Data Assimilation: New Challenges in Random and Stochastic Dynamical Systems

Daniel Sanz-Alonso & Andrew Stuart

D Blömker (Augsburg), D Kelly (NYU), KJH Law (KAUST), A. Shukla (Warwick), KC Zygalakis (Southampton)

SIAM Dynamical Systems 2015
Snowbird, Utah, May 18^{th} 2015

Funded by EPSRC, ERC and ONR

Enabling Quantification of EQUIP Uncertainty for Inverse Problems

THE UNIVERSITY OF WARWICK

Outline

Data Assimilation: New Challenges in Random and Stochastic Dynamical Systems

Daniel Sanz-Alonso & Andrew Stuart

D Blömker (Augsburg), D Kelly (NYU), KJH Law (KAUST), A. Shukla (Warwick), KC Zygalakis (Southampton)

SIAM Dynamical Systems 2015
Snowbird, Utah, May 18^{th} 2015

Funded by EPSRC, ERC and ONR

Enabling Quantification of EQUIP Uncertainty for Inverse Problems

THE UNIVERSITY OF WARWICK

Table of Contents

Signal: Unpredictability

Consider the following map with random initial conditions:

$$v_{j+1} = \Psi(v_j), \qquad v_0 \sim \mu_0.$$

Example – Lorenz '63 $(a, b, r) = (10, 8/3, 28)$.

$$\frac{dv^{(1)}}{dt} = -a(v^{(1)} - v^{(2)}),$$

$$\frac{dv^{(2)}}{dt} = -av^{(1)} - v^{(2)} - v^{(1)}v^{(3)},$$

$$\frac{dv^{(3)}}{dt} = -bv^{(3)} + v^{(1)}v^{(2)} - b(r + a).$$

Set $v = (v^{(1)}, v^{(2)}, v^{(3)})$, $v_j = v(jh)$ to put in framework above.

Signal and Observation: Control Unpredictability?

Signal, **deterministic** (chaotic) dynamics on Hilbert space $(\mathcal{H}, \langle\cdot,\cdot\rangle, |\cdot|)$:

Signal Process

$$v_{j+1} = \Psi(v_j), \qquad v_0 \sim \mu_0.$$

Signal and Observation: Control Unpredictability?

Signal, **deterministic** (chaotic) dynamics on Hilbert space $(\mathcal{H}, \langle\cdot,\cdot\rangle, |\cdot|)$:

Signal Process

$$v_{j+1} = \Psi(v_j), \qquad v_0 \sim \mu_0.$$

Observations, **partial and noisy**, $P : \mathcal{H} \to \mathcal{H}$ is a **projection**.

Observation Process

$$y_{j+1} = Pv_{j+1} + \epsilon\xi_{j+1}, \qquad \mathbb{E}\xi_j = 0, \quad \mathbb{E}|\xi_j|^2 = 1, \quad \text{i.i.d.}$$

Signal and Observation: Control Unpredictability?

Signal, **deterministic** (chaotic) dynamics on Hilbert space $(\mathcal{H}, \langle\cdot,\cdot\rangle, |\cdot|)$:

Signal Process

$$v_{j+1} = \Psi(v_j), \qquad v_0 \sim \mu_0.$$

Observations, **partial and noisy**, $P : \mathcal{H} \to \mathcal{H}$ is a **projection**.

Observation Process

$$y_{j+1} = Pv_{j+1} + \epsilon\xi_{j+1}, \qquad \mathbb{E}\xi_j = 0, \quad \mathbb{E}|\xi_j|^2 = 1, \quad \text{i.i.d.}$$

Filter: probability distribution of v_j given observations to time j:

Filter

$$\mu_j(A) = \mathbb{P}(v_j \in A|\mathcal{F}_j), \quad \mathcal{F}_j = \sigma(y_1, \dots, y_j).$$

Goal (Cerou [3], SIAM J. Cont. Opt. 2000)

Key Question: For which Ψ and P does the filter μ^j concentrate on the true signal, up to error ϵ, in the large-time limit?

Goal (Cerou [3], SIAM J. Cont. Opt. 2000)

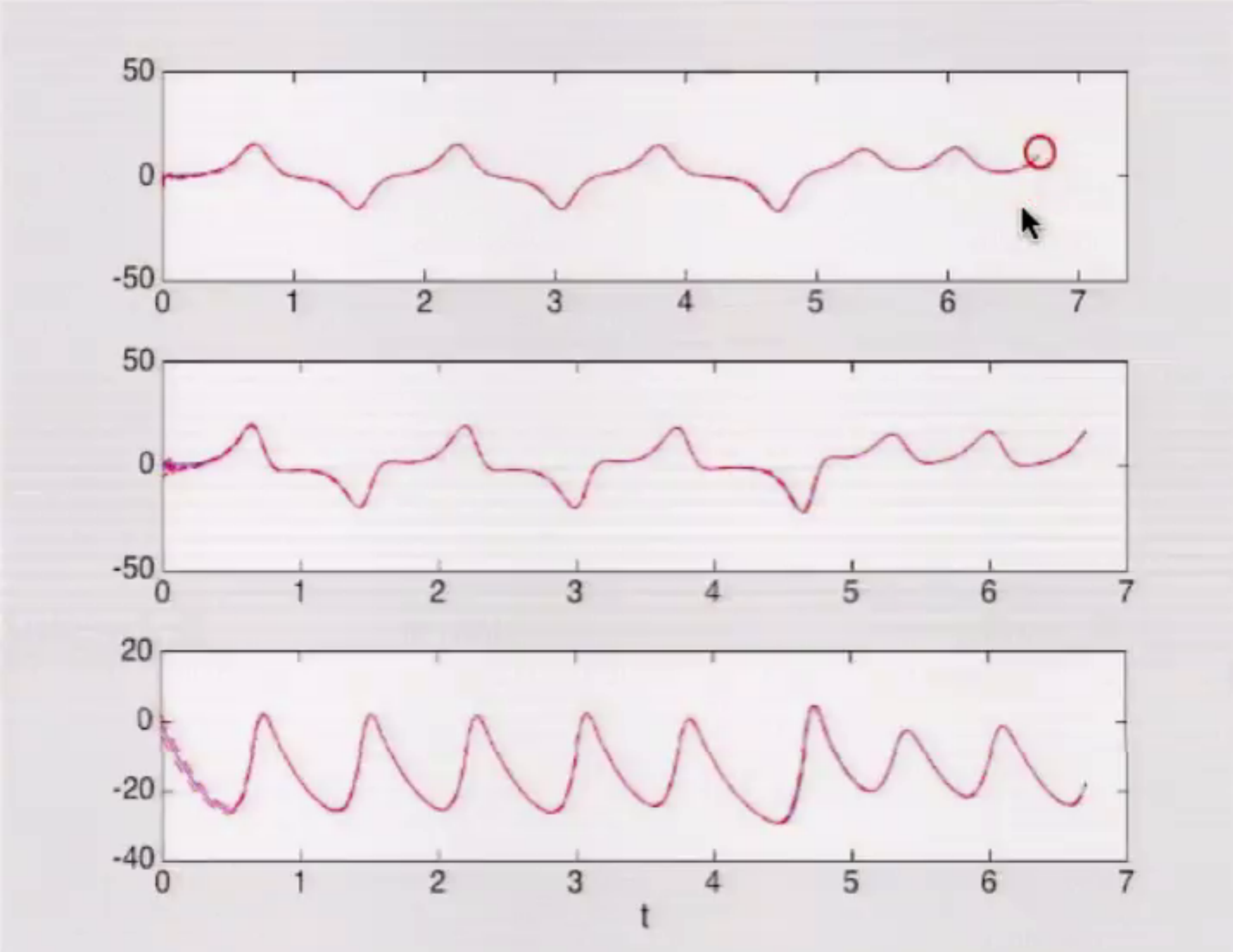
Key Question: For which Ψ and P does the filter μ^j concentrate on the true signal, up to error ϵ, in the large-time limit?

