI (EnKF-CAI Cor.). Each statistic is generated from 100 independent trials. The benchmark RMSE is 12.93.*

In Table 7.5 we display the results of a similar experiment, now varying the observation interval h keeping a fixed inflation strength $\rho = 0.1$. Again we display frequency of catastrophic filter divergence for EnKF-CI and the and pattern correlation RMSE. Each statistic is generated from 100 independent trials.

We found that the frequnecy of catastophic filter divergence is not monotonic in h, but rather displays a peak at $h = 0.1$. This agrees with the findings of [10]. We also find that the accuracy of EnKF-CAI is not monotonic in h but rather, there is an optimal choice around $h = 0.1$. As with the inflation strength experiment, the optimal parameter choice for the accuracy of EnKF-CAI occurs when EnKF-CI is exhibiting frequent catastrophic filter divergence. The fact that filter performance decreased with decreasing h can be attributed to a violation of filter observability conditions when the time-step becomes too small [4].

h	0.01	0.02	0.05	0.1	0.2	0.5
EnKF-CI Cata. Div.	0%	1%	18%	25%	5%	0%
EnKF-CAI RMSE	25.75	20.71	11.91	6.43	14.09	14.80
EnKF-CAI Cor.	0.31	0.37	0.69	0.64	0.50	0.36

TABLE 7.5. *In the $F = 16$ regime we display for different choice of observation interval h, the frequency of catastrophic filter divergence for EnKF-CI (EnKF-CI Cata. Div.), the RMSE for EnKF-CAI and the pattern correlation for EnKF-CAI (EnKF-CAI Cor.). Each statistic is generated from 100 independent trials. The benchmark RMSE is 12.93.*

7.3. Dependence on the numerical integrator.   In this section, we address the question of whether the catastrophic filter divergence can be avoided by simply using a more stable integrator in EnKF and EnKF-CI. We will see that with implicit methods and variable time step methods, one can avoid catastrophic filter divergence, but at prohibitive computational cost. Moreover, although they avoid catastrophic filter divergence, they do not avoid ordinary filter divergence; in particular we will see that EnKF and EnKF-CI with implicit / variable time-steps do not exhibit any

filtering skill, in contrast with EnKF-AI and EnKF-CAI. This corroborates the findings of [9, Chapter 2], where it was shown that numerical model error from implicit methods leads to reduced filtering skill.

In Table 7.6, we display the results from 100 trials, comparing the performance of EnKF-AI with explicit Euler (step size $\Delta = 10^{-4}$) against EnKF with (respectively) explicit Euler ($\Delta = 10^{-4}$), standard 4th order Runge–Kutta ($\Delta = 2.5 \times 10^{-3}$), `ode45`, and implicit Euler ($\Delta = 10^{-2}$). `ode45` is a native MATLAB integrator, which is based on a Runge-Kutta (4,5) method with variable step size. The implicit equation in implicit Euler is solved using the native MATLAB function `fslove`. The average time is also shown to indicate the computational expense associated with each method.

We see that `ode45` and implicit Euler are the only methods that do not exhibit catastrophic filter divergence, but each trial is enormously expensive when compared to the explicit Euler based methods. Moreover, `ode45` and Implicit Euler display zero filtering skill, as judged by the RMSE and pattern correlation (the benchmark RMSE is 12.93). As discussed earlier, EnKF-AI also exhibits little skill in this regime, but we shall see that the skill is dramatically improved by introducing additional constant inflation. Unsurprisingly, the explicit Euler based methods are much faster than both the higher order and the implicit methods and the EnKF-AI method is only fractionally more expensive than EnKF.

	EnKF-AI	EnKF			
Integrator	Explicit	Explicit	RK4	`ode45`	Implicit
Cata. Div.	0%	100%	92%	0%	0%
RMSE	24.48	NaN	NaN	44.66	20.38
Pattern Cor.	0.23	NaN	NaN	0.22	0.11
Avg. Time	2.49	2.39	19.53	41.00	662.48

TABLE 7.6. *The performance and time cost of EnKF-AI (explicit Euler, $\Delta = 10^{-4}$) against EnKF in the $F = 16$ regime. EnKF is implemented by explicit Euler ($\Delta = 10^{-4}$), RK4 ($\Delta = 2.5 \times 10^{-3}$), `ode45`, and implicit Euler ($\Delta = 10^{-2}$). The benchmark RMSE is 12.93. The time cost is measured in seconds.*

