

# Concrete ensemble Kalman filters with rigorous catastrophic filter divergence

David Kelly<sup>a</sup>, Andrew J. Majda<sup>a,b,1</sup>, and Xin T. Tong<sup>a,b,1</sup>

<sup>a</sup>Courant Institute of Mathematical Sciences, New York University, New York, NY 10012; and <sup>b</sup>Center for Atmosphere Ocean Science, New York University, New York, NY 10012

Contributed by Andrew J. Majda, June 15, 2015 (sent for review May 12, 2015; reviewed by Jeffrey Anderson and Sebastian Reich)

The ensemble Kalman filter and ensemble square root filters are data assimilation methods used to combine high-dimensional, nonlinear dynamical models with observed data. Ensemble methods are indispensable tools in science and engineering and have enjoyed great success in geophysical sciences, because they allow for computationally cheap low-ensemble-state approximation for extremely high-dimensional turbulent forecast models. From a theoretical perspective, the dynamical properties of these methods are poorly understood. One of the central mysteries is the numerical phenomenon known as catastrophic filter divergence, whereby ensemble-state estimates explode to machine infinity, despite the true state remaining in a bounded region. In this article we provide a breakthrough insight into the phenomenon, by introducing a simple and natural forecast model that transparently exhibits catastrophic filter divergence under all ensemble methods and a large set of initializations. For this model, catastrophic filter divergence is not an artifact of numerical instability, but rather a true dynamical property of the filter. The divergence is not only validated numerically but also proven rigorously. The model cleanly illustrates mechanisms that give rise to catastrophic divergence and confirms intuitive accounts of the phenomena given in past literature.

data assimilation | ensemble Kalman filter | filter divergence

With the growing importance of accurate weather forecasting and the expanding availability of geophysical measurements, data assimilation has never been more vital to society. Ensemble-based assimilation methods, including the ensemble Kalman filter (EnKF) (1) and ensemble square root filters (ESRF) (2, 3), are crucial components of data assimilation that are applied ubiquitously across the geophysical sciences (4, 5). Despite their widespread application, the theoretical understanding of these methods remains underdeveloped. Recent efforts have been made to understand the dynamical properties of EnKF/ESRF in the practical setting of high-dimensional turbulent forecast models with low ensemble size, focusing on well-posedness (6) and stability.

One of the main motivations for these theoretical studies was the curious numerical phenomenon known as catastrophic filter divergence (7, 8). In refs. 7 and 8, it was numerically demonstrated that state estimates provided by ensemble-based methods can explode to machine infinity, despite the forecast model's being dissipative and satisfying the absorbing ball property (9), which demands that the true state be always absorbed by a bounded region of the state space. The genesis of this phenomenon has previously been attributed to an interplay between stiffness in the forecast models and forecast ensemble alignment. Nevertheless, until now there has been no concrete example that transparently illustrates the mechanism behind catastrophic filter divergence. Without an explicit example or concrete understanding of the phenomenon, it is difficult to identify which models are vulnerable to it and how we should modify the ensemble methods to overcome it.

In this article, we bridge the gap by providing an extremely simple forecast model that has the absorbing ball property but regardless exhibits catastrophic filter divergence under both EnKF and ESRF methods. The mechanism behind the divergence is straightforward and is explained with a simple diagram (Fig. 1).

The divergence is not a pathological example, but rather occurs for a large set of filter initializations and parameter values. We not only demonstrate the blowup numerically, we also prove it rigorously in *Theorem 2*. The mechanism behind the blowup validates the intuition from ref. 8 and involves an interplay between model stiffness and ensemble alignment.

When encountered in practical models, catastrophic filter divergence is often attributed to numerical instabilities. Such numerical instabilities are common in practical geophysical simulations, where Courant–Friedrichs–Lewy conditions are routinely violated (10). We emphasize that for our model catastrophic filter divergence is not a numerical artifact but rather a genuine and mathematically verifiable property of the filters.

## EnKF and ESRF Formulation

We now briefly describe the EnKF and ESRF algorithms. Let  $\Psi$  be the forecast model, with the true model state satisfying  $U_n = \Psi(U_{n-1})$  for all integers  $n \geq 1$  and with some given (possibly random) initial state  $U_0$ . At each time step we make an observation  $Z_n = HU_n + \xi_n$ , where  $H$  is the observation matrix and  $\xi_n$  are independent and identically distributed (iid) random variables distributed as  $\mathcal{N}(0, \Gamma)$  and for simplicity we take  $\Gamma = I$ . The objective of data assimilation is to use the forecast model to combine an estimate of the previous state  $U_{n-1}$  with the new observational data  $Z_n$  to produce an estimate of the current state  $U_n$ .

