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)

EQUADIFF 2015
Lyon, France, July 6th 2015
Funded by EPSRC, ERC and ONR

Enabling Quantification of Uncertainty for Inverse Problems

THE UNIVERSITY OF WARWICK
# Table of Contents

1. **INTRODUCTION**

2. **THREE IDEAS**

3. **DISCRETE TIME: THEORY**

4. **CONTINUOUS TIME: DIFFUSION LIMITS**

5. **CONCLUSIONS**
Consider the following map on Hilbert space \((\mathcal{H}, \langle \cdot , \cdot \rangle, | \cdot |)\):

**Signal Dynamics**

\[ v_{j+1} = \Psi(v_j), \quad v_0 \sim \mu_0. \]
Consider the following map on Hilbert space \((\mathcal{H}, \langle \cdot, \cdot \rangle, |\cdot|)\):

**Signal Dynamics**

\[ v_{j+1} = \Psi(v_j), \quad v_0 \sim \mu_0. \]

Assume **dissipativity**:

**Absorbing Set**

Compact \(B\) in \(\mathcal{H}\) with the property that, for \(|v_0| \leq R\), there is \(J = J(R) > 0\) such that, for all \(j \geq J\), \(v_j \in B\).
Consider the following map on Hilbert space $(\mathcal{H}, \langle \cdot, \cdot \rangle, | \cdot |)$:

**Signal Dynamics**

\[ v_{j+1} = \Psi(v_j), \quad v_0 \sim \mu_0. \]

Assume **dissipativity**:

**Absorbing Set**

Compact $B$ in $\mathcal{H}$ with the property that, for $|v_0| \leq R$, there is $J = J(R) > 0$ such that, for all $j \geq J$, $v_j \in B$.

**Limited predictability**:

**Global Attractor**

\[ d(v_j, \mathcal{A}) \to 0, \text{ as } j \to \infty. \]
Signal and Observation

Random initial condition:

Signal Process

\[ v_{j+1} = \Psi(v_j), \quad v_0 \sim \mu_0. \]
Signal and Observation

**Random initial condition:**

**Signal Process**

\[ v_{j+1} = \Psi(v_j), \quad v_0 \sim \mu_0. \]

**Observations, partial and noisy, \( P : \mathcal{H} \to \mathbb{R}^J :**

**Observation Process**

\[ y_{j+1} = P v_{j+1} + \epsilon \xi_{j+1}, \quad \mathbb{E} \xi_j = 0, \quad \mathbb{E} |\xi_j|^2 = 1, \text{ i.i.d. w/pdf } \rho. \]
Signal and Observation

Random initial condition:

Signal Process

\[ v_{j+1} = \Psi(v_j), \quad v_0 \sim \mu_0. \]

Observations, partial and noisy, \( P : \mathcal{H} \to \mathbb{R}^J \):

Observation Process

\[ y_{j+1} = P v_{j+1} + \epsilon \xi_{j+1}, \quad \mathbb{E} \xi_j = 0, \quad \mathbb{E} |\xi_j|^2 = 1, \text{ i.i.d. w/pdf } \rho. \]

Filter: probability distribution of \( v_j \) given observations to time \( j \):

Filter

\[ \mu_j(A) = \mathbb{P}(v_j \in A | \mathcal{F}_j), \quad \mathcal{F}_j = \sigma(y_1, \ldots, y_j). \]
Signal and Observation: Control Unpredictability?

**Pushforward under dynamics:**

**Signal Process**

\[ \hat{\mu}_{j+1} = \Psi \ast \mu_j. \]
Signal and Observation: Control Unpredictability?

**Pushforward under dynamics:**

**Signal Process**

\[ \hat{\mu}_{j+1} = \Psi \ast \mu_j. \]

**Observation Process**

\[ \mu_{j+1}(A) = \frac{\int_A \rho(\epsilon^{-1}(y_{j+1} - P\nu)) \hat{\mu}_{j+1}(dv)}{\int_\mathcal{H} \rho(\epsilon^{-1}(y_{j+1} - P\nu)) \hat{\mu}_{j+1}(dv)}. \]
Signal and Observation: Control Unpredictability?

**Pushforward under dynamics:**

**Signal Process**

\[ \hat{\mu}_{j+1} = \Psi \ast \mu_j. \]

**Observation Process**

\[ \mu_{j+1}(A) = \frac{\int_A \rho(\epsilon^{-1}(y_{j+1} - P\nu)) \hat{\mu}_{j+1}(d\nu)}{\int_{\mathcal{H}} \rho(\epsilon^{-1}(y_{j+1} - P\nu)) \hat{\mu}_{j+1}(d\nu)}. \]

**When is the filter predictable:**

**Filter Accuracy**

\[ \mu_j \approx \delta_{\nu^*_j} \text{ as } j \to \infty. \]

**Key Question:** For which $\Psi$ and $P$ does the filter $\mu^j$ concentrate on the true signal, up to error $\epsilon$, in the large-time limit?
Key Question: For which $\Psi$ and $P$ does the filter $\mu^j$ concentrate on the true signal, up to error $\epsilon$, in the large-time limit?