Key Problem: Ψ may expand

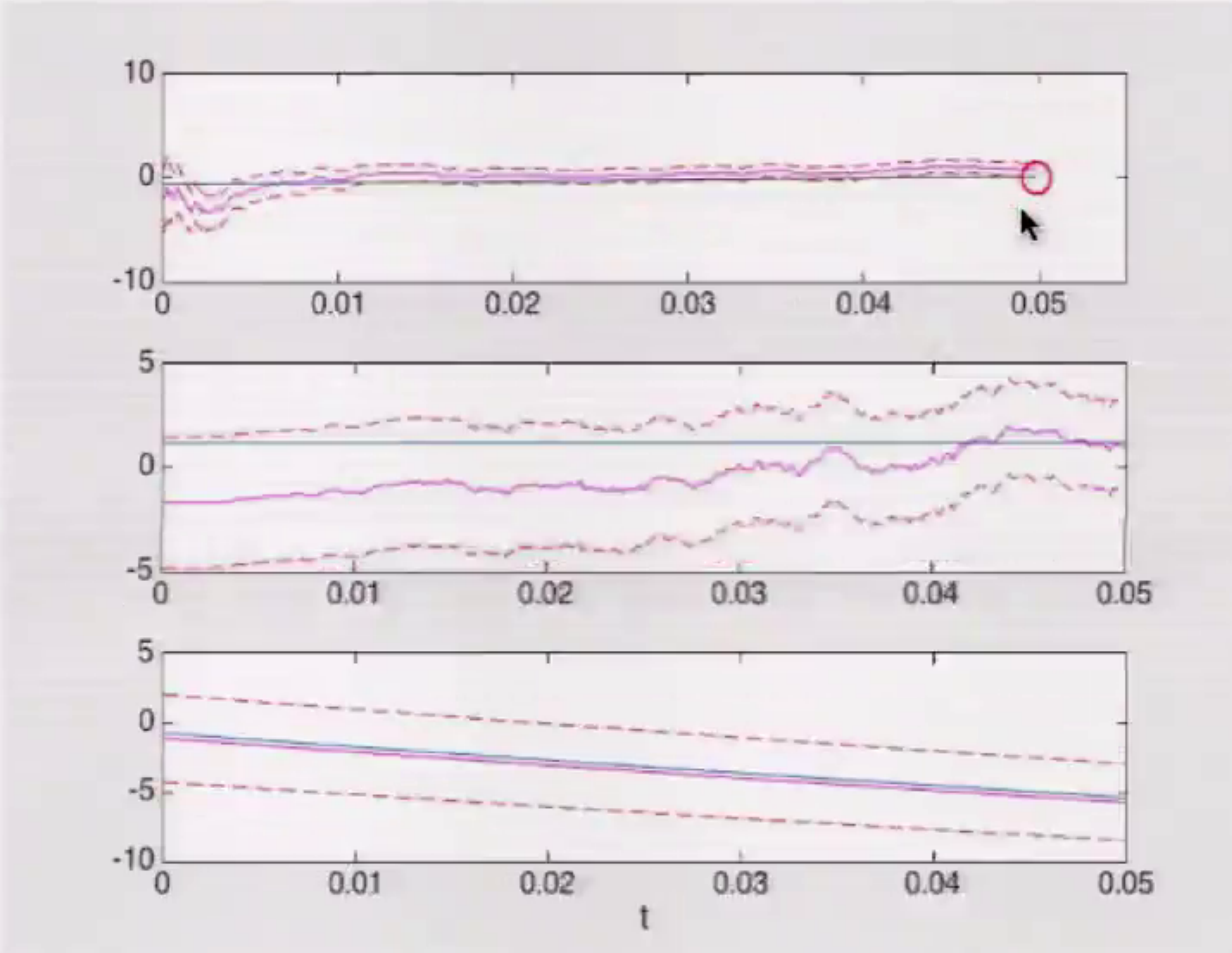
Define $Q = I - P$.

Key Idea: $Q\Psi$ should contract

Filtering Lorenz '63 (Observe First Component)



Filtering Lorenz '63 (Observe First Component)



Goal (Cerou [3], SIAM J. Cont. Opt. 2000)

Key Question: For which Ψ and P does the filter μ^j concentrate on the true signal, up to error ϵ, in the large-time limit?

Key Problem: Ψ may expand

Signal and Observation: Control Unpredictability?

Signal, **deterministic** (chaotic) dynamics on Hilbert space $(\mathcal{H}, \langle\cdot,\cdot\rangle, |\cdot|)$:

Signal Process

$$v_{j+1} = \Psi(v_j), \qquad v_0 \sim \mu_0.$$

Observations, **partial and noisy**, $P : \mathcal{H} \to \mathcal{H}$ is a **projection**.

Observation Process

$$y_{j+1} = Pv_{j+1} + \epsilon\xi_{j+1}, \qquad \mathbb{E}\xi_j = 0, \quad \mathbb{E}|\xi_j|^2 = 1, \quad \text{i.i.d.}$$

Filter: probability distribution of v_j given observations to time j:

Filter

$$\mu_j(A) = \mathbb{P}(v_j \in A|\mathcal{F}_j), \quad \mathcal{F}_j = \sigma(y_1, \dots, y_j).$$

Goal (Cerou [3], SIAM J. Cont. Opt. 2000)

Key Question: For which Ψ and P does the filter μ^j concentrate on the true signal, up to error ϵ, in the large-time limit?