In both EnKF and ESRF algorithms, an ensemble  $\{V_n^{(k)}\}_{k=1}^K$  is used as an empirical estimate of the posterior distribution of the state  $U_n$  given the history of observations  $Z_1, \dots, Z_n$ . The empirical

## Significance

Understanding ensemble-based data assimilation methods, including their performance when applied to high-dimensional nonlinear models with low ensemble size, is a crucial problem in science and engineering. Catastrophic filter divergence is a well-documented but mechanistically mysterious phenomenon whereby ensemble-state estimates explode to machine infinity despite the true state remaining in a bounded region. We provide breakthrough insight into the phenomenon by proposing a simple forecast model that experiences catastrophic filter divergence under all ensemble-based methods. This is the first instance to our knowledge of a forecast model that plainly and rigorously illustrates that simple mechanisms can lead to such a drastic filter malfunction and thereby sheds light on when catastrophic filter divergence should be expected and how it can be avoided.

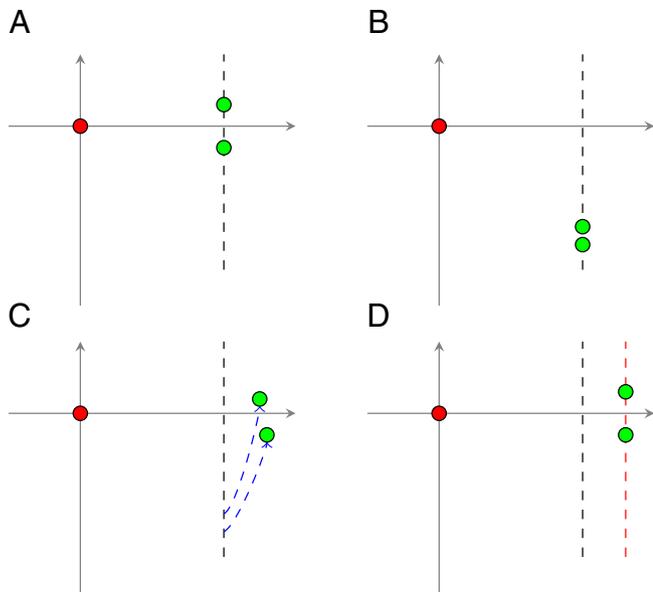
Author contributions: D.K., A.J.M., and X.T.T. designed research, performed research, contributed new reagents/analytic tools, analyzed data, and wrote the paper.

Reviewers: J.A., National Center for Atmospheric Research; and S.R., Universität Potsdam. The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

<sup>1</sup>To whom correspondence may be addressed. Email: jonjon@cims.nyu.edu or tong@cims.nyu.edu.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1511063112/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1511063112/-DCSupplemental).



**Fig. 1.** (A) The forecast ensemble (green circles) and the truth (red circle). (B) In the analysis step, the posterior ensemble remains in the subspace spanned by the forecast ensemble. Owing to the observation matrix  $H$ , all observations will shift the ensemble downward by an amount (approximately) proportional to the  $x$  coordinate of the forecast ensemble. (C) The map  $\Psi_\theta$  is applied and the flow lines of  $\Psi_\theta$  are shown (blue dashed). (D) The map  $\Psi_{\text{lock}}$  is applied, which locks the ensemble into new alignment (red dashed line). The ensemble has returned to the aligned state of A, but with larger  $x$  value. The increase in  $x$  is proportional to the original  $x$  value. Catastrophic filter divergence occurs when this cycle repeats itself.

mean of the ensemble is a useful estimate of the state of the model and the empirical covariance matrix provides a quantification of uncertainty.

The EnKF algorithm is the iteration of two steps, the forecast step and the analysis step. In the forecast step, the time  $n-1$  posterior ensemble  $\{V_{n-1}^{(k)}\}_{k=1}^K$  is evolved to the forecast ensemble  $\{\hat{V}_n^{(k)}\}_{k=1}^K$ , where  $\hat{V}_n^{(k)} = \Psi(V_{n-1}^{(k)})$ . The primary function of the forecast ensemble is to represent uncertainty in the forecast model; this uncertainty is quantified via the empirical covariance matrix  $\hat{C}_n$ :

$$\hat{C}_n = \frac{1}{K-1} \sum_{k=1}^K (\hat{V}_n^{(k)} - \bar{V}_n) \otimes (\hat{V}_n^{(k)} - \bar{V}_n), \quad [1]$$

where  $\bar{V}_n = K^{-1} \sum_{k=1}^K \hat{V}_n^{(k)}$ . In the analysis step, the time  $n$  observation  $Z_n$  is assimilated with the forecast ensemble to produce the posterior ensemble  $\{V_n^{(k)}\}_{k=1}^K$ . The assimilation update of each ensemble member is described by a Kalman-type update in a possibly nonlinear setting. The update uses a perturbed observation  $Z_n^{(k)} = Z_n + \xi_n^{(k)}$ , where  $\xi_n^{(k)}$  is an iid sequence distributed identically to  $\xi_n$ . The Kalman update is then given by

$$V_n^{(k)} = \hat{V}_n^{(k)} - \hat{C}_n H^T (I + H \hat{C}_n H^T)^{-1} (H \hat{V}_n^{(k)} - Z_n^{(k)}). \quad [2]$$

The noise perturbations  $\xi_n^{(k)}$  are introduced to maintain the classical Kalman prior–posterior covariance relation

$$C_n = \hat{C}_n - \hat{C}_n H^T (I + H \hat{C}_n H^T)^{-1} H \hat{C}_n \quad [3]$$

in an average sense, where  $C_n$  denotes the sample covariance of the posterior ensemble.

