

# A mechanism for catastrophic filter divergence in data assimilation for sparse observation networks

Georg A. Gottwald<sup>1</sup> and A. J. Majda<sup>2</sup>

<sup>1</sup>School of Mathematics and Statistics, University of Sydney, NSW 2006, Australia, georg.gottwald@sydney.edu.au

<sup>2</sup>Department of Mathematics and Center for Atmosphere Ocean Science, Courant Institute of Mathematical Sciences New York University, U.S.A., jonjon@cims.nyu.edu

Correspondence to: Georg A. Gottwald  
(georg.gottwald@sydney.edu.au)

**Abstract.** We study catastrophic filter divergence in data assimilation procedures whereby the forecast model develops severe numerical instabilities leading to a blow up of the solution. Catastrophic filter divergence occurs in sparse observational grids with small observational noise for intermediate observation intervals and finite ensemble sizes. Using a minimal five dimensional model we establish that catastrophic filter divergence is caused by the filtering procedure producing analyses which are not consistent with the true dynamics, and stiffness caused by the fast attraction of the inconsistent analyses towards the attractor during the forecast step.

---

**Keywords.** Data assimilation; Ensemble Kalman filter; Filter divergence

## 1 Introduction

Data assimilation is the procedure to find the best estimation of the state of a dynamical system given a forecast model with possible model error and noisy observations at discrete observation intervals (Kalnay, 2002; Majda and Harlim, 2012). The presence of the often chaotic nature of the underlying nonlinear dynamics as well as the sparseness of the observational network significantly complicate this process. In the setting of ensemble based filters (Evensen, 1994, 2006) finite ensemble sizes may introduce additional sources of error (see for example Ehrendorfer (2007)). Insufficient ensemble size typically causes an underestimation of the error covariances which may ultimately lead to filter divergence when the filter trusts its own forecast and ignores the information provided by the observations. This filter divergence is caused by ensemble members aligning with the most unstable directions (Ng *et al.*, 2011) and is

exacerbated by large observational noise. In (spatially) sparse observational networks finite size effects may lead to spurious overestimating correlations between otherwise uncorrelated variables (Hamill *et al.*, 2001; Whitaker *et al.*, 2004; Liu *et al.*, 2008; Sacher and Bartello, 2008; Whitaker *et al.*, 2009), spoiling the overall analysis skill.

Harlim and Majda (2010) and Gottwald *et al.* (2011) documented a new type of filter divergence which is characterized by the forecast model diverging to machine infinity. It was shown that this catastrophic filter divergence occurs in sparse observational networks with small observational noise for moderate observation intervals., in contrast to the classical filter divergence described in the previous paragraph.

We will establish here the mechanism leading to this instability in a minimal low dimensional model: In a sparse observational grid, finite ensemble sizes cause the ensemble to align, and in the case of small observational noise generate analyses which are not consistent with the actual dynamics and are located in phase space off the attractor. If the attraction towards the attractor is sufficiently strong, the subsequent forecast step attempts to integrate a stiff dynamical system which may cause the integrator to develop numerical instabilities.

In Section 2 we introduce the minimal model for which catastrophic filter divergence is studied. We briefly describe ensemble Kalman filters in Section 3. Numerical results are presented in Section 4 and the mechanism for catastrophic filter divergence is established. We conclude with a discussion in Section 5.

## 2 A minimal model

We study the (Lorenz, 1996) Lorenz-96 model

$$\dot{z}_i = z_{i-1}(z_{i+1} - z_{i-2}) - z_i + F \quad i = 1, \dots, D \quad (1)$$