Key Problem: Ψ may expand

Define $Q = I - P$.

Key Idea: $Q\Psi$ should contract

Running Example: Geophysical Applications

Consider the ODE for $v(\cdot) \in \mathcal{H}$, given $u, f \in \mathcal{H}$:

$$\frac{dv}{dt} + Av + B(v, v) = f, \quad v(0) = u.$$

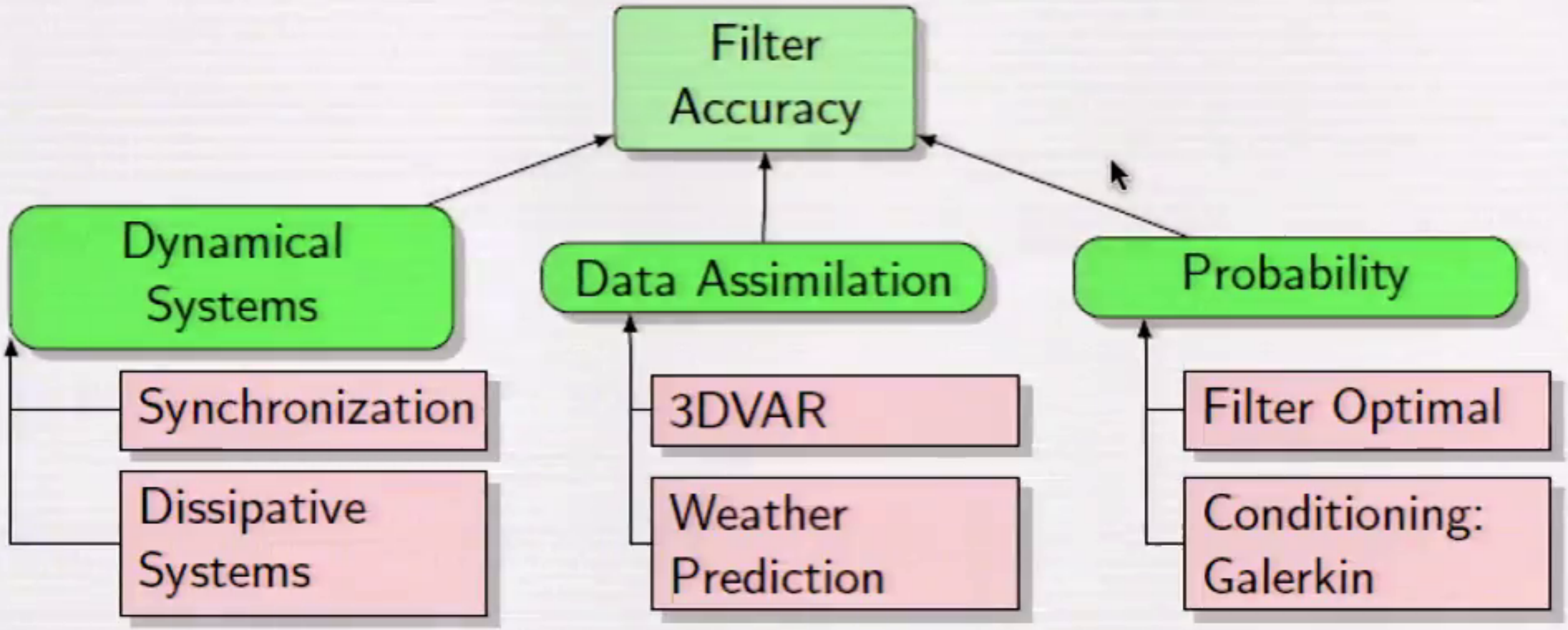
Assume dissipative with energy-conserving nonlinearity i.e. $\exists \lambda > 0$:

$$\langle Av, v \rangle \geq \lambda |v|^2, \quad \langle B(v, v), v \rangle = 0.$$

Dissipative semigroup Ψ denoting time h flow $\big(v_j = v(jh)\big)$:

$$v_{j+1} = \Psi(v_j), \quad v_0 = u.$$

Table of Contents


Filter Accuracy
Dynamical Systems
Data Assimilation
Probability
Synchronization
Dissipative Systems
3DVAR
Weather Prediction
Filter Optimal
Conditioning: Galerkin

Idea 1: Synchronization (Foias and Prodi [4], RSM Padova 1967 Pecora and Carroll [10], PRL 1990.)

Truth $v^\dagger = (p^\dagger, q^\dagger)$	Synchronization Filter $m = (p, q)$
$p^\dagger_{j+1} = P\Psi(p^\dagger_j, q^\dagger_j),$	$p_{j+1} = p^\dagger_{j+1},$
$q^\dagger_{j+1} = Q\Psi(p^\dagger_j, q^\dagger_j),$	$q_{j+1} = Q\Psi(p^\dagger_j, q_j);$
$---$	$---$
$v^\dagger_{j+1} = \Psi(v^\dagger_j),$	$m_{j+1} = Q\Psi(m_j) + p^\dagger_{j+1}.$

Idea 1: Synchronization (Foias and Prodi [4], RSM Padova 1967 Pecora and Carroll [10], PRL 1990.)

Truth $v^\dagger = (p^\dagger, q^\dagger)$ — **Synchronization Filter $m = (p, q)$**

$$p^\dagger_{j+1} = P\Psi(p^\dagger_j, q^\dagger_j), \qquad p_{j+1} = p^\dagger_{j+1},$$

$$q^\dagger_{j+1} = Q\Psi(p^\dagger_j, q^\dagger_j), \qquad q_{j+1} = Q\Psi(p^\dagger_j, q_j);$$

— — —

$$v^\dagger_{j+1} = \Psi(v^\dagger_j), \qquad m_{j+1} = Q\Psi(m_j) + p^\dagger_{j+1}.$$

Synchronization for various chaotic dynamical systems [5, 10, 11]:

$$|m_j - v^\dagger_j| \to 0, \text{ as } j \to \infty.$$

Idea 2: 3DVAR (Lorenc [9] Q. J. R. Met. Soc 1986)

3DVAR Filter. $|\cdot|_C = |C^{-\frac{1}{2}}\cdot|$.

$$m_{j+1} = \operatorname{argmin}_{m\in\mathcal{H}}\{|m-\Psi(m_j)|_C^2 + \epsilon^{-2}|y_{j+1} - Pm|_\Gamma^2\}.$$