The ESRF algorithms considered in this article are the ensemble transform Kalman filter (ETKF) (2) and the ensemble adjustment Kalman filter (EAKF) (3). Both filters use the same forecast step as EnKF but differ from EnKF (and from each other) in the analysis step. In both ETKF and EAKF, the posterior ensemble mean  $\bar{V}_n$  is updated from the forecast mean  $\bar{V}_n$  via

$$\bar{V}_n = \bar{V}_n - \hat{C}_n H^T (I + H \hat{C}_n H^T)^{-1} (H \bar{V}_n - Z_n). \quad [4]$$

Given the updated mean, to compute the update for each ensemble member it is sufficient to compute the posterior ensemble spread matrix  $S_n = [V_n^{(1)} - \bar{V}_n, \dots, V_n^{(K)} - \bar{V}_n]$ . This is computed using the similarly defined forecast spread matrix  $\hat{S}_n$ . Any reasonable choice of  $S_n$  should satisfy the Kalman covariance identity

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

where, by definition,  $C_n = S_n S_n^T$ . The ETKF algorithm achieves Eq. 3 by setting  $S_n = \hat{S}_n T_n$  where  $T_n$  is the transform matrix defined by

$$T_n = (I + (K-1)^{-1} \hat{S}_n^T H^T H \hat{S}_n)^{-1/2}.$$

The EAKF algorithm achieves Eq. 3 by setting  $S_n = A_n \hat{S}_n$ , where  $A_n$  is the adjustment matrix defined by

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

Here  $Q \Lambda R$  is the singular value decomposition (SVD) of  $\hat{S}_n$  and  $G^T D G$  is the diagonalization of  $(K-1)^{-1} \Lambda^\dagger Q^T H^T H Q \Lambda^\dagger$ , and  $\dagger$  indicates pseudo inverse of a matrix. For more details on EnKF and ESRF see ref. 4.

From this point on, for our specific model-filter setup we will assume  $K$  to be even and divide the ensemble members into two groups  $\{V_n^{+, (k)}\}_{k=1}^{K_0}$  and  $\{V_n^{-, (k)}\}_{k=1}^{K_0}$ , where  $K = 2K_0$ . The difference between the two groups is only in how the ensemble members are initialized. As will become apparent later, this division serves as a convenient way of extending a  $K = 2$  catastrophic divergence result to an arbitrary even  $K$ .

### Emergence of Catastrophic Filter Divergence

When filtering a dissipative nonlinear system with EnKF or ESRF, the root of catastrophic filter divergence must be in the analysis step. The dissipation guarantees that the forecast step will reduce the energy of the ensemble, so the only opportunity to increase the energy is in the analysis step. In fact, the Kalman update formulas, Eqs. 2 and 4, have no general mechanisms that preserve the energy, and we will see that alignment of forecast ensemble members can lead to large increases of energy in the analysis step.

In refs. 7 and 8, the authors argue that catastrophic filter divergence is strongly associated with alignment of forecast ensemble members. Specifically, in ref. 8 the forecast model is a five-dimensional Lorenz-96 model with one observed variable. Alignment of the forecast ensemble is caused by stiffness in the forecast ordinary differential equation, as evidenced by a strongly negative Lyapunov exponent. When the ensemble aligns in a subspace perpendicular to the observed direction, the analysis update can shift the ensemble (within the subspace) to points that lie on higher energy trajectories of the forecast model. Because the observation is perpendicular to the ensemble subspace, the observation does not contain any useful information and hence

the analysis update can potentially affect the performance of the filter. In the next forecast step, the stiffness of the forecast leads to realignment of the ensemble, but on much higher energy trajectories than in the previous forecast step.

In the next section, we construct a simple and concrete forecast model that exhibits the energy growth described above. By an appropriate choice of parameters in the model we can further show catastrophic filter divergence rigorously exists.