with  $\mathbf{z} = (z_1, \dots, z_D)$  and periodic  $z_{i+D} = z_i$  in a five dimensional setting. We use negative forcing here which allows strong mixing with small dimension  $D$ . We consider here  $D=5$  with  $F=-16$  which was coined the 5-mode Anti-Lorenz system by Abramov and Majda (2006). For these parameters we find as Lyapunov exponents  $\lambda = (2.72, 0.09, -0.09, -1.83, -5.89)$  for an integration lasting 250 time units. Note that  $\sum_{i=1}^5 \lambda_i = \lim_{t \rightarrow \infty} \frac{1}{t} \int \text{Tr}(M(t)) dt$  where  $M$  is the linearized vectorfield of (1), and hence  $\sum_{i=1}^5 \lambda_i = -5$ . Using the Kaplan-Yorke dimension (see for example Schuster and Just (2005)) this suggests that the attractor has a fractal dimension of  $D_{\text{attr}} = 4.15$  and trajectories are on average attracted to this manifold with the fast rate  $\lambda_5 = -5.89$ . The climatic mean and variance is estimated from a long time trajectory as  $\bar{z} = -2.47$  and  $\sigma_{\text{clim}}^2 = 33.7$ , respectively. The decay rate of the autocorrelation is estimated as  $\tau_{\text{corr}} = 0.14$  and the first zero-crossing of the autocorrelation function is around  $\tau_0 = 0.75$ .

We assume that observations of the variables are given at equally spaced discrete observation times  $t_i$  with observation interval  $\Delta t_{\text{obs}}$ . We observe only one variable  $z_1$ . It is well known that the Kalman filter is suboptimal for dynamical systems which are nonlinear and involve non-Gaussian statistics. It is pertinent to mention that although the five-dimensional Lorenz system (1) is highly nonlinear its probability density function is near-Gaussian. The Lorenz system (1) is assimilated using an Ensemble Transform Kalman filter (ETKF) (Tippett *et al.*, 2003; Wang *et al.*, 2004) which is briefly described in the following section.

## 3 Ensemble Kalman filter

In an ensemble Kalman filter (EnKF) (Evensen, 2006) an ensemble with  $k$  members  $\mathbf{z}_k \in \mathbf{R}^D$

$$\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_k] \in \mathbf{R}^{D \times k}$$

is propagated by the full nonlinear dynamics

$$\dot{\mathbf{Z}} = \mathbf{f}(\mathbf{Z}), \quad \mathbf{f}(\mathbf{Z}) = [f(\mathbf{z}_1), f(\mathbf{z}_2), \dots, f(\mathbf{z}_k)] \in \mathbf{R}^{D \times k}. \quad (2)$$

The ensemble is split into its mean

$$\bar{\mathbf{z}} = \frac{1}{k} \sum_{i=1}^k \mathbf{z}_i$$

and its ensemble deviation matrix

$$\mathbf{Z}' = \mathbf{Z} - \bar{\mathbf{z}}\mathbf{e}^T,$$

where  $\mathbf{e} = [1, \dots, 1]^T \in \mathbf{R}^k$ . The ensemble deviation matrix  $\mathbf{Z}'$  is used to provide a Monte-Carlo estimate of the forecast covariance matrix

$$\mathbf{P}_f(t) = \frac{1}{k-1} \mathbf{Z}'(t) \mathbf{Z}'(t)^T \in \mathbf{R}^{D \times D}.$$

Note that  $\mathbf{P}_f(t)$  is rank-deficient if the ensemble size  $k$  is smaller than the rank of the covariance matrix. The rank is generically not known in atmospheric models with  $D$  of the order of  $10^9$ , but is believed to be orders of magnitudes smaller than  $D$  and orders of magnitude larger than 100, the typical ensemble size in numerical weather prediction.

Given the forecast ensemble  $\mathbf{Z}_f = \mathbf{Z}(t_i - \epsilon)$  and the associated forecast error covariance matrix (or the *prior*)  $\mathbf{P}_f(t_i - \epsilon)$ , the actual Kalman analysis (Kalnay, 2002; Evensen, 2006; Simon, 2006) updates a forecast into a so-called analysis (or the *posterior*). Variables at times  $t = t_i - \epsilon$  are evaluated before taking observations  $\mathbf{y}_o$  into account in the analysis step, and variables at times  $t = t_i + \epsilon$  are evaluated after the analysis step when the observations have been taken into account. In the first step of the analysis the forecast mean  $\bar{\mathbf{z}}_f$  is updated to the analysis mean

$$\bar{\mathbf{z}}_a = \bar{\mathbf{z}}_f - \mathbf{K}_o [\mathbf{H}\bar{\mathbf{z}}_f - \mathbf{y}_o], \quad (3)$$

where the Kalman gain matrix is defined as

$$\mathbf{K}_o = \mathbf{P}_f \mathbf{H}^T (\mathbf{H} \mathbf{P}_f \mathbf{H}^T + \mathbf{R}_o)^{-1}. \quad (4)$$