In Table 7.7 we display the results from a similar experiment, now comparing EnKF-CAI (with explicit Euler) against EnKF-CI with the same four choices of integrator. Catastrophic filter divergence is overall less prevalent, with `ode45` and implicit Euler once again exhibiting no catastrophic filter divergence. In contrast with the previous experiment, EnKF-CAI exhibits good filtering skill as judged by the RMSE and pattern correlation (once again, the benchmark RMSE is 12.93). EnKF-CI with `ode45` exhibits less filtering skill than EnKF-CAI, with RMSE not beating the benchmark and EnKF-CI with implicit Euler exhibits even less skill. As in the first experiment, the adaptive inflation is more accurate and far cheaper than the higher order and implicit methods.

7.4. The distribution of Θ_n and Ξ_n. The filter ensemble statistics Θ_n and Ξ_n, defined in Section 3 are used to determine when the adaptive inflation of EnKF-AI should be triggered. In particular, the mechanism is triggered whenever either of the two statistics passes an appropriately chosen threshold. In this section we look at the distribution of these statistics in order to evaluate the choice of thresholds introduced in Section 6. We have collected the data of Θ_n and Ξ_n from the EnKF-AI scheme from all the analysis steps. In Table 7.8 we compare the distributions of Θ_n and Ξ_n in the three different turbulent regimes and compare with the choice of threshold.

	EnKF-CAI	EnKF-CI			
Integrator	Explicit	Explicit	RK4	ode45	Implicit
Cata. Div.	0%	18%	5%	0%	0%
RMSE	11.91	NaN	NaN	14.95	16.29
Pattern Cor.	0.69	NaN	NaN	0.71	0.42
Avg. Time	2.65	2.34	21.48	37.02	650.31

TABLE 7.7. *The performance and time cost of EnKF-CAI (explicit Euler, $\Delta = 10^{-4}$)) against EnKF-CI in the $F = 16$ regime. EnKF is implemented by explicit Euler ($\Delta = 10^{-4}$), RK4 ($\Delta = 2.5 \times 10^{-3}$), ode45, and implicit Euler ($\Delta = 10^{-2}$). The benchmark RMSE is 12.93. The time cost is measured in seconds.*

F	4	8	16
M_1	32.5	69.6	127.6
Average Θ_n	2.95	20.28	70
$\mathbf{P}(\Theta_n > M_1)$	0.0285%	3.025%	9.73%
M_2	6.2	28.8	81.4
Average Ξ_n	0.001	0.0437	0.22
$\mathbf{P}(\Xi_n > M_2)$	0%	0%	0%

TABLE 7.8. *For the three turbulent regimes ($F = 4, 8, 16$), we list the thresholds M_1 and M_2 chosen according to the method of Section 6. We also list the average of Θ_n and Ξ_n over 100 independent trials and also give the probability that the statistics pass their respective threshold.*

It is clear from the table that stronger turbulence significantly increase the size of both Θ_n and Ξ_n. The thresholds M_1 and M_2 are always larger than the average of Θ_n and Ξ_n, but the ratios $M_1/\langle\Theta_n\rangle$ and $M_2/\langle\Xi_n\rangle$ significantly decreases as the turbulence gets stronger. As a result, the percentage of Θ_n and Ξ_n that are beyond the thresholds significantly increases. Connecting this observation with the performance of EnKF-AI, which is good in $F = 4$, slightly off the benchmark in $F = 8$, and very bad in $F = 16$, we can conclude that whether Θ_n and Ξ_n are beyond the thresholds is a good indicator of the malfunctioning. As intended, the adaptive inflation is rarely triggered when the filter operates properly, but frequently triggered when the situation is chaotic.

The histograms of Θ_n and Ξ_n in three regimes are presented in Figure 7.4. Note that Θ_n has a Gaussian like tail in the weak turbulence regime and exponential like in strong turbulence. The tail of Ξ_n is much heavier and polynomial-like. The thresholds we compute from the benchmarks are fairly large comparing with the distribution. These two posterior observations suggest that using even an aggressive threshold strategy as 6.4, the outliers are relatively scarce so adaptive inflation does not hamper the performance of the filters.