Solve Variational Equations (with $C = \epsilon^2(\eta^{-2}\Gamma P + Q)$)

$$m_{j+1} = (I-K)\Psi(m_j) + Ky_{j+1}, \quad K = (1+\eta^2)^{-1}P,$$

Variance Inflation (from weather prediction) $\eta \ll 1$

$m_{j+1} = Q\Psi(m_j) + Py_{j+1}, \quad \eta = 0.$ **Synchronization Filter.**

Idea 3: Filter Optimality (Folklore, but see e.g. Williams ⋯)

Recall $\mathcal{F}_j = \sigma(y_1, \ldots, y_j)$ and define the mean of the filter:

$$\hat{v}_j := \mathbb{E}(v_j|\mathcal{F}_j) = \mathbb{E}^{\mu_j}(v_j).$$

Use Galerkin orthogonality wrt conditional expectation

For any $\mathcal{F}_j$ measurable m_j :

$$\mathbb{E}|v_j - \hat{v}_j|^2 \le \mathbb{E}|v_j - m_j|^2.$$

Take m_j from **3DVAR** to get bounds on the mean of the filter.
Similar bounds apply to the variance of the filter. (Not shown.)

Table of Contents

Assumptions

There are two equivalent Hilbert spaces $(\mathcal{H}, \langle\cdot,\cdot\rangle, |\cdot|)$ and $(\mathcal{V}, \langle\cdot,\cdot\rangle_{\mathcal{V}}, \|\cdot\|)$:

Assumption 1: Absorbing Ball Property

There is $R_0 > 0$ such that:

- for $B(R_0) := \{x \in \mathcal{H} : |x| \le R_0\}$, $\Psi(B(R_0)) \subset B(R_0)$;
- for any bounded set $S \subset \mathcal{H}$ $\exists J = J(S) : \Psi^J(S) \subset B(R_0)$.

Assumption 2: Squeezing Property

There is $\alpha(R_0) \in (0,1)$ such that, for all $u, v \in B(R_0)$,

$$\|Q(\Psi(u) - \Psi(v))\|^2 \le \alpha(\mathbf{R_0})\|u - v\|^2.$$

Theorem (Sanz-Alonso and S, 2014, [12])

Let Assumptions 1,2 hold. Then there is a constant $c > 0$ independent of the noise strength ϵ such that

$$\limsup_{j\to\infty} \mathbb{E}|v_j - \hat{v}_j|^2 \le c\epsilon^2$$

.

Idea of proof:

- Fix $m_0 \in B(R_0)$ and let $\mathcal{P}$ denote the $\mathcal{H}-$projection onto $B(R_0)$. Define the **modified 3DVAR**:

$$m_{j+1} = \mathcal{P}\big(Q\Psi(m_j) + y_{j+1}\big).$$

- Prove

$$\limsup_{j\to\infty} \mathbb{E}|v_j - m_j|^2 \le c\epsilon^2.$$

- Use the L^2 optimality of the filtering distribution.

Idea of proof (sketch, Ψ globally Lipschitz):

$$
\begin{aligned}
m_{j+1} &= Q\Psi(m_j) && + \overbrace{P\Psi(v_j) + \epsilon\xi_{j+1}}^{\mathbf{y_{j+1}}}, \\
v_{j+1} &= Q\Psi(v_j) && + P\Psi(v_j).
\end{aligned}
$$

Subtract and use independence plus contractivity of $Q\Psi$:

$$
\begin{aligned}
\mathbb{E}\|v_{j+1} - m_{j+1}\|^2 &= \mathbb{E}\|Q\left(\Psi(v_j) - \Psi(m_j)\right) - \epsilon\xi_{j+1}\|^2 \\
&\le \mathbb{E}\|Q\left(\Psi(v_j) - \Psi(m_j)\right)\|^2 + \epsilon^2\mathbb{E}\|\xi_{j+1}\|^2 \\
&\le \alpha\mathbb{E}\|v_j - m_j\|^2 + \epsilon^2\mathbb{E}\|\xi_{j+1}\|^2.
\end{aligned}
$$

Lorenz '63 (Hayden, Olson and Titi [5], Physica D 2011.)

$$\begin{aligned}
\frac{dv^{(1)}}{dt} + a(v^{(1)} - v^{(2)}) &= 0\\
\frac{dv^{(2)}}{dt} + av^{(1)} + v^{(2)} + v^{(1)}v^{(3)} &= 0\\
\underbrace{\frac{dv^{(3)}}{dt}}_{\frac{\mathbf{dv}}{\mathbf{dt}}} + \underbrace{bv^{(3)}}_{\mathbf{Av}} \underbrace{- v^{(1)}v^{(2)}}_{\mathbf{B(v,v)}} &= \underbrace{-b(r+a)}_{\mathbf{f}}
\end{aligned}$$

Standard parameter values: $(a, b, r) = (10, 8/3, 28)$.

Observation matrix

$$P := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Theory applicable with $\|\cdot\|^2 := |P\cdot|^2 + |\cdot|^2$: [5], [8].

Lorenz '96 (Law, Sanz-Alonso, Shukla and S [11], arXiv 2014.)

Consider the following system, subject to the periodicity boundary conditions $v_0 = v_{3J}, v_{-1} = v_{3J-1}, v_{3J+1} = v_1$:

$$\underbrace{\frac{dv^{(j)}}{dt}}_{\frac{\mathbf{dv}}{\mathbf{dt}}} + \underbrace{v^{(j)}}_{\mathbf{Av}} + \underbrace{v^{(j-1)}(v^{(j+1)} - v^{(j-2)})}_{\mathbf{B(v,v)}} = \underbrace{F}_{\mathbf{f}}, \quad j = 1, 2, \dots, 3J.$$

Observation matrix P: observe 2 out of every 3 points. Theory applicable with $\|\cdot\|^2 := |P\cdot|^2 + |\cdot|^2$: [11].

Lorenz '63 (Hayden, Olson and Titi [5], Physica D 2011.)

$$\begin{aligned}
\frac{dv^{(1)}}{dt} &+ a(v^{(1)} - v^{(2)}) && = 0\\
\frac{dv^{(2)}}{dt} &+ \quad av^{(1)} + v^{(2)} + v^{(1)}v^{(3)} && = 0\\
\underbrace{\frac{dv^{(3)}}{dt}}_{\frac{dv}{dt}} &+ \underbrace{\quad bv^{(3)}\quad}_{Av} \underbrace{- v^{(1)}v^{(2)}}_{B(v,v)} && = \underbrace{-b(r+a)}_{f}
\end{aligned}$$

Standard parameter values: $(a, b, r) = (10, 8/3, 28)$.