### A Forecast and Observational Model

The forecast model  $\Psi: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is defined as the composition of two maps, a linear rotation map  $\Psi_\theta$  and a “locking map”  $\Psi_{\text{lock}}$ . Hence, from this point on we will refer to  $\Psi$  as the rotation-lock map. The map  $\Psi_\theta$  is simply the linear transformation

$$\Psi_\theta(x, y) = \begin{pmatrix} \rho \cos \theta & -\rho \sin \theta \\ \rho \sin \theta & \rho \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

for some angle  $\theta \in [0, 2\pi)$  and with  $\rho \in (0, 1)$ . Hence, the dynamics described by  $\Psi_\theta$  is anticlockwise rotation by angle  $\theta$  with attraction toward the origin, with contraction parameter  $\rho$ . The map  $\Psi_\theta$  defines the macroscopic dynamics of  $\Psi$  and moreover ensures the dissipation of energy for  $\Psi$ . The locking map  $\Psi_{\text{lock}}$  is defined by mapping the input vector to the nearest member of the grid

$$\mathcal{G} = \{(m, (2n + 1)\varepsilon) \in \mathbb{R}^2 : m, n \in \mathbb{Z}\}.$$

To be precise, we have

$$\Psi_{\text{lock}}(x, y) = (r(x), 2\varepsilon r((y + \varepsilon)/2\varepsilon) + \varepsilon),$$

where  $r$  is the function that rounds off a real number to the nearest integer, with the “tie-breaker” constraint  $r(n + 1/2) = n + 1/2$  for any integer  $n$ . The map  $\Psi_{\text{lock}}$  refines the microscopic structure of  $\Psi$ ; it generates the stiffness that forces the ensemble to align perfectly in a nontrivial subspace of the state space. In this 2D setting, we will see that the nontrivial subspace is simply a vertical line.

The forecast model is then given by

$$\Psi(x, y) = \Psi_{\text{lock}}(\Psi_\theta(x, y)), \quad [5]$$

hence we first apply the rotation and then lock the resulting vector to the grid  $\mathcal{G}$ . The model is given by

$$U_n = \Psi(U_{n-1})$$

We will typically choose  $U_0 = 0$  as initialization for our truth model, which, due to the definition of the round-off function  $r$ , is still a fixed point for the dynamics  $\Psi$ .

At large scales, the locking map makes very little difference to the qualitative and quantitative behavior of the forecast. Indeed, the following result shows that the forecast model still satisfies energy dissipation and consequently has the absorbing ball property (9). The proof is contained in [Supporting Information](#).

**Proposition 1.** *The nonlinear system generated by the rotation-lock map  $\Psi$  in 5 satisfies an energy principle:*

$$|\Psi(u)|^2 \leq \frac{1 + \rho^2}{2} |u|^2 + \frac{1 + \rho^2}{1 - \rho^2},$$

for all  $u \in \mathbb{R}^2$  and hence satisfies the absorbing ball property. Moreover, the origin is a fixed point of  $\Psi$ .

The observational model is of the standard form  $Z_n = HU_n + \xi_n$ , with the observation matrix

$$H = \begin{pmatrix} 1 & 0 \\ \varepsilon^{-2} & 1 \end{pmatrix}$$

and where the observation noise  $\xi_n$  are iid  $\mathcal{N}(0, I_2)$  random variables. Hence, the observations matrix is full rank, but when  $\varepsilon$  is small the condition number becomes quite large. In particular, the  $y$  observation becomes very poor when  $\varepsilon$  is small. When the forecast ensemble is aligned in the  $y$  direction with an erroneous  $x$  coordinate, the Kalman update with matrix  $H$  cannot correct this error but will instead magnify the error and transport it to the  $y$  coordinate. This is the catalyst for the exponential explosion of the EnKF/EAKF ensembles.

### The Energy Blow-Up Cycle

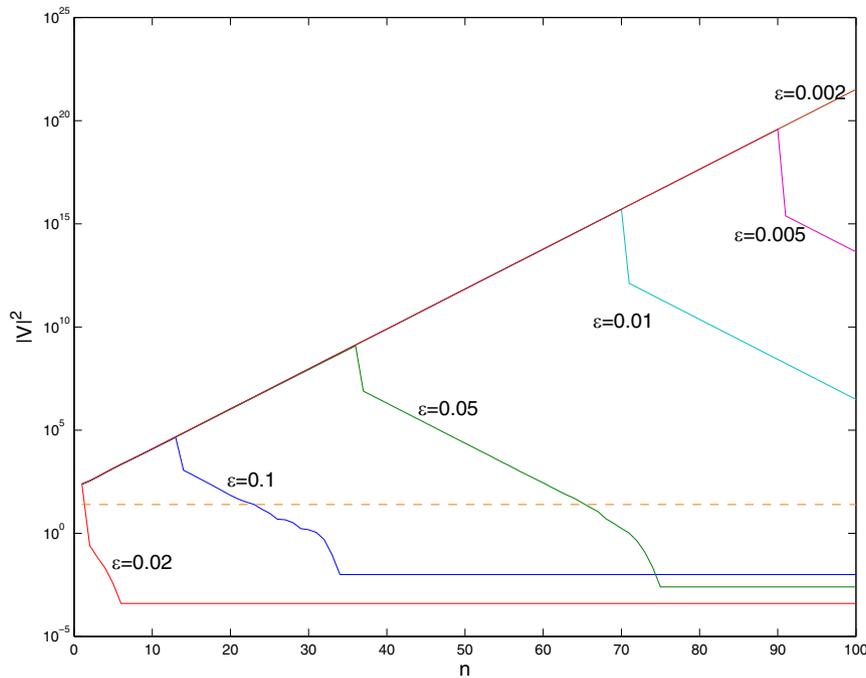
With the rotation-lock dynamics  $\Psi$  concretely defined, the mechanism behind filter divergence can now be explained graphically. In Fig. 1 we illustrate one cycle of this mechanism, each iteration of the cycle leads to an exponential increase in the ensemble energy. In this illustration there are only two ensemble members, hence we will simply denote them  $V_n^+$  and  $V_n^-$ . The same picture holds equally well for the case with arbitrary  $K = 2K_0$  members and will be explained at the end of the section. The truth is chosen to be the trivial solution  $U_n = 0$  for all times  $n$ .

