

EnKF and Catastrophic filter divergence

David Kelly

Andrew Stuart

Mathematics Institute
University of Warwick
Coventry UK CV4 7AL
dtbkelly@gmail.com

October 30, 2013

DAS 13, University of Maryland.

The set-up for EnKF

We have a **deterministic model**

$$\frac{d\mathbf{v}}{dt} = F(\mathbf{v}) \quad \text{with } \mathbf{v}_0 \sim N(m_0, C_0).$$

We will denote $\mathbf{v}(t) = \Psi_t(\mathbf{v}_0)$.

We want to **estimate** $\mathbf{v}_j = \mathbf{v}(jh)$ for some $h > 0$ and $j = 0, 1, \dots, J$ given the **observations**

$$y_{j+1} = H\mathbf{v}_{j+1} + \xi_{j+1} \quad \text{for } \xi_{j+1} \text{ iid } N(0, \Gamma).$$

The set-up for EnKF

We estimate using an **ensemble** of particles $\{u^{(k)}\}_{k=1}^K$. Each particle is a statistical **representative** of the **posterior**.

For each particle, we have an **artificial observation**

$$y_{j+1}^{(k)} = y_{j+1} + \xi_{j+1}^{(k)} \quad , \quad \xi_{j+1}^{(k)} \text{ iid } N(0, \Gamma).$$

We update each particle using the **Kalman update**

$$u_{j+1}^{(k)} = \Psi_h(u_j^{(k)}) + G(u_j) \left(y_{j+1}^{(k)} - H\Psi_h(u_j^{(k)}) \right) \quad ,$$

where $G(u_j)$ is the **Kalman gain** computed using the **forecasted ensemble covariance**

$$\hat{C}_{j+1} = \frac{1}{K} \sum_{k=1}^K (\Psi_h(u_j^{(k)}) - \overline{\Psi_h(u_j)})^T (\Psi_h(u_j^{(k)}) - \overline{\Psi_h(u_j)}) \quad .$$

Filter divergence

It has been observed (★) that the ensemble can **blow-up** (ie. reach machine-infinity) in **finite time**, even when the model has nice bounded solutions.

This is known as **catastrophic filter divergence**.

It is suggested in (★) that this is caused by numerically integrating a stiff-system. Our aim is to “prove” this.

★ Harlim, Majda (2010), Gottwald (2011), Gottwald, Majda (2013).

Discrete time results

We make a “**dissipativity**” assumption on F . Namely that

$$F(\cdot) = A \cdot + B(\cdot, \cdot) \quad (\dagger)$$

with A linear elliptic and B bilinear, satisfying certain estimates and symmetries.

Eg. 2d-Navier-Stokes, Lorenz-63, Lorenz-96.

Theorem (AS,DK)

If $H = I$ and $\Gamma = \gamma^2 I$, then there exists constant β, K such that

$$\mathbf{E}|u_j^{(k)}|^2 \leq e^{2\beta jh} \mathbf{E}|u_0^{(k)}|^2 + 2K\gamma^2 \left(\frac{e^{2\beta jh} - 1}{e^{2\beta h} - 1} \right)$$

Rmk. This becomes useless as $h \rightarrow 0$

The EnKF equations look like a discretization

Recall the ensemble update equation

$$\begin{aligned}u_{j+1}^{(k)} &= \Psi_h(u_j^{(k)}) + G(u_j) \left(y_{j+1}^{(k)} - H\Psi_h(u_j^{(k)}) \right) \\ &= \Psi_h(u_j^{(k)}) + \hat{C}_{j+1}H^T(H^T\hat{C}_{j+1}H + \Gamma)^{-1} \left(y_{j+1}^{(k)} - H\Psi_h(u_j^{(k)}) \right)\end{aligned}$$

Subtract $u_j^{(k)}$ from both sides and divide by h

$$\begin{aligned}\frac{u_{j+1}^{(k)} - u_j^{(k)}}{h} &= \frac{\Psi_h(u_j^{(k)}) - u_j^{(k)}}{h} \\ &\quad + \hat{C}_{j+1}H^T(hH^T\hat{C}_{j+1}H + h\Gamma)^{-1} \left(y_{j+1}^{(k)} - H\Psi_h(u_j^{(k)}) \right)\end{aligned}$$

Clearly we need to rescale the noise (ie. Γ).