Observation matrix

$$P := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Theory applicable with $\|\cdot\|^2 := |P\cdot|^2 + |\cdot|^2$: [5], [8].

Lorenz '96 (Law, Sanz-Alonso, Shukla and S [11], arXiv 2014.)

Consider the following system, subject to the periodicity boundary conditions $v_0 = v_{3J}, v_{-1} = v_{3J-1}, v_{3J+1} = v_1$:

$$\underbrace{\frac{dv^{(j)}}{dt}}_{\frac{\mathbf{dv}}{\mathbf{dt}}} + \underbrace{v^{(j)}}_{\mathbf{Av}} + \underbrace{v^{(j-1)}(v^{(j+1)} - v^{(j-2)})}_{\mathbf{B(v,v)}} = \underbrace{F}_{\mathbf{f}}, \quad j = 1, 2, \ldots, 3J.$$

Observation matrix P: observe 2 out of every 3 points. Theory applicable with $\|\cdot\|^2 := |P\cdot|^2 + |\cdot|^2$: [11].

Navier-Stokes Equation on a 2D Torus

(Hayden, Olson and Titi [5], Physica D 2011.)

P_{leray} denotes the Leray projector:

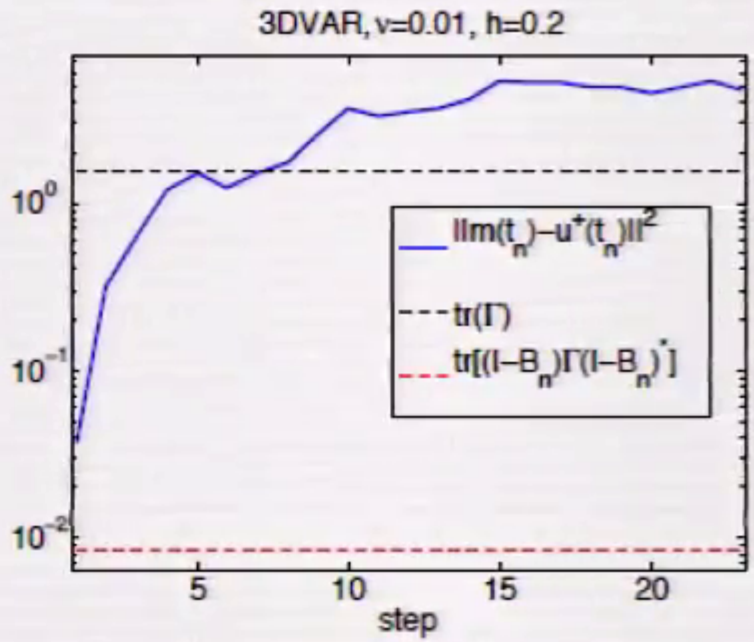
$$Au = -\nu P_{\text{leray}} \Delta u, \qquad B(u,v) = \frac{1}{2} P_{\text{leray}}[u \cdot \nabla v] + \frac{1}{2} P_{\text{leray}}[v \cdot \nabla u].$$

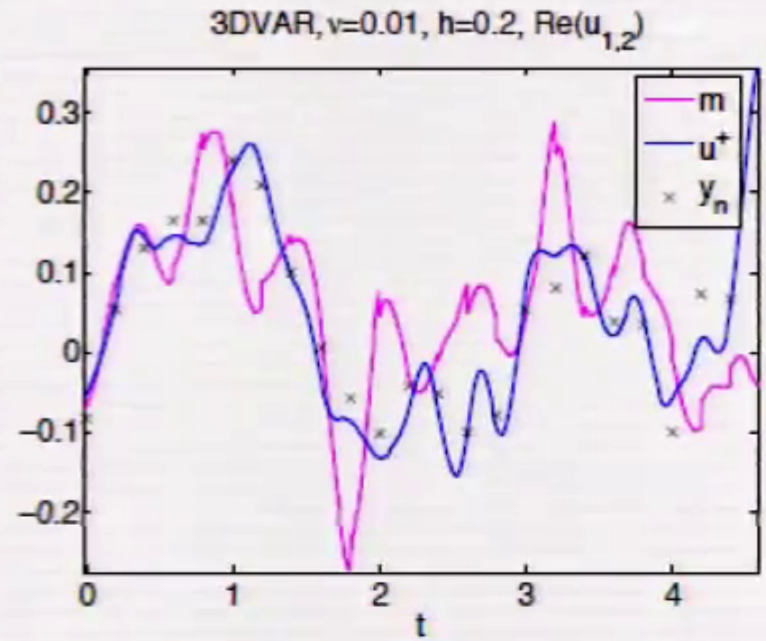
Observation operator in (divergence-free) Fourier space:

$$Pu = \sum_{|k| \leq k_{\max}} u_k e^{ikx}.$$

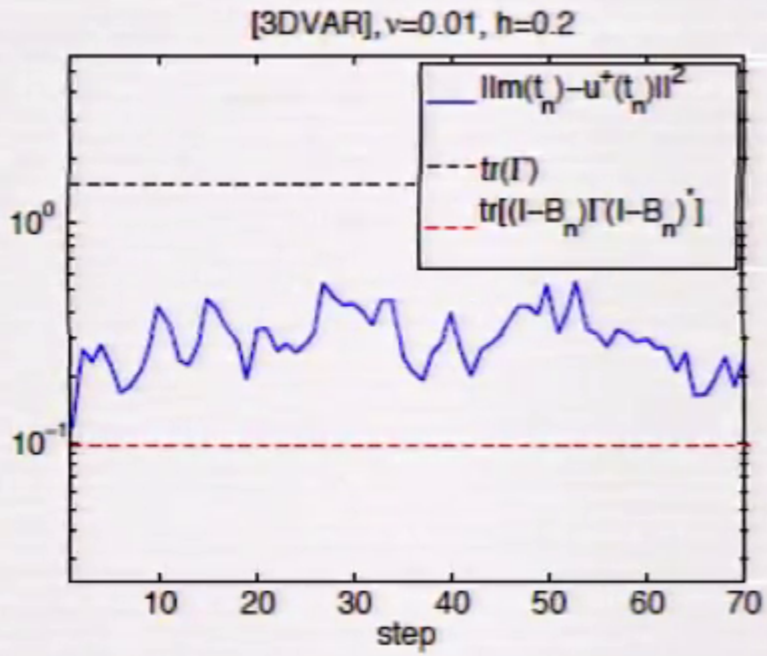
Theory applicable with $\mathcal{H} = \mathcal{V} := H^1_{\text{div}}(\mathbb{T}^2)$ and $k_{\max}$ sufficiently large/h sufficiently small: [2], [5].