In Fig. 1A we show the forecast ensemble (green circles) and the truth signal (red circle). The forecast ensemble members have respective coordinates  $\hat{V}_n^+ = (\hat{x}_n, \varepsilon)$  and  $\hat{V}_n^- = (\hat{x}_n, -\varepsilon)$ . Owing to the locking map  $\Psi_{\text{lock}}$  it is natural for the forecast ensemble to be aligned in such a way. In Fig. 1B, an observation  $Z_n$  is made and the analysis step is performed; hence, the green dots now represent the updated posterior ensemble members. Owing to the configuration of the ensemble and the choice of observation matrix, the actual location of the observation is not particularly relevant. In fact, the analysis step will always move the ensemble members downward within the space spanned by the ensemble (black dashed line) by an amount approximately proportional to  $\hat{x}_n$ . Indeed, from the analysis update (Eq. 2), we can explicitly calculate

$$V_n^\pm \approx (\hat{x}_n, \pm\varepsilon - 2\hat{x}_n / (1 + 2\varepsilon^2))$$

where we ignore  $O(\varepsilon^2)$  terms. The details of this computation are given in [Supporting Information](#) before Eq. S2 with  $K_0 = 1$ . In Fig. 1C, we perform the first half of the forecast map by applying the rotation  $\Psi_\theta$ . Here  $\rho, \theta$  must be chosen in such a way that  $V_n^+$  and  $V_n^-$  get rotated to the attracting region of some grid points  $(\hat{x}_{n+1}, \varepsilon)$ ,  $(\hat{x}_{n+1}, -\varepsilon)$ , respectively. In Fig. 1D, we complete the forecast step by applying  $\Psi_{\text{lock}}$ . Because the rotation leaves the ensemble members in the appropriate attracting regions, we will have  $\hat{V}_{n+1}^+ = (\hat{x}_{n+1}, \varepsilon)$  and  $\hat{V}_{n+1}^- = (\hat{x}_{n+1}, -\varepsilon)$ . Hence, we are in a position similar to that in Fig. 1A, but having increased the value of the  $x$  coordinate in both ensembles. Each time this cycle occurs, the energy of the ensemble increases exponentially. By tuning the algorithm, we can make this cycle repeat itself many times, resulting in catastrophic filter divergence.

In the case of more than two ensemble members,  $K_0 > 1$ , the picture is largely unchanged. In Fig. 1A the forecast ensemble members  $\{\hat{V}^{+, (k)}\}_{k=1}^{K_0}$  are clustered together at the upper green dot and  $\{\hat{V}^{-, (k)}\}_{k=1}^{K_0}$  at the lower green dot. As with the two member case, the locking mechanism makes this a natural forecast configuration. In Fig. 1B, when the analysis step is performed, the ensemble members are all shifted down along the ensemble subspace (black dashed line), but now the clusters break up due to the presence of additional perturbative noise.



**Fig. 2.** Exponential energy growth of one ensemble member of EnKF. The rotation-lock model  $\Psi$  is filtered using  $2K_0=6$  ensembles, with initialization  $\rho=0.8, M_0=10, R_0=0$ , and  $\varepsilon$  takes value from 0.1 to 0.002.

However, the model is tuned in such a way that the declustering will be undone by the locking map applied in Fig. 1D, and hence the ensemble returns to the same initial picture, but on higher energy trajectories.

**Mathematically Rigorous Divergence**

In this section, we will specify the precise choices of parameters in the model that guarantee catastrophic filter divergence and we will list the rigorous mathematical results. The proofs can be found in *Supporting Information*.