Key Problem: $\Psi$ may expand

**Key Question:** For which $\Psi$ and $P$ does the filter $\mu^j$ concentrate on the true signal, up to error $\epsilon$, in the large-time limit?

**Key Problem:** $\Psi$ may expand

View $P$ as a **projection** on $\mathcal{H}$. Define $Q = I - P$.

**Key Idea:** $Q\Psi$ should contract
A Large Class of Examples

Geophysical Applications

\[ \frac{dv}{dt} + Au + B(u, u) = f. \]

Dissipative with energy conserving nonlinearity

- \( \exists \lambda > 0 : \langle Av, v \rangle \geq \lambda |v|^2. \)
- \( \langle B(v, v), v \rangle = 0. \)
- \( f \in L^2_{\text{loc}}(\mathbb{R}^+; \mathcal{H}). \)

Examples

- Lorenz '63
- Lorenz '96
- Incompressible 2D Navier-Stokes equation on a torus
INTRODUCTION
THREE IDEAS
DISCRETE TIME: THEORY
CONTINUOUS TIME: DIFFUSION LIMITS
CONCLUSIONS
INTRODUCTION  THREE IDEAS  DISCRETE TIME: THEORY  CONTINUOUS TIME: DIFFUSION LIMITS  CONCLUSIONS

Filter  Accuracy

Dynamical Systems

Synchronization

Dissipative Systems

Data Assimilation

3DVAR

Weather Prediction

Probability

Filter Optimal

Conditioning: Galerkin
Idea 1: Synchronization  
(Foias and Prodi [7], RSM Padova 1967 Pecora and Carroll [13], PRL 1990.)

<table>
<thead>
<tr>
<th>Truth $v^\dagger = (p^\dagger, q^\dagger)$</th>
<th>Synchronization Filter $m = (p, q)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_{j+1}^\dagger = P\psi(p_j^\dagger, q_j^\dagger)$,</td>
<td>$p_{j+1} = p_{j+1}^\dagger$,</td>
</tr>
<tr>
<td>$q_{j+1}^\dagger = Q\psi(p_j^\dagger, q_j^\dagger)$,</td>
<td>$q_{j+1} = Q\psi(p_j^\dagger, q_j)$;</td>
</tr>
<tr>
<td>$v_{j+1}^\dagger = \psi(v_j^\dagger)$,</td>
<td>$m_{j+1} = Q\psi(m_j) + p_{j+1}^\dagger$.</td>
</tr>
</tbody>
</table>
### Idea 1: Synchronization

(Foias and Prodi [7], RSM Padova 1967 Pecora and Carroll [13], PRL 1990.)

<table>
<thead>
<tr>
<th>Truth $v^\dagger = (p^\dagger, q^\dagger)$</th>
<th>Synchronization Filter $m = (p, q)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p^\dagger_{j+1} = P\Psi(p^\dagger_j, q^\dagger_j)$,</td>
<td></td>
</tr>
<tr>
<td>$q^\dagger_{j+1} = Q\Psi(p^\dagger_j, q^\dagger_j)$,</td>
<td></td>
</tr>
<tr>
<td>$v^\dagger_{j+1} = \Psi(v^\dagger_j)$,</td>
<td></td>
</tr>
<tr>
<td>$m_{j+1} = Q\Psi(m_j) + p^\dagger_{j+1}$.</td>
<td></td>
</tr>
<tr>
<td>$p_{j+1} = p^\dagger_{j+1}$,</td>
<td></td>
</tr>
<tr>
<td>$q_{j+1} = Q\Psi(p^\dagger_j, q_j)$;</td>
<td></td>
</tr>
</tbody>
</table>

Synchronization for various chaotic dynamical systems (including the three canonical examples above [8, 13, 4, 14]):

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

Cycled 3DVAR Filter.  \( | \cdot |_A = |A^{-\frac{1}{2}} \cdot |. \)

\[
m_{j+1} = \arg\min_{m \in \mathcal{H}} \{ |m - \Psi(m_j)|^2_C + \epsilon^{-2} |y_{j+1} - Pm|^2 \}.
\]

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) \( \eta \ll 1 \)

\[
m_{j+1} = Q\Psi(m_j) + Py_{j+1}, \quad \eta = 0. \quad \text{Synchronization Filter.}
\]
Inaccurate: $\eta$ too large. (NSE torus) Law and S [10], Monthly Weather Review, 2012