The analysis covariance  $\mathbf{P}_a$  is given by

$$\mathbf{P}_a = (\mathbf{I} - \mathbf{K}_o \mathbf{H}) \mathbf{P}_f. \quad (5)$$

To calculate an ensemble  $\mathbf{Z}_a$  which is consistent with the error covariance after the observation  $\mathbf{P}_a$ , and which therefore needs to satisfy

$$\mathbf{P}_a = \frac{1}{k-1} \mathbf{Z}_a \mathbf{Z}_a^T,$$

we use the method of ensemble square root filters (Simon, 2006) which expresses the analysis ensemble as a linear combination of the forecast ensemble. In particular we use the method proposed in (Tippett *et al.*, 2003; Wang *et al.*, 2004), the so called ensemble transform Kalman filter (ETKF). Alternatively one could have chosen the ensemble adjustment filter (Anderson, 2001) or the continuous Kalman-Bucy filter which does not require the inversion of matrix inverses (Bergemann *et al.*, 2009). A new forecast  $\mathbf{Z}(t_{i+1} - \epsilon)$  is then obtained by propagating  $\mathbf{Z}_a$  with the full nonlinear dynamics to the next time of observation. The numerical results presented in the next section are obtained with this method.

#### 4 The genesis of catastrophic filter divergence

150 We observe only one of the five variables  $z_i$  (wlog we use  $z_1$ ) and generate observations  $y_o$  from the truth by adding Gaussian observational noise with small observational error covariance  $\mathbf{R}_o = 0.01$  after equal observation intervals  $\Delta t_{\text{obs}}$ . We have used two methods as numerical integration schemes to integrate forward in time the system (1) during the forecasting step; a first order in time backward Euler step and a second-order in time implicit midpoint rule which is unconditionally stable for the system (1) (Stuart and Humphries, 1996).

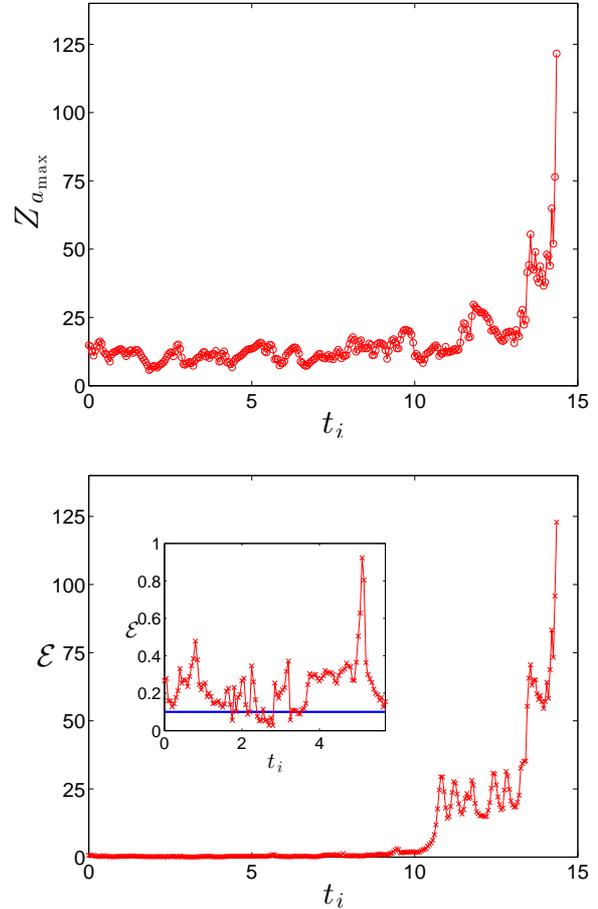
160 Although the implicit midpoint scheme is unconditionally stable if run in free forecast mode, it appears that the analysis step is able to push the radius of convergence further and further out. In the following we employ the implicit midpoint solver where we limit the number of iterations to 1000 in each integration time step; in the regimes when there is no catastrophic filter divergence typically less than 10 iterations are sufficient for convergence; we note that if we do not restrict the number of iterations, the integrator does not converge but instead reaches a periodic orbit at amplitudes of  $10^5$ .