Before giving the precise statement of catastrophic filter divergence, we first explain the flavor of the result. Let  $N > 0$  be an arbitrary integer and let  $p \in (0,1)$ . We will show that there is an exponentially increasing sequence  $M_n > 0$ , such that with probability  $1-p$  the EnKF/ESRF ensemble  $\{V_n^{(k)}\}_{k=1}^K$  is bounded below by  $M_n$  for every  $n \leq N$ . In particular, by choosing parameters appropriately, we can ensure that the ensemble will diverge exponentially quickly for an arbitrary long time  $N$  and with arbitrarily high probability  $1-p$ .

We will parameterize the forecast model and EnKF/ESRF algorithm in the following way. First, we pick the angle  $\theta$  such that

$$\tan\theta = \frac{2K_0}{2K_0 - 1 + 2K_0\varepsilon^2}.$$

This choice of rotation ensures the ensemble transportation illustrated by Fig. 1C takes place. We also require that  $\rho$  be a rational number with  $\rho > 1/\sqrt{2}$  and that  $K_0 \leq 100$ , but  $K_0 \neq 2$ . The requirement of rationality is a technicality in the proof that is clearly of no consequence for numerical implementation. Both the rationality condition and the restrictions on  $K_0$  can be replaced with a much more general (but much more technical) condition. In particular, one can use any  $K_0$  and replace the rationality condition with a more specific  $K_0$ -dependent condition. For the sake of exposition we will defer the fully general treatment to *Supporting Information*.

We define the truth signal with  $U_0 = 0$  and hence  $U_n = 0$  for all  $n \geq 0$ . The posterior ensemble members will be initialized as

$$V_0^{\pm, (k)} = \left( M_0, -\tan\theta M_0 \pm \varepsilon + \varepsilon \tan\theta R_0^{\pm, (k)} \right) \quad [6]$$

with any  $M_0$  satisfying  $M_0(\sqrt{2}\rho - 1) > 1$  and with any  $R_0^{\pm, (k)} \in (-1,1)$ . This choice of initialization ensures the ensemble starts from Fig. 1C and has iterative exponential growth. Owing to the freedom in  $M_0$  and  $R_0$ , this initialization is not a pathological case, but rather a positive volume set. The exponentially increasing positive sequence  $M_n$  is defined by the recursion

$$M_n = r(\rho \sec\theta M_{n-1}).$$

With the above choice of  $\rho, \theta$  it is easy to check that  $M_n$  is indeed increasing exponentially (verified in *Supporting Information*). In the following,  $\mathbf{P}$  denotes the probability measure for the probability space on which the observational noise is defined.

**Theorem 2.** *Let  $\{V_n^{\pm, (k)}\}_{k=1}^{K_0}$  denote either the EnKF or ESRF ensemble with the forecast model and initialization described above. For any integer  $N > 0$  and any  $p \in (0,1)$ , we can find  $\varepsilon_0$  such that for any  $\varepsilon < \varepsilon_0$  we have*

$$\mathbf{P}\left(|V_n^{\pm, (k)}| \geq M_n, \text{ for all } n \leq N\right) \geq 1-p.$$

*In particular, with probability  $1-p$  the ensemble is increasing exponentially with rate at least  $\sqrt{2}\rho$  for at least  $N$  consecutive steps.*

The proof works exactly as the illustration in Fig. 1. The parameters have been chosen in such a way that the cycle described in Fig. 1 will repeat itself  $N$  times, with probability  $1-p$ . The details are given in *Supporting Information*.

**Numerical Evidence**

Here we use elementary numerical experiments to demonstrate the claimed catastrophic filter divergence phenomenon. We use  $2K_0=6$  EnKF ensemble members to filter the rotation-lock

system  $\Psi$  with contraction ratio  $\rho = 0.8 > 1/\sqrt{2}$ . With  $\varepsilon \rightarrow 0$ , the angle of rotation  $\theta$  converges to  $\arctan \frac{6}{5}$ . From Proposition 1 one can deduce that the attractors of  $\Psi$  lie inside the bounded set

$$\mathcal{A} = \left\{ u \in \mathbb{R}^2 : |u|^2 \leq \frac{2(\rho^2 + 1)}{(1 - \rho^2)^2} \leq 25.31 \right\}.$$

The true model is initialized at the fixed point  $U_0 = 0$ . If the EnKF is working properly, the filter ensemble will be attracted to 0 or at least stay close to  $\mathcal{A}$ .