Continuous-time limit

If we set $\Gamma = h^{-1}\Gamma_0$ and substitute $y_{j+1}^{(k)}$, we obtain

$$\frac{u_{j+1}^{(k)} - u_j^{(k)}}{h} = \frac{\Psi_h(u_j^{(k)}) - u_j^{(k)}}{h} + \widehat{C}_{j+1}H^T(hH^T\widehat{C}_{j+1}H + \Gamma_0)^{-1} \\ \left(H\mathbf{v} + h^{-1/2}\Gamma_0^{1/2}\xi_{j+1} + h^{-1/2}\Gamma_0^{1/2}\xi_{j+1}^{(k)} - H\Psi_h(u_j^{(k)}) \right)$$

But we know that

$$\Psi_h(u_j^{(k)}) = u_j^{(k)} + O(h)$$

and

$$\widehat{C}_{j+1} = \frac{1}{K} \sum_{k=1}^K (\Psi_h(u_j^{(k)}) - \overline{\Psi_h(u_j)})^T (\Psi_h(u_j^{(k)}) - \overline{\Psi_h(u_j)}) \\ = \frac{1}{K} \sum_{k=1}^K (u_j^{(k)} - \overline{u_j})^T (u_j^{(k)} - \overline{u_j}) + O(h) = C(u_j) + O(h)$$

Continuous-time limit

We end up with

$$\begin{aligned} \frac{u_{j+1}^{(k)} - u_j^{(k)}}{h} &= \frac{\Psi_h(u_j^{(k)}) - u_j^{(k)}}{h} - C(u_j)H^T\Gamma_0^{-1}H(u_j^{(k)} - v_j) \\ &\quad + C(u_j)H^T\Gamma_0^{-1} \left(h^{-1/2}\xi_{j+1} + h^{-1/2}\xi_{j+1}^{(k)} \right) + O(h) \end{aligned}$$

This looks like a **numerical scheme** for

$$\begin{aligned} \frac{d\mathbf{u}^{(k)}}{dt} &= F(\mathbf{u}^{(k)}) - C(\mathbf{u})H^T\Gamma_0^{-1}H(\mathbf{u}^{(k)} - \mathbf{v}) \quad (\bullet) \\ &\quad + C(\mathbf{u})H^T\Gamma_0^{-1/2} \left(\frac{d\mathbf{W}^{(k)}}{dt} + \frac{d\mathbf{B}}{dt} \right). \end{aligned}$$

Rmk. The extra dissipation term **only sees differences in observed space** and **only dissipates in the space spanned by ensemble**.

Continuous-time results

Theorem (AS,DK)

Suppose the model v satisfies (\dagger) and $\{u^{(k)}\}_{k=1}^K$ satisfy (\bullet) . Let

$$e^{(k)} = u^{(k)} - v .$$

If $H = I$ and $\Gamma = \gamma^2 I$, then there exists constant β, K such that

$$\mathbf{E} \sum_{k=1}^K |e^{(k)}(t)|^2 \leq \mathbf{E} \sum_{k=1}^K |e^{(k)}(0)|^2 \exp(\beta t) .$$

Summary + Future Work

- (1) Writing down an SDE/SPDE allows us to see the **important quantities** in the algorithm.
- (2) Does not “prove” that filter divergence is a numerical phenomenon, but is a decent starting point.
 - (1) Improve the condition on H .
 - (2) If we can **measure** the important quantities, then we can test the performance during the algorithm.
 - (3) Suggests new EnKF-like algorithms, for instance by discretising the stochastic PDE in a more **numerically stable** way.