It is worth noting that Ξ_n is usually a small number and very rarely exceeds the threshold M_2. This suggests that the filter performance would be unaffected if the adaptive inflation was changed to $\lambda_n = c_\varphi \Theta_n$ so that only Θ_n is used to trigger the inflation. We have tested this simpler adaptive inflation mechanism with both EnKF-AI and EnKF-CAI in the setting of the previous subsections, and the performances are much as the same. From a theoretical perspective however, both statistics must be included to obtain a rigorously stable filter.
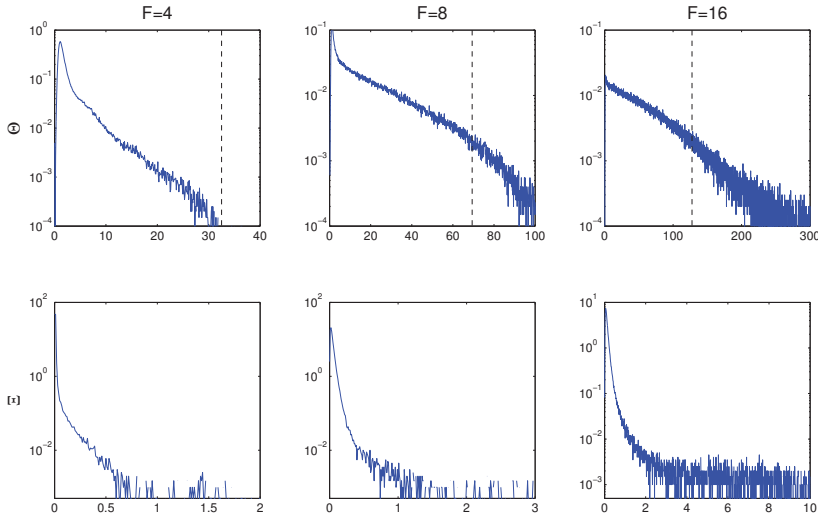
FIG. 7.4. *Histograms of Θ_n (upper panel) and Ξ_n (lower panel) in $F = 4, 8, 16$ regimes. The thresholds M_1 are marked as black dashes. The thresholds M_2 lie out side the x-axis for Ξ_n. Note the logarithmic scale on the vertical axis.*

7.5. Discussion of experiments.   In Section 7.1, the numerical simulations convincingly demonstrate that catastrophic filter divergence is prevalent in EnKF and EnKF-CI, particularly when using an unstable integrator in the forecast step. By comparing with EnKF-AI and EnKF-CAI, adaptive inflation is shown to successfully avoid catastrophic filter divergence, without sacrificing any filtering skill. In the weakly turbulent regime, the adaptive inflation mechanism is rarely turned on and EnKF-AI performs similarly to EnKF and likewise for the constant inflation counterparts. In more turbulent regimes, the adaptive inflation is regularly triggered and the adaptive inflation filters display filtering skill, typically beating the RMSE benchmark. In Section 7.2, it is also shown that this skill can be optimized by tuning constant inflation strength.

We saw in Section 7.3 that catastrophic filter divergence can be avoided by using implicit or variable time step integrators, but at the expense of prohibitive computational cost and the loss of filtering skill. Indeed, one would be better off using the naive estimator developed for the RMSE benchmark.

The adaptive inflation method permits for the effective use of very cheap integrators such as explicit Euler in the forecast step. As evidenced by EnKF, it is impossible to use such integrators in turbulent models, due to prevalence of catastrophic filter divergence. The adaptive inflation completely avoids this issue.

In Section 7.4, we provide the distribution analysis of Θ_n and Ξ_n which indicates the thresholds are achieving our purpose. We also saw that the threshold for Ξ_n was never triggered, indicating that a simpler adaptive inflation only using the statistic Θ_n will have practically the same performance. This has been validated, but is not presented here.

8. Conclusion and discussion
In this article we have proposed a novel modification of ensemble based methods using an adaptive covariance inflation. This modified method resolves common issues

with ensemble based methods, which were raised in the introduction, namely that there is a lack of understanding of the stability properties of EnKF, with catastrophic filter divergence drawing stark attention to this fact. The EnKF-AI and related filtering methods (ETKF-AI, EAKF-AI, EnKF-CAI, etc.) resolve this stability issue entirely.