According to Theorem 2, with  $\varepsilon$  sufficiently small, if we start the ensemble from positions given by **6** with  $M_0 = 10 > 1/(\sqrt{2}\rho - 1)$  and  $R_0^{\pm, (k)} = 0$ , the filter ensemble will diverge to machine infinity exponentially fast for a finite number of steps  $N$ , with high probability. Here we test the claim with a range of  $\varepsilon$  and  $M_0$ . Fig. 2 shows a log-plot of the energy of the first ensemble member  $|V_n^{+, (1)}|^2$  against iteration count  $n$  with  $\varepsilon$  taken from the range  $\{0.1, 0.05, 0.02, 0.01, 0.005, 0.002\}$ . The dashed line indicates the upper bound for the attractors  $\mathcal{A}$  of the forecast model  $\Psi$ . The experiment confirms the claim of Theorem 2; in Fig. 2, for  $\varepsilon$  that are sufficiently small ( $\leq 0.01$ ), the energy of the ensemble grows exponentially fast for significantly many iterations ( $> 70$ ) to an extremely high level ( $10^{15}$  to  $10^{21}$ ), despite the attractors of the system being in a bounded region (with energy  $\leq 25.31$ ). With the exception of  $\varepsilon = 0.02$ , taking  $\varepsilon$  smaller leads to a longer continuation of exponential growth, which is in agreement with Theorem 2. In the  $\varepsilon = 0.02$  case, the ensemble does accurately track the true signal and no divergence occurs. This is a testament to the fact that, even though moderately small choices, like  $\varepsilon = 0.1$ , can lead to exponential growth to several orders of magnitude, it cannot be guaranteed unless  $\varepsilon$  is taken very small.

In Fig. 3 we show a log-plot of a similar experiment, but with  $\varepsilon = 0.02$  fixed and  $M_0$  chosen from the range  $\{10, 20, 50, 100, 200, 500\}$ . On closer examination, even though the filter is stable with  $M_0 = 10$ , choosing an  $M_0$  initialization that is only one or two

orders of magnitude higher can lead to many orders of magnitude in exponential growth.

We have also tested different combinations of  $K_0, \rho, M_0, R_0$  with  $\varepsilon$ . For small  $\varepsilon$  such as  $\varepsilon = 0.002$ , the long-lasting exponential growth of ensemble energy is extremely stable with respect to changes in other parameters. For relatively large  $\varepsilon$ , the exponential growth phenomenon is rather unstable. Take  $\varepsilon = 0.02$  for example. In the setting of Fig. 2, there is no filter divergence, but when the initialization parameter  $M_0$  is chosen differently from 20 to 500, there is exponential growth of length ranging from 15 to 90 steps. As above, this is another testament to the fact that, although sustained exponential growth occurs with moderately small  $\varepsilon$ , it must be taken very small to guarantee filter divergence.

The numerical experiments above have also been performed with both EAKF and ETKF methods and the results are unchanged. In the EAKF setting, one must make the seemingly innocuous assumption that the diagonalization of an already diagonal matrix is the trivial one  $D = |DI|^{-1}$ . In some cases, numerical packages will not adhere to this rule. Interestingly, a nontrivial choice of diagonalization will lead to stabilization of the filter in the EAKF case.

### Discussion

In the preceding text we have constructed an elementary forecast model that exhibits the absorbing ball property and that nevertheless leads to drastic malfunction of ensemble-based Kalman filters. In particular, even when the true model signal is stationary at the origin, the filter estimate diverges exponentially fast to infinity in a phenomena known as catastrophic filter divergence (7, 8). The mechanism that creates this filter malfunction is very intuitive and is illustrated by a simple picture. Moreover, the mechanism builds on generic dynamical properties that are true of much more complex dissipative systems. Namely, the locking mechanism present in our model can be reproduced in more general models via stiffness due to strongly negative Lyapunov exponents, as present in the five-dimensional Lorenz-96 model (8). Our claims are not only intuitively clear but are backed by rigorous results guaranteeing that the filter will diverge with arbitrarily high