175 In Figure 1 we show an instant of a catastrophic filter divergence for  $dt = 0.025$  and  $\Delta t_{\text{obs}} = 0.05$  where we used  $k = 6$  ensemble members (so the forecast error covariance matrix is not necessarily rank deficient). Besides the maximal amplitude of the analysis ensemble, we show the norm-error  $\mathcal{E}$  of the analysis

$$\mathcal{E}(t_i) = \|\bar{\mathbf{z}}_a(t_i) - \mathbf{z}_t(t_i)\| \quad (6)$$

180 evaluated at each analysis cycle  $t_i$  between the truth  $\mathbf{z}_t$  and the ensemble mean  $\bar{\mathbf{z}}_a$ . After  $t_i = 14.5$  the norm-error becomes machine infinity, due to the forecast model developing a numerical instability. The genesis of the blow-up is clearly seen from Figure 1: Until  $t_1 \approx 10$  the filter is stable and the analysis is tracking where the norm error may be even smaller than the observational error. This is followed by a non-tracking episode lasting to  $t_i \approx 13$  in which the norm error evolves around a mean value of approximately  $\langle \mathcal{E} \rangle^2 \approx 20.5 \approx \sqrt{\langle \mathcal{E}^2 \rangle - \text{Var}[\mathcal{E}]} = \sqrt{2D\sigma_{\text{clim}}^2 - \text{Var}[\mathcal{E}]}$ , suggesting 205 that the analysis is uncorrelated from the truth and not controlled by the observations anymore. This episode precedes the actual blow-up episode of the forecast integrator in which the norm-error grows to machine infinity.

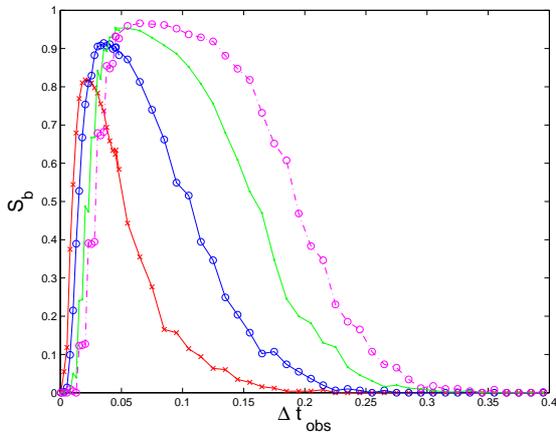
210 In order to get meaningful statistics on the occurrence of blow-ups we count the number  $N_b$  of blow-ups that occur before a total of 500 simulations have terminated without blow-up. The proportions of blow-ups for the respective filters is then given by  $S_b = N_b / (N_b + 500)$ . We define blow-up as 215 instances where the maximal value of either an analysis or forecast exceeds  $100\sigma_{\text{clim}} \approx 580$ . We have checked that once this threshold is crossed the numerical instability develops



**Fig. 1.** (Top): Maximal absolute value  $\mathbf{Z}_{a,\text{max}}$  of the analysis ensemble over all  $D = 5$  components for  $dt = 0.025$  and  $\Delta t_{\text{obs}} = 0.05$ . (Bottom): The error norm  $\mathcal{E}$  as a function of analyses cycles. The continuous line (online blue) in the inset shows the observational error  $\sqrt{R}$ .

machine infinity within a few more time steps of the forecast model. In Figure 2 we show  $S_b$  as a function of the observation interval  $\Delta t_{\text{obs}}$  for several values of the integration time step  $dt$ . It is seen that blow up occurs for moderate observation time intervals. No blow-up occurs for sufficiently small or sufficiently large values of  $\Delta t_{\text{obs}}$ . Furthermore, in line with the stability of the implicit midpoint rule, the percentage of blow-ups as well as the range of  $\Delta t_{\text{obs}}$  for which blow-up occurs is reduced by reducing the integration time step.