In particular we can develop a complete stability theory for these filters. In Section 4, we have proved time-uniform mean-square upper bounds for EnKF-AI and related filters, using a Lyapunov argument. This is a considerable improvement on the known results for EnKF [7] since the observable energy criterion is no longer required and the resulting Lyapunov function for the filter always has compact sub-level sets. In Section 5, we exploit this fact to prove geometric ergodicity of the signal-ensemble process for EnKF-AI and related methods. The ergodicity result is a type of stability which ensures that filter initialization errors are dissipated exponentially quickly by the method.

In Section 7, we provide numerical evidence which supports the stability theory developed for EnKF-AI and related methods. In particular, catastrophic filter divergence is completely eliminated by adaptive inflation, even when cheap unstable integrators are used in the forecast step. For such integrators, catastrophic filter divergence is unavoidable in EnKF and EnFK-CI. Moreover, adaptive inflation does not sacrifice any filtering skill, with EnKF-AI and EnKF-CAI displaying good accuracy when compared to benchmark accuracy scores, even in highly turbulent regimes. This suggests that adaptive inflation may in fact make filtering algorithms faster, since they allow for such cheap, typically unstable forecast integrators.

In recent years, there have been many methods introduced featuring a type of adaptive inflation filtering. These methods are motivated by very different principles. For example, [12, 13] introduce a method that uses the norm of innovation to filter the optimal parameter for multiplicative inflation. Two different ways to find the covariances of the system and observation noises, assuming they are not known, are introduced in [14, 15]. A covariance relaxation method, so the posterior ensemble is pulled back to the forecast ensemble with a linear factor obtained from the innovation process is designed in [16]. In [30] a modification of 3DVAR is proposed which relies on projection back to a stable region, this modification has been theoretically analyzed from the perspective of long-time accuracy. On one hand, these methods are quite different from the EnKF-AI method we introduced. With the exception of [30], none have been rigorously proved to be stable in the sense of our theorems here. On the other hand, it is very possible that our theoretical framework here can be extended to these methods to build a stability framework, at least in some simple settings. The compelling reason here is that all these methods use the innovation sequence to decide the inflation, relaxation or noise covariance, which all grow with respect to Θ_n. From this perspective, our framework here offers an explanation of the good performance of these data assimilation methods, as they are more resilient against filter divergence; and when the filter is on the edge of malfunctioning, the adaptive mechanisms can pull the ensemble back to more reasonable states.

Appendix A. Elementary claims.

LEMMA A.1. *By Young's inequality, for any $\epsilon > 0$ and $x, y \in \mathbb{R}^d$, the following holds:*

$$|x+y|^2 = |x|^2 + |y|^2 + 2\langle x, y \rangle \le (1+\epsilon^2)|x|^2 + (1+\epsilon^{-2})|y|^2.$$

LEMMA A.2. *For any group of vectors $v_1, \dots, v_n$, the following holds:*

$$\arg\min_v \sum_{k=1}^n |v_i - v|^2 = \frac{1}{n} \sum_{k=1}^n v_i.$$

Proof. One can verify this by using the KKT condition. □

Appendix B. Stability for ETKF-AI, EAKF-AI.

Proof. (Proof of Theorem 4.5.)   We use the observed-unobserved notation introduced in Section 2.3. Notice that the mean update of (3.5) is the ensemble averaged version of (3.1) without the artificial perturbations $\xi_n^{(k)}$. Applying Jensen's inequality

$$\mathbb{E}_{n-1}|\overline{X}_n|^2 \leq \frac{1}{K} \sum_{k=1}^K \mathbb{E}_{n-1}|X_n^{(k)}|^2$$

to (4.3), we find the following with a constant D_1,

$$\mathbb{E}_{n-1}|\overline{X}_n|^2 - 2\rho_0^{-1}\|H_0\|^2 \mathbb{E}_{n-1}|HU_n|^2 \leq D_1.$$