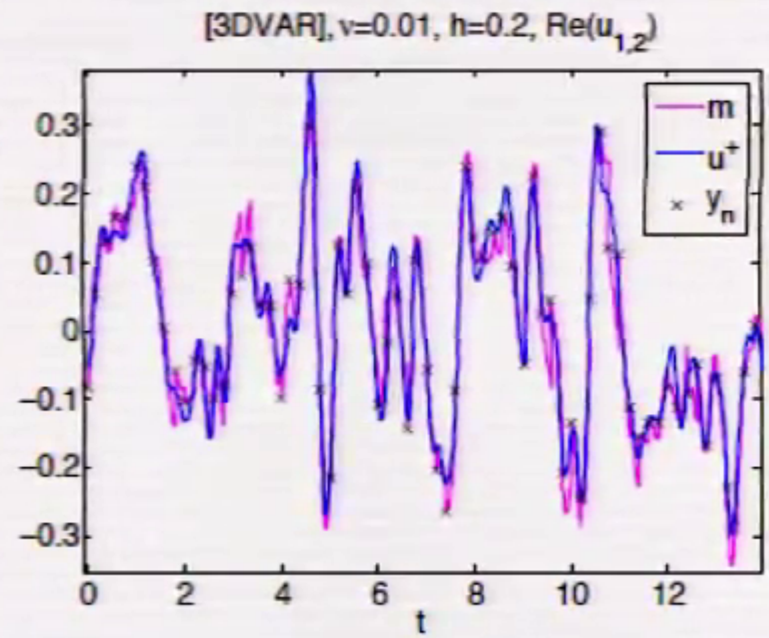
Inaccurate: η too large. (NSE torus) Law and S [7], Monthly Weather Review, 2012





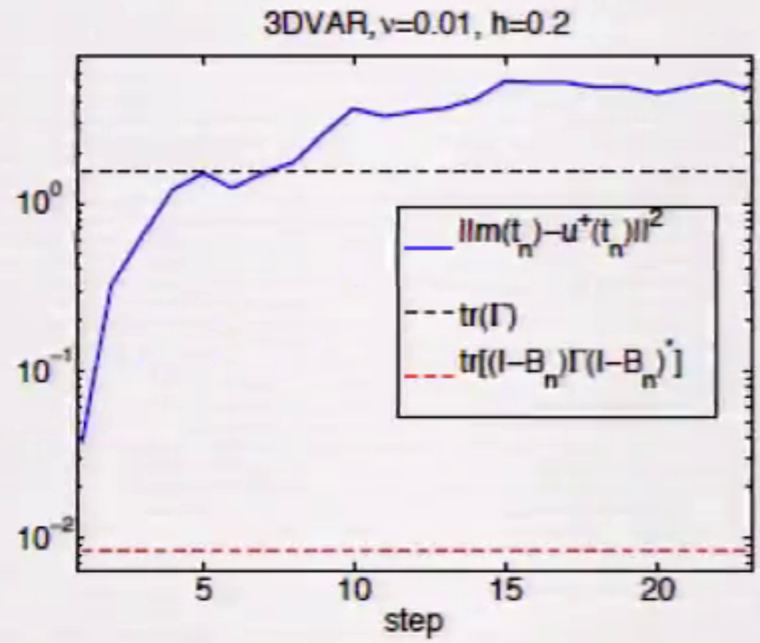
Accurate: smaller η. (NSE torus) Law and S [7], Monthly Weather Review, 2012

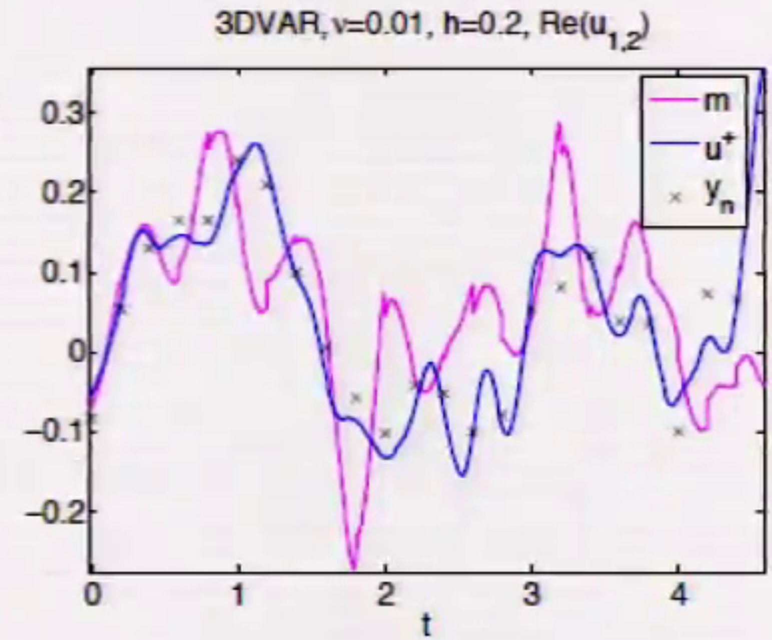




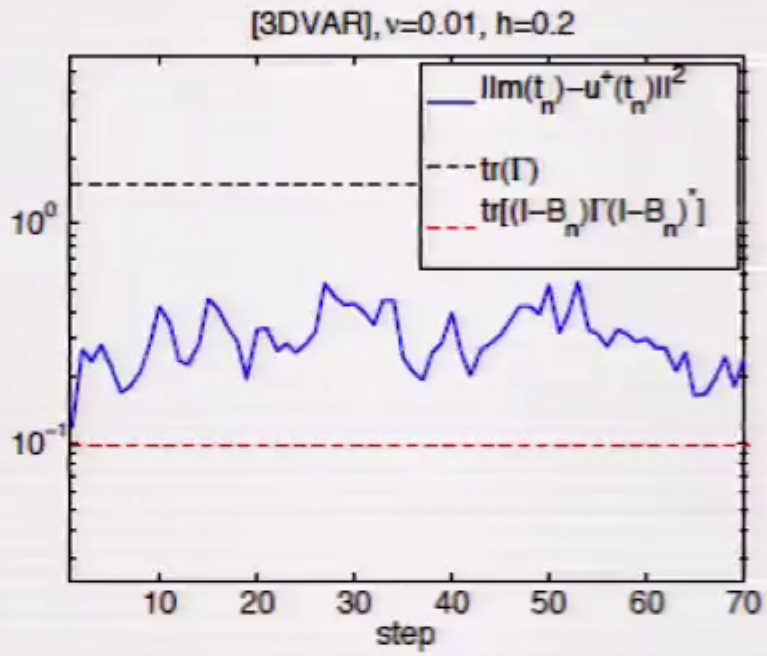
Inaccurate: η too large. (NSE torus)