To understand the mechanism by which catastrophic filter divergence is initiated in an ensemble Kalman filter, we first focus on the non-tracking episode preceding the actual blow-up episode (i.e. the period  $t_i \in (10, 13)$  in Figure 1). We propose that catastrophic filter divergence is caused by insufficient ensemble sizes paired with small observational noise.



**Fig. 2.** Percentage  $S_b$  of blow-ups as a function of the observation interval  $\Delta t_{\text{obs}}$  for several values of the integration time step  $dt$ . We plot results for  $dt = 0.001$  (crosses, online red),  $dt = 0.0025$  (open circles, online blue),  $dt = 0.005$  (dots, online green) and  $dt = 0.0075$  (open circles with dashed line, online magenta).

We have checked that by increasing the ensemble size to im-  
 220 practically high values of  $k = 400$  we were able to avoid  
 catastrophic blow-up. To monitor the ensemble spread we  
 consider the ensemble dimension  $D_{\text{ens}}$  as defined in (Patil  
*et al.*, 2001; Pazó *et al.*, 2011)

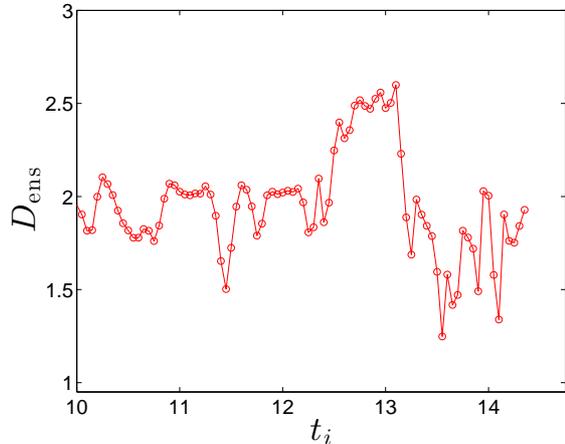
$$D_{\text{ens}} = \frac{\left(\sum_{i=1}^k \sqrt{\mu_i}\right)^2}{\sum_{i=1}^k \mu_i},$$

where  $\mu_i$  denotes the  $i$ th eigenvalue of the  $k \times k$  covariance  
 260 matrix

$$C = X_f^T X_f.$$

Note that  $D_{\text{ens}}$  takes values between 1 and  $\min(k, D)$ ,  
 230 depending on whether the ensemble members are all aligned  
 or are orthogonal to each other. In Figure 3 we show the  
 ensemble dimension as a function of time for an ensemble  
 with  $k = 6$  members corresponding to the blow-up presented  
 235 in Figure 1. It is seen that  $D_{\text{ens}} \approx 2$  during the stable tracking  
 episode, indicating that the ensemble is not spanning all  
 270 directions on the attractor (we recall the fractal attractor  
 dimension to be  $D_{\text{attr}} = 4.15$ ) but instead is aligning with  
 the first two Lyapunov vectors (cf. (Ng *et al.*, 2011)). On the  
 240 other hand, for ensemble sizes of  $k = 400$  we observe that  
 mostly  $D_{\text{ens}} > 4$ . Around  $t_i = 12.5$  the ensemble dimension  
 275 increases before falling below values of 2, indicating ensemble  
 collapse.

245 Finite ensemble sizes and the associated loss of ensemble  
 spread are known to cause non-catastrophic filter divergence  
 280 in which the filter trusts the wrong forecasts ignoring