The corresponding part for (4.4) will become

$$|\overline{Y}_n|^2 \leq (1 + \frac{1}{2}\beta_h)|\overline{\widetilde{Y}}_n|^2 + (1 + 2\beta_h^{-1})|\widehat{B}_n^T H_0^T (I + H_0 \widetilde{C}_n^X H_0^T)^{-1}(H_0 \overline{\widehat{X}}_n - Z_n)|^2.$$

Applying Jensen's inequality to (4.5), we find the second part above is bounded by

$$|\widehat{B}_n^T H_0^T (I + H_0 \widetilde{C}_n^X H_0^T)^{-1}(H_0 \overline{\widehat{X}}_n - Z_n)|^2 \leq K\|H_0\|^2 \max\{M_1^2 M_2^2, \rho_0^{-2} c_\varphi^{-2}\}.$$

So there is a constant D_2 such that

$$\mathbb{E}_{n-1}|\overline{V}_n|^2 - 2\rho_0^{-1}\|H\|^2 \mathbb{E}_{n-1}|HU_n|^2 \leq (1 + \frac{1}{2}\beta_h)|\overline{\widehat{V}}_n|^2 + D_2,$$

where we use the fact that $\|H_0\|^2 = \|H\|^2$. Then notice that the posterior covariance follows $C_n \preceq \widehat{C}_n$ by (2.9). Therefore, using

$$\sum_{k=1}^K |V_n^{(k)}|^2 = K|\overline{V}_n|^2 + \mathrm{tr}(C_n) \leq K|\overline{V}_n|^2 + \mathrm{tr}(\widehat{C}_n),$$

and $\sum_{k=1}^K |\widehat{V}_n^{(k)}|^2 = K|\overline{\widehat{V}}_n|^2 + \mathrm{tr}(\widehat{C}_n)$, we obtain from the previous inequality

$$\mathbb{E}_{n-1} \sum_{k=1}^K |V_n^{(k)}|^2 - 2K\rho_0^{-1}\|H\|^2 \mathbb{E}_{n-1}|HU_n|^2 \leq (1 + \frac{1}{2}\beta_h)\mathbb{E}_{n-1} \sum_{k=1}^K |\widehat{V}_n^{(k)}|^2 + D_2.$$

Assumption 4.1 leads to

$$(1 + \frac{1}{2}\beta_h)\mathbb{E}_{n-1} \sum_{k=1}^K |\widehat{V}_n^{(k)}|^2 \leq (1 - \frac{1}{2}\beta_h) \sum_{k=1}^K |V_{n-1}^{(k)}|^2 + (1 + \frac{1}{2}\beta_h)KK_h,$$

and moreover

$$\mathbb{E}_{n-1}\sum_{k=1}^{K}|V_n^{(k)}|^2 - 2K\rho_0^{-1}\|H\|^2\mathbb{E}_{n-1}|HU_n|^2$$

$$\leq (1-\frac{1}{2}\beta_h)\sum_{k=1}^{K}|V_{n-1}^{(k)}|^2 + (1+\frac{1}{2}\beta_h)KK_h + D_2.$$

Then notice a multiple of Assumption 4.1 is

$$4K\beta_h^{-1}\rho_0^{-1}\|H\|^2\mathbb{E}_{n-1}|U_n|^2 \leq 4K\beta_h^{-1}\rho_0^{-1}\|H\|^2(1-\beta_h)|U_{n-1}|^2 + 4K\beta_h^{-1}\rho_0^{-1}\|H\|^2K_h.$$

The sum of the two previous two inequalities yields

$$\mathbb{E}_{n-1}\mathcal{E}_n = \sum_{k=1}^{K}\mathbb{E}_{n-1}|V_n^{(k)}|^2 + 4K\beta_h^{-1}\rho_0^{-1}(1-\frac{1}{2}\beta_h)\|H\|^2\mathbb{E}_{n-1}|U_n|^2$$

$$\leq (1-\frac{1}{2}\beta_h)\sum_{k=1}^{K}|V_{n-1}^{(k)}|^2 + 4K\beta_h^{-1}\rho_0^{-1}\|H\|^2(1-\beta_h)|U_{n-1}|^2 + D,$$

with $D := 4K\beta_h^{-1}\rho_0^{-1}\|H\|^2K_h + (1+\frac{1}{2}\beta_h)KK_h + D_2$. Then notice that $(1-\frac{1}{2}\beta_h)^2 \geq (1-\beta_h)$, therefore

$$\mathbb{E}_{n-1}\mathcal{E}_n \leq (1-\frac{1}{2}\beta_h)\mathcal{E}_{n-1} + D.$$