Law and S [7], Monthly Weather Review, 2012





Accurate: smaller η. (NSE torus) Law and S [7], Monthly Weather Review, 2012



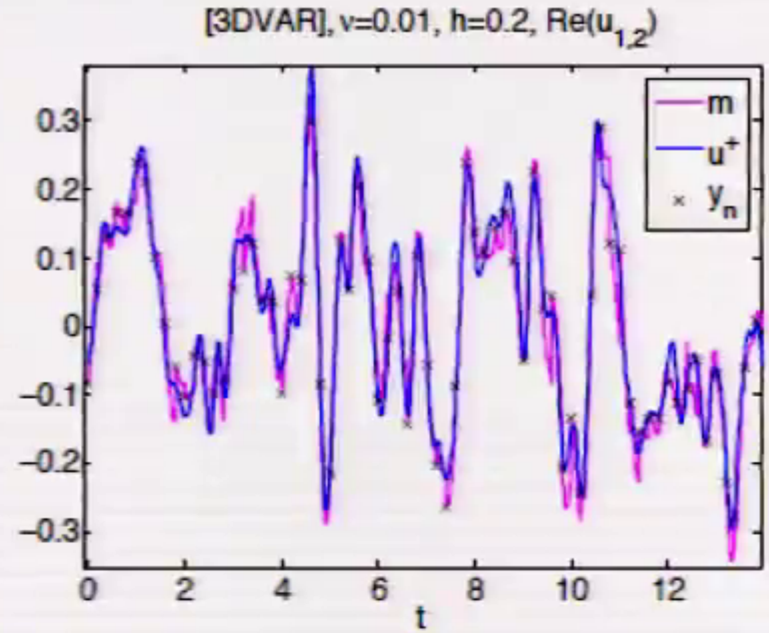


Table of Contents

S(P)DE Limits with Blömker 2012 [1], Kelly 2014 [6]; also Tong, Majda, Kelly 2015 [13] .

High Frequency Data Limit – 3DVAR

$$\frac{dm}{dt} + Am + B(m,m) + CP^*\Gamma^{-1}\Big(P(m-v) + \Gamma^{\frac{1}{2}}\frac{dW}{dt}\Big) = f$$

S(P)DE Limits with Blömker 2012 [1], Kelly 2014 [6]; also Tong, Majda, Kelly 2015 [13] .

High Frequency Data Limit – 3DVAR

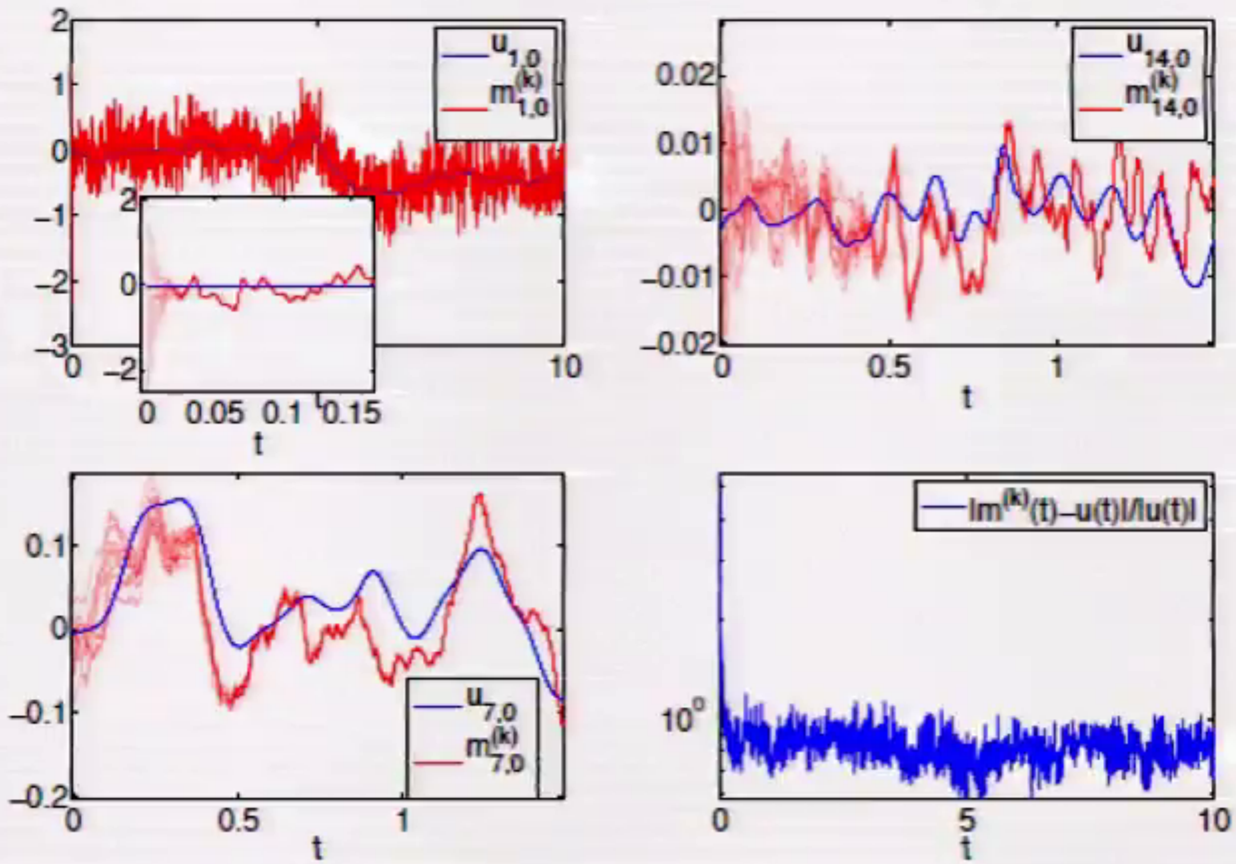
$$\frac{dm}{dt} + Am + B(m,m) + CP^*\Gamma^{-1}\Big(P(m-v) + \Gamma^{\frac{1}{2}}\frac{dW}{dt}\Big) = f$$