**Fig. 3.** Ensemble dimension  $D_{\text{ens}}$  as a function of analyses cycles. Parameters as in Figure 1

error-correcting observations (Houtekamer and Mitchell,  
 1998; Hamill *et al.*, 2001; Sacher and Bartello, 2008;  
 Ng *et al.*, 2011). Finite ensemble sizes cause the forecast  
 error covariance  $\mathbf{P}_f$  to exhibit on one hand small diagonal  
 covariances and on the other hand off-diagonal entries values  
 of unrealistically large absolute value (Hamill *et al.*, 2001).  
 This leads to unrealistic innovations of the unobserved  
 255 variables towards the observation of the observed distant  
 variable. Gottwald *et al.* (2011) showed that catastrophic  
 filter divergencies are suppressed by a variance limiting  
 Kalman filter (VLKF) which controls overestimation of the  
 analysis error covariance. To further check that an unreal-  
 istic overestimation of off-diagonal entries of the forecast  
 covariance matrix  $\mathbf{P}_f$  is responsible for the catastrophic  
 filter divergence, we have checked that blow-ups can be  
 avoided by employing covariance localisation into the data  
 assimilation procedure. Houtekamer and Mitchell (2001)  
 and Hamill *et al.* (2001) achieved covariance localisation  
 by Schur-multiplication of the forecast error covariance  
 $\mathbf{P}_f$  with a localisation matrix  $C_{\text{loc}}$ . We used the compactly  
 supported localisation function introduced by Gaspari and  
 Cohn (1999), and found that catastrophic filter divergence is  
 suppressed. We remark that the actual truth, however, does  
 indeed exhibit nontrivial correlations between all variables  
 for our parameters in this low-dimensional with  $D = 5$ .  
 Furthermore, we have employed for all our numerical sim-  
 ulations a 5% multiplicative covariance inflation, a standard  
 remedy to control underestimation of covariances (Anderson  
 and Anderson, 1999). However, multiplicative inflation does  
 not alter the subspace spanned by the ensemble and therefore  
 does not prevent catastrophic filter divergence.

The destructive interplay of sparse accurate observations  
 and finite size ensembles can be illustrated as follows. The

Kalman filter produces innovations according to (3) which read for our case as

$$\bar{z}_{ai} = \bar{z}_{fi} - \frac{\mathbf{P}_{f_{i1}}}{\mathbf{P}_{f_{i1}} + \mathbf{R}_o} [\bar{z}_{fi} - \mathbf{y}_o],$$

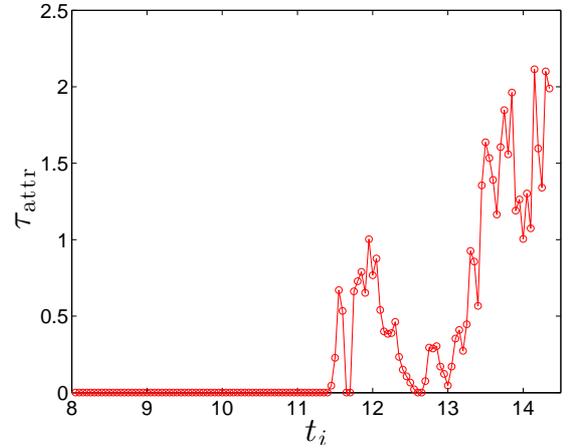
for  $i = 1, \dots, 5$ . The combination of small ensemble sizes causing small values of  $\mathbf{P}_{f_{i1}}$  and large absolute values of  $\mathbf{P}_{f_{i1}}$  for  $i > 1$  with small observational noise  $\mathbf{R}_o$  leads to analyses which are significantly influenced by the observation  $\mathbf{y}_o$  at site  $i = 1$ , irrespective of the actual physical correlations present in the dynamics. The resulting analysis is therefore not dynamically consistent but lies in phase space off the attractor. This is confirmed in Figure 4 where we show that during the non-tracking period  $t_i > 10$  the analyses are not situated on the attractor. We measure the distance of the analyses to the attractor by propagating the analysis forward in time and estimating the time  $\tau_{\text{attr}}$  taken for the trajectory to reach an Euclidean distance  $\theta$  from the attractor. We created an approximation of the attractor by storing  $2 \cdot 10^6$  data points sampled at 0.005 time units. We choose  $\theta = 1$ . It is clearly seen that in the non-tracking episode the analyses tend to lie off the attractor. The forecast model, initialised with an analysis lying off the attractor, then tries to follow the stable manifold towards the globally attracting set with a rate which is in our case very fast on average with a Lyapunov exponent of  $-5.89$ . This renders the dynamical system stiff developing numerical instabilities for sufficiently large time steps  $dt$  causing the filter to catastrophically diverge to machine infinity.