□

Appendix C. Ergodicity formulas. Here we present the concrete formulas that are used in Section 5 for ETKF-AI and EAKF-AI. We omit unneccessary details concerning the transform and adjustment matrices, which can be found in [2–4, 7]. In both ESRF methods, the Markov kernel $\Phi: \mathcal{X} \times \mathcal{B}(\mathcal{Y}) \to [0,1]$ with $\mathcal{E} := \mathbb{R}^d \times \mathbb{R}^{d \times K}$ and $\mathcal{Y} := \mathbb{R}^d \times \mathbb{R}^{d \times K} \times \mathbb{R}^q$ is described by

$$(U_{n-1}, V_{n-1}^{(1)}, \ldots, V_{n-1}^{(K)}) \mapsto (U_n, \widehat{V}_n^{(1)}, \ldots, \widehat{V}_n^{(K)}, Z_n).$$

The deterministic step is given by the map $\Gamma(U,V,Z) = (U, \Gamma^{(1)}, \ldots, \Gamma^{(K)})$ where

$$\Gamma^{(k)} = \overline{\widehat{V}} - \widetilde{C}H^T(I + H\widetilde{C}H^T)^{-1}(H\overline{\widehat{V}} - Z) + S^{(k)} \tag{C.1}$$

with $\overline{V} = \frac{1}{K}\sum_{k=1}^{K}\widehat{V}^{(k)}$ and $\widetilde{C} = \widehat{C} = \frac{1}{K-1}\sum_{k=1}^{K}(\widehat{V}^{(k)} - \overline{V}) \otimes (\widehat{V}^{(k)} - \overline{V})$ and $S^{(k)}$ is the k-th column of the updated spread matrix S. S is given by $\widehat{S}T(\widehat{S})$ in the ETKF method, and $A(\widehat{S})\widehat{S}$ in the EAKF method, where $\widehat{S}$ is the forecast spread matrix $\widehat{S} = (\widehat{V}^{(1)} - \overline{\widehat{V}}, \ldots, \widehat{V}^{(K)} - \overline{\widehat{V}})$.

With the ETKF method, there are a few way to define the transformation matrix $T(\widehat{S})$, one reasonable choice is taking the matrix square root

$$T(\widehat{S}) = \left(I_K + (K-1)^{-1}\widehat{S}^T H^T H\widehat{S}\right)^{-\frac{1}{2}} = \left(I_K - (K-1)^{-1}\widehat{S}^T H^T (I + H\widehat{C}H^T)^{-1}H\widehat{S}^T\right)^{\frac{1}{2}}.$$

As for the EAKF method, the construction of the adjustment matrix $A(\widehat{S})$ is slightly more complicated as the following

$$A(\widehat{S}) = Q\Lambda G^T(I + D)^{-1/2}\Lambda^\dagger Q^T.$$

Here $Q\Lambda R$ is the SVD decomposition of $\widehat{S}$ and $G^T D G$ is the diagonalization of

$$(K-1)^{-1}\Lambda^T Q^T H^T H^T Q\Lambda^T,$$

and † indicates pseudo inverse of a matrix.

And for the ETKF-AI and EAKF-AI, everything is the same as their counterpart without adaptive inflation, except that we use the following in (C.1)

$$\widetilde{C} = \frac{1}{K-1}\sum_{k=1}^{K}(\widehat{V}^{(k)} - \overline{\widehat{V}}) \otimes (\widehat{V}^{(k)} - \overline{\widehat{V}}) + c_\varphi\Theta_n(1+\Xi_n)\mathbb{1}_{\Theta_n > M_1 \text{ or } \Xi_n > M_2}I.$$

REFERENCES

[1] G. Evensen, *The ensemble Kalman filter: Theoretical formulation and practical implementation*, Ocean dynamics, 53(4), 343–367, 2003.

[2] 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(3), 420–436, 2001.

[3] J.L. Anderson, *An ensemble adjustment Kalman filter for data assimilation*, Mon. Weather Rev., 129(12), 2884–2903, 2001.