High Frequency Data Limit – Ensemble Kalman Filter

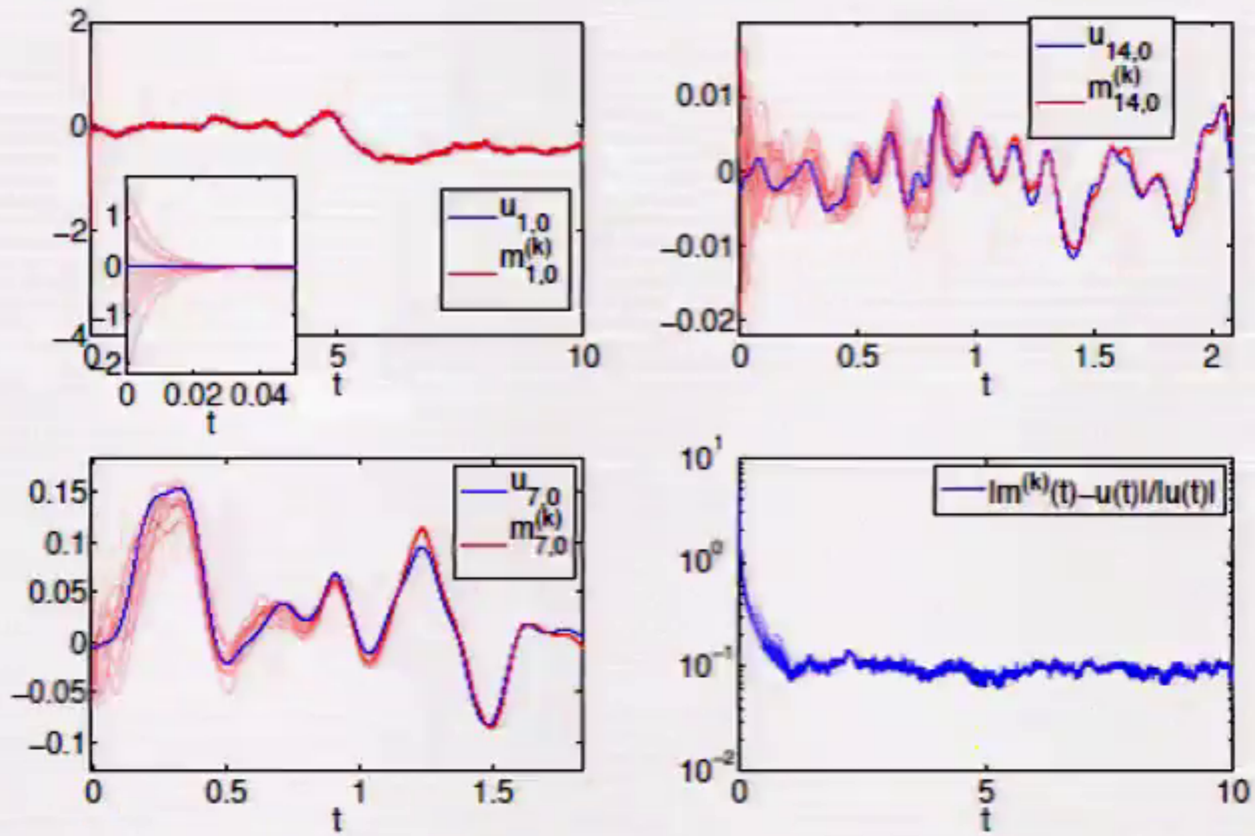
$$\frac{dm^{(j)}}{dt} + Am^{(j)} + B(m^{(j)}, m^{(j)}) + CP^*\Gamma^{-1}\Big(P(m^{(j)} - v) + \Gamma^{\frac{1}{2}}\frac{dW}{dt}\Big) = f,$$

$$\overline{m} = \frac{1}{J}\sum_{j=1}^{J} m^{(j)}, \quad C = C(m) = \frac{1}{J}\sum_{j=1}^{J}(m^{(j)} - \overline{m}) \otimes (m^{(j)} - \overline{m}).$$

SPDE Inaccurate (NSE Torus) (Blömker et al [1])



SPDE Accurate (NSE Torus) (Blömker et al [1])



Summary

- **Chaos** – and resulting **unpredictability** – is the enemy in many scientific and engineering applications.
- Its study has led to a great deal of interesting mathematics over the last century.
- **Data** – when combined with **models** – can have a massive positive impact on prediction in all of these scientific and engineering applications.
- The emerging new field, in which **model and data are analyzed simultaneously**, will lead to interesting new mathematics over the next century.
- **Data Assimilation** needs input from **Dynamical Systems**.

References I

D. Blömker, K.J.H. Law, A.M. Stuart, and K.C. Zygalakis.
Accuracy and stability of the continuous-time 3DVAR filter for the Navier-Stokes equation.
Nonlinearity, 26:2193–2219, 2013.

C.E.A. Brett, K.F. Lam, K.J.H. Law, D.S. McCormick, M.R. Scott, and A.M. Stuart.
Accuracy and Stability of Filters for the Navier-Stokes equation.
Physica D: Nonlinear Phenomena, 245:34–45, 2013.

F. Cérou.
Long time behavior for some dynamical noise free nonlinear filtering problems.
SIAM Journal on Control and Optimization, 38(4):1086–1101, 2000.

References II

C. Foias and G. Prodi.
Sur le comportment global des solutions non statiennaires des equations de Navier–Stokes en dimension 2.
Rend. Sem. Mat. Univ. Padova, 39, 1967.

K. Hayden, E. Olson, and E.S. Titi.
Discrete Data Assimilation in the Lorenz and 2D Navier–Stokes equations.
Physica D: Nonlinear Phenomena, 240:1416–1425, 2011.

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

K.Law and A.M. Stuart.
Evaluating data assimilation algorithms.
Monthly Weather Review, 140:3757–3782, 2012.

References III

K.J.H. Law, A. Shukla, and A.M. Stuart.
Analysis of the 3DVAR Filter for the Partially Observed Lorenz '63 Model.
Discrete and Continuous Dynamical Systems A, 34:1061–1078, 2014.

A.C. Lorenc.
Analysis methods for numerical weather prediction.
Quarterly Journal of the Royal Meteorological Society, 112(474):1177–1194, 1986.

L.M. Pecora and T.L. Carroll.
Synchronization in chaotic systems.
Physical review letters, 64(8):821, 1990.

D. Sanz-Alonso, K.J.H. Law, A. Shukla, and A.M. Stuart.
Filter accuracy for chaotic dynamical systems: fixed versus adaptive observation operators.
arxiv.org/abs/1411.3113, 2014.

References IV

D. Sanz-Alonso and A.M. Stuart.
Long-time asymptotics of the filtering distribution for partially observed chaotic deterministic dynamical systems.
arxiv.org/abs/1411.6510, 2014.

X.T. Tong, A.J. Majda, and D.T.B. Kelly.
Nonlinear stability and ergodicity of ensemble based Kalman filters.
NYU preprint 2015.