As seen in Figure 2 there is no filter divergence for sufficiently small and sufficiently large observation intervals  $\Delta t_{\text{obs}}$ . This can now be understood as follows: For too small observation intervals the forecast model will not have sufficiently propagated the analysis away from dynamically realistic values, whereas for sufficiently large values of  $\Delta t_{\text{obs}} \gg \tau_{\text{corr}}$  the ensemble will have acquired sufficient spread with  $D_{\text{ens}} \geq 4$  exploring the whole of the attractor.

It is pertinent to mention that the existence of alignment of the ensemble and the occurrence of off-attractor analyses does not necessarily cause catastrophic filter divergence (cf Figures 3 and 4 at  $t_i = 11.5$ ).

## 5 Discussion

We have numerically established that catastrophic filter divergence is caused by the interplay of finite size effects and sparse observations with small observational noise producing analyses which may be situated in phase space far away from the actual attractor. The subsequent attraction back towards the attractor by the forecast model may develop numerical instabilities if the attraction rate is sufficiently large. This suggests that blow-up is to be expected in sparse observational networks involving observables which exhibit a large degree of variance. If those high variance fields are measured accu-



**Fig. 4.** Distance between analysis and the attractor as measured by  $\tau_{\text{attr}}$ . Parameters as in Figure 1

rately, catastrophic filter divergence is possible. This is, for example, the case in data assimilation of small scale intermittent turbulent fields or in situations where sparse accurate observations of variables exhibiting strong spatial gradients such as jets can cause numerical instabilities to occur (Anderson, 2012).

Our work established a dynamical genesis of catastrophic filter divergence. To avoid this type of filter divergence it was argued that multiplicative covariance inflation was not sufficient. Besides an impractical reduction of the integration time step (or an increase of the limit of iterations required in an implicit method), to control the stiffness of the dynamical system, or an impractical increase of the number of ensembles to eliminate finite size sampling errors, covariance localisation was found to be effective in mitigating catastrophic filter divergence. Our study shows that choosing too small observation covariances can lead to filter blow ups as observed in climate models (Shlyueva, 2012).

These results point towards the necessity of using methods with judicious model error to avoid such divergences in filtering turbulent systems (Harlim and Majda, 2010; Majda and Harlim, 2012; Keating *et al.*, 2012) or using variance limiting strategies as proposed in Gottwald *et al.* (2011); Mitchell and Gottwald (2012).

*Acknowledgements.* We are grateful to Jeff Anderson, Heikki Jarvinen, Balu Nadiga and Anna Shlyueva for sharing their experiences with catastrophic filter divergences in their numerical simulations. GAG acknowledges support from the Australian Research Council, and AJM by ONR grants N0014-11-1-0306, DRI grant, N00014-10-1-0554, and MURI grant, N00014-12-1-0912.