[4] A.J. Majda and J. Harlim, *Filtering Complex Turbulent Systems*, Cambridge University Press, Cambridge, UK, 2012.

[5] E. Kalnay, *Atmospheric Modeling, Data Assimilation, and Predictability*, Cambridge University Press, 2003.

[6] D.T.B. Kelly, K.J. Law, and A.M. Stuart, *Well-posedness and accuracy of the ensemble Kalman filter in discrete and continuous time*, Nonlinearity, 27, 2579–2603, 2014.

[7] X.T. Tong, A.J. Majda, and D. Kelly, *Nonlinear stability and ergodicity of ensemble based Kalman filters*, arXiv:1507.08307, 2015.

[8] C. Gonzales-Tokman and B.R. Hunt, *Ensemble data assimilation for hyperbolic systems*, Physica D, 116(1), 128–142, 2013.

[9] A.J. Majda and J. Harlim, *Catastrophic filter divergence in filtering nonlinear dissipative systems*, Commun. Math. Sci., 8, 27–43, 2008.

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

[11] D. Kelly, A.J. Majda, and X.T. Tong, *Concrete ensemble Kalman filters with rigorous catastrophic filter divergence*, Proc. Natl. Acad. Sci., 112(34), 10589–10594, 2015.

[12] J.L. Anderson, *An adaptive covariance inflation error correction algorithms for ensemble filters*, Tellus A, 59, 210–224, 2007.

[13] J.L. Anderson, *Spatially and temporally varing adaptive covariance inflation for ensemble filters*, Tellus A, 61A:72–83, 2009.

[14] T. Berry and T. Sauer, *Adaptive ensemble Kalman filtering of nonlinear systems*, Tellus A, 65(20331), 2013.

[15] Y. Zhen and J. Harlim, *Adaptive error covariance estimation methods for ensemble Kalman filters*, J. Comput. Phys., 294, 619–638, 2014.

[16] Y. Ying and F. Zhang, *An adaptive covariance relaxation method for ensemble data assimilation*, Quarterly Journal of the Royal Meteorological Society, 2015.

[17] J. Mandel, L. Cobb, and J.D. Beezley, *On the convergence of the ensemble Kalman filter*, Appl. Math., 56(6), 533–541, 2011.

[18] F. Le Gland, V. Monbet, and V.D. Tran, *Large Sample Asymptotics for the Ensemble Kalman Filter*, Oxford Handbook of Nonlinear Filtering, 2010.

[19] O.G. Ernst, B. Sprungk, and H.J. Starkloff, *Analysis of the ensemble and polynomial chaos kalman filters in bayesian inverse problems*, arXiv:1504.03529, 2015.

[20] G. Evensen, *Sampling strategies and square root analysis schemes for the EnKF*, Ocean Dynamics, 54(6), 539–560, 2004.

[21] 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, 2741–2758, 1999.

[22] T.M. Hamill, C. Whitaker, and C. Snyder, *Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter*, Mon. Weather Rev., 129, 2776–2790, 2001.

[23] R. Furrer and T. Bengtsson, *Estimation of hign-dimensional prior and posterior covariance matrices in Kalman filter variants*, J. Multi. Anal., 98, 227–255, 2007.

[24] A.H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1972.

[25] X.T. Tong and R. Van Handel, *Ergodicity and stability of the conditional distributions of nondegenerate markov chains*, Ann. Appl. Probab., 22(4), 1495–1540, 2012.

[26] S. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability*, Springer-Verlag, 1993.

[27] J.C. Mattingly and A.M. Stuart, *Geometric ergodicity of some hypo-elliptic diffustion for particle motions*, Markov Process. Related Fields, 8(2), 199–214, 2002.

[28] A.J. Majda, R.V. Abramov, and M.J. Grote, *Information Theory and Stochastics for Multiscale Nonlinear Systems*, American Mathematical Soc., 25, 2005.

[29] A.J. Majda and X. Wang, *Nonlinear Dynamics and Statistical Theories for Basic Geophysical Flows*, Cambridge University Press, Cambridge, UK, 2006.

[30] D. Sanz-Alonso and A.M. Stuart, *Long-time asymptotics of the filtering distribution for partially observed chaotic dynamical systems*, arXiv:1411.6510, 2014.