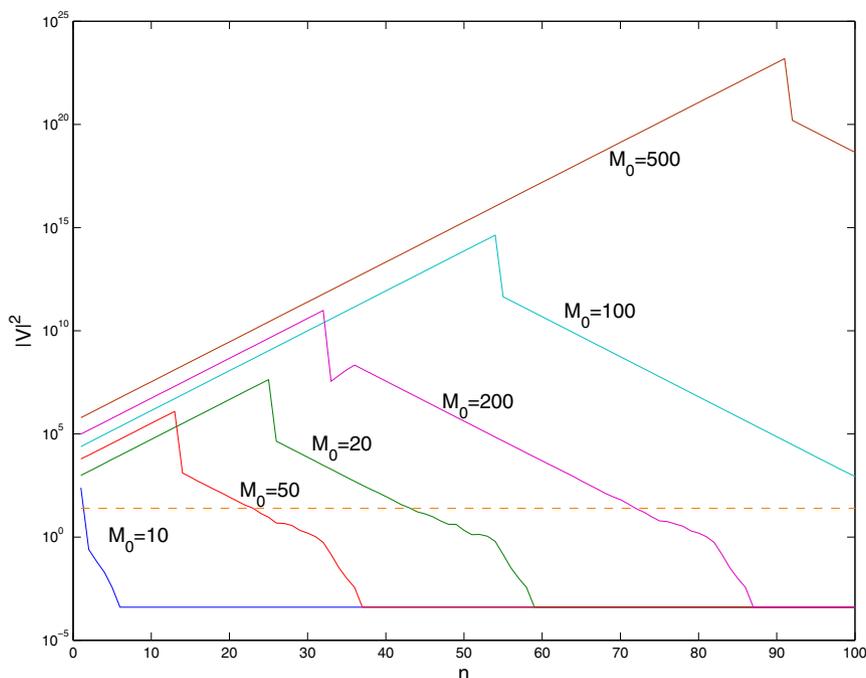


Fig. 3. Exponential energy growth of one ensemble member of EnKF. The rotation-lock system is filtered using  $2K_0 = 6$  ensembles, and  $\rho = 0.8$ ,  $\varepsilon = 0.02$ ,  $R_0 = 0$ , and  $M_0$  takes value from 10 to 500.

probability. The theoretical results are in complete agreement with the numerical experiments provided.

The benefits of this work extend beyond simply proving that catastrophic filter divergence is a genuine mathematical phenomenon and not a numerical artifact. By documenting a simple mechanism that leads to catastrophic filter divergence, we better understand what types of systems are susceptible to malfunction. In particular, the user should be cautious of catastrophic filter divergence in unstable nonlinear systems with strongly negative Lyapunov exponents, because this can lead to an ensemble alignment mechanism, as found in refs. 7 and 8. An interesting project for future research is to develop the intuition illustrated here into a test to identify and correct catastrophic filter divergence before it destroys the state estimate entirely.

Moreover, the present article implies that the stability properties of ensemble-based filters, including boundedness on an infinite time horizon and geometric ergodicity, have delicate dependence over both the dynamical properties of the nonlinear system and the form of observations. In fact, the authors have discovered a simple criterion for nonlinear dissipative systems that guarantees the prescribed stability properties. This criterion is known as observable energy dissipation and essentially requires that the model dissipate energy in the observed directions. The

present article suggests that this criterion is sharp; the rotation-locking map is an example of a forecast model that satisfies an energy criterion but not an observable energy criterion and for which ensemble-based methods exhibit no stability properties whatsoever. The reason this model-filter combination does not satisfy the observable energy criterion is simply due to the large condition number of the matrix  $H$ . More details on the criterion and the relation to the condition number can be found in ref. 11.

The catastrophic filter divergence experienced here can be easily averted through additive variance inflation. Intuitively, additive variance inflation will cause the ensemble to dealign itself and the cycle described by Fig. 1 will not iterate. In ref. 12, the authors have used the intuition presented in this article to design a notion of adaptive additive covariance inflation that not only guarantees stability of ensemble-based filters for the above forecast model, but indeed for any model that possesses an energy dissipation principle.

**ACKNOWLEDGMENTS.** D.K. is supported as a Courant instructor. This work was supported by Multidisciplinary University Research Initiative Grant N-000-1412-10912 (to A.J.M. and X.T.T.). A.J.M. is the principal investigator, and X.T.T. is supported as a postdoctoral fellow.

1. Evensen G (2003) The ensemble kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn* 53(4):343–367.
2. Bishop CH, Etherton BJ, Majumdar SJ (2001) Adaptive sampling with the ensemble transform kalman filter. Part i: Theoretical aspects. *Mon Weather Rev* 129(3):420–436.
3. Anderson JL (2001) An ensemble adjustment kalman filter for data assimilation. *Mon Weather Rev* 129(12):2884–2903.
4. Majda AJ, Harlim J (2012) *Filtering Complex Turbulent Systems* (Cambridge Univ Press, Cambridge, UK).
5. Kalnay E (2003) *Atmospheric Modeling, Data Assimilation, and Predictability* (Cambridge Univ Press, Cambridge, UK).
6. Kelly D, Law KJ, Stuart AM (2014) Well-posedness and accuracy of the ensemble kalman filter in discrete and continuous time. *Nonlinearity* 27:2579–2603.
7. Majda AJ, Harlim J (2008) Catastrophic filter divergence in filtering nonlinear dissipative systems. *Commun Math Sci* 8:27–43.
8. Gottwald G, Majda AJ (2013) A mechanism for catastrophic filter divergence in data assimilation for sparse observation networks. *Nonlinear Process Geophys* 20:705–712.
9. Temam R (1997) *Infinite Dimensional Dynamical Systems in Mechanics and Physics*. Applied Mathematical Sciences, eds Marsden JE, Sirovich L, John F (Springer, New York), Vol 68.
10. Grote MJ, Majda AJ (2006) Stable time filtering of strongly unstable spatially extended systems. *Proc Natl Acad Sci USA* 103(20):7548–7553.
11. Tong XT, Majda AJ, Kelly D (2015) Nonlinear stability and ergodicity of ensemble based Kalman filters. arXiv:1507.08307.
12. Tong XT, Majda AJ, Kelly D (2015) Nonlinear stability of the ensemble Kalman filter with adaptive covariance inflation. arXiv:1507.08319.