## References

- Abramov, R. V. and Majda, A. J. (2006). New Approximations and Tests of Linear Fluctuation-Response for Chaotic Nonlinear Forced-Dissipative Dynamical Systems. *Journal of Nonlinear Science*, **18**(3), 303–341.
- Anderson, J. L. (2001). An ensemble adjustment Kalman filter for data assimilation. *Monthly Weather Review*, **129**(12), 2884–2903.
- Anderson, J. L. (2012). personal communication.
- Anderson, J. L. and Anderson, S. L. (1999). A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Review*, **127**(12), 2741–2758.
- Bergemann, K., Gottwald, G. A., and Reich, S. (2009). Ensemble propagation and continuous matrix factorization algorithms. *Quarterly Journal of the Royal Meteorological Society*, **135**(643), 1560–1572.
- Ehrendorfer, M. (2007). A review of issues in ensemble-based Kalman filtering. *Meteorologische Zeitschrift*, **16**(6), 795–818.
- Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, **99**(C5), 10143–10162.
- Evensen, G. (2006). *Data Assimilation: The Ensemble Kalman Filter*. Springer, New York.
- Gaspari, G. and Cohn, S. E. (1999). Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, **125**(554), 723–757.
- Gottwald, G. A., Mitchell, L., and Reich, S. (2011). Controlling overestimation of error covariance in ensemble Kalman filters with sparse observations: A variance limiting Kalman filter. *Monthly Weather Review*, **139**(8), 2650–2667.
- Hamill, T. M., Whitaker, J. S., and Snyder, C. (2001). Distance-dependent filtering of background covariance estimates in an ensemble Kalman filter. *Monthly Weather Review*, **129**(11), 2776–2790.
- Harlim, J. and Majda, A. J. (2010). Catastrophic filter divergence in filtering nonlinear dissipative systems. *Communications in Mathematical Sciences*, **8**(1), 27–43.
- Houtekamer, P. L. and Mitchell, H. L. (1998). Data assimilation using an ensemble Kalman filter technique. *Monthly Weather Review*, **126**(3), 796–811.
- Houtekamer, P. L. and Mitchell, H. L. (2001). A sequential ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, **129**(1), 123–136.
- Kalnay, E. (2002). *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, Cambridge.
- Keating, S. R., Majda, A. J., and Smith, K. S. (2012). New methods for estimating ocean eddy heat transport using satellite altimetry. *Monthly Weather Review*, **140**, 1703–1722.
- Liu, J., Fertig, E. J., Li, H., Kalnay, E., Hunt, B. R., Kostelich, E. J., Szunyogh, I., and Todling, R. (2008). Comparison between local ensemble transform Kalman filter and PSAS in the NASA finite volume GCM - perfect model experiments. *Nonlinear Processes in Geophysics*, **15**(4), 645–659.
- Lorenz, E. N. (1996). Predictability - a problem partly solved. In T. Palmer, editor, *Predictability*. European Centre for Medium-Range Weather Forecast, Shinfield Park, Reading, UK.
- Majda, A. J. and Harlim, J. (2012). *Filtering Complex Turbulent Systems*. Cambridge University Press, Cambridge.
- Mitchell, L. and Gottwald, G. A. (2012). Controlling model error of underdamped forecast models in sparse observational networks using a variance-limiting kalman filter. *Quarterly Journal of the Royal Meteorological Society*, **139**, 212–225.
- Ng, G.-H. C., McLaughlin, D., Entekhabi, D., and Ahanin, A. (2011). The role of model dynamics in ensemble Kalman filter performance for chaotic systems. *Tellus A*, **63**, 958–977.
- Patil, D. J., Hunt, B., Kalnay, E., Yorke, J., and Ott, E. (2001). Local low dimensionality of atmospheric dynamics. *Physical Review Letters*, **86**(26), 5878–5881.
- Pazó, D., Rodriguez, M. A., and López, J. M. (2011). Maximizing the statistical diversity of an ensemble of bred vectors by using the geometric norm. *Journal of the Atmospheric Sciences*, **68**(7), 1507–1512.
- Sacher, W. and Bartello, P. (2008). Sampling errors in ensemble Kalman filtering. Part I: Theory. *Monthly Weather Review*, **136**, 3035–3049.
- Schuster, H. G. and Just, W. (2005). *Deterministic Chaos: An Introduction*. John Wiley & Sons, 4th edition.
- Shlyueva, A. (2012). personal communication.
- Simon, D. J. (2006). *Optimal State Estimation*. John Wiley & Sons, Inc., New York.
- Stuart, A. M. and Humphries, A. (1996). *Dynamical Systems and Numerical Analysis*. Cambridge University Press, Cambridge.
- Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M., and Whitaker, J. S. (2003). Ensemble square root filters. *Monthly Weather Review*, **131**(7), 1485–1490.
- Wang, X., Bishop, C. H., and Julier, S. J. (2004). Which is better, an ensemble of positive-negative pairs or a centered spherical simplex ensemble? *Monthly Weather Review*, **132**(7), 1590–1505.
- Whitaker, J. S., Compo, G. P., Wei, X., and Hamill, T. M. (2004). Reanalysis without radiosondes using ensemble data assimilation. *Monthly Weather Review*, **132**(5), 1190–1200.
- Whitaker, J. S., Compo, G. P., and Thépaut, J. N. (2009). A comparison of variational and ensemble-based data assimilation systems for reanalysis of sparse observations. *Monthly Weather Review*, **137**(6), 1991–1999.