3DVAR, $\nu=0.01$, $h=0.2$

$||m(t_n) - u^+(t_n)||^2$

$\text{tr} (\Gamma)$

$\text{tr} [(I-B_n)\Gamma(I-B_n)^*]$
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 \( m_j \):

\[
\mathbb{E}|v_j - \hat{v}_j|^2 \leq \mathbb{E}|v_j - m_j|^2.
\]

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.)
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 $B(R_0) := \{x \in \mathcal{H} : |x| \leq R_0\}$, $\Psi(B(R_0)) \subset B(R_0)$;
- 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 $u, v \in B(R_0)$,

$$\|Q(\Psi(u) - \Psi(v))\|^2 \leq \alpha(R_0)\|u - v\|^2.$$
Theorem (Sanz-Alonso and S, 2014, [15])

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

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

Idea of proof:

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

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

- Prove

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

- Use the $L^2$ optimality of the filtering distribution.
Idea of proof (sketch, $\Psi$ globally Lipschitz):

\[
\begin{align*}
m_{j+1} &= Q\Psi(m_j) + P\Psi(v_j) + \epsilon \xi_{j+1}, \\
v_{j+1} &= Q\Psi(v_j) + P\Psi(v_j).
\end{align*}
\]

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

\[
\mathbb{E}\|v_{j+1} - m_{j+1}\|^2 = \mathbb{E}\|Q(\Psi(v_j) - \Psi(m_j)) - \epsilon \xi_{j+1}\|^2
\]

\[
\leq \mathbb{E}\|Q(\Psi(v_j) - \Psi(m_j))\|^2 + \epsilon^2 \mathbb{E}\|\xi_{j+1}\|^2
\]

\[
\leq \alpha \mathbb{E}\|v_j - m_j\|^2 + \epsilon^2 \mathbb{E}\|\xi_{j+1}\|^2.
\]

Use Gronwall.
Lorenz '63 (uses noiseless synchronization filter analysis in Hayden, Olson and Titi [8], Physica D 2011.)

\[
\begin{align*}
\frac{dv^{(1)}}{dt} + a(v^{(1)} - v^{(2)}) &= 0 \\
\frac{dv^{(2)}}{dt} + av^{(1)} + v^{(2)} + v^{(1)}v^{(3)} &= 0 \\
\frac{dv^{(3)}}{dt} + bv^{(3)} - v^{(1)}v^{(2)} &= -b(r + a)
\end{align*}
\]

Observation matrix

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

Theory applicable with \(\| \cdot \|^2 := |P \cdot|^2 + |\cdot|^2\) for \(h\) sufficiently small: [8], [11].
Consider the following system, subject to the periodicity boundary conditions \( v_0 = v_{3J}, \ v_{-1} = v_{3J-1}, \ v_{3J+1} = v_1 \):

\[
\frac{dv^{(j)}}{dt} + v^{(j)} + v^{(j-1)}(v^{(j+1)} - v^{(j-2)}) = F, \quad j = 1, 2, \ldots, 3J.
\]

Observation matrix \( P \): observe 2 out of every 3 points. Theory applicable with \( \| \cdot \|^2 := |P \cdot|^2 + |\cdot|^2 \) for \( h \) sufficiently small: [14].
$P_{\text{leray}}$ denotes the Leray projector:

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

Observation operator in (divergence-free) Fourier space:

$$Pu = \sum_{|k| \leq k_{\text{max}}} u_k \frac{k^\perp}{|k|} e^{ik \cdot x}.$$

Theory applicable with $\mathcal{H} = \mathcal{V} := H^1_{\text{div}} (\mathbb{T}^2)$ and $k_{\text{max}}$ sufficiently large/$h$ sufficiently small: [3], [8].
### Observations control unpredictability in these cases:

<table>
<thead>
<tr>
<th>ODE</th>
<th>Dimension of $\nu$</th>
<th>Rank($P$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lorenz '63</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Lorenz '96</td>
<td>3J</td>
<td>2J</td>
</tr>
<tr>
<td>NSE on torus</td>
<td>$\infty$</td>
<td>Finite</td>
</tr>
</tbody>
</table>
Table of Contents

1. INTRODUCTION
2. THREE IDEAS
3. DISCRETE TIME: THEORY
4. CONTINUOUS TIME: DIFFUSION LIMITS
5. CONCLUSIONS

**Ensemble Kalman Filter.** For $n = 1, \ldots, N$:

$$v_{j+1}^{(n)} = \arg\min_{v \in \mathcal{H}} \{ |v - \Psi(v_j^{(n)})|^2 C_j + \epsilon^{-2} |y_{j+1}^{(n)} - Pv|^2 \}.$$

**Empirical Covariance**

$$\bar{v}_j = \frac{1}{N} \sum_{n=1}^{N} v_j^{(n)}, \quad C_j = \frac{1}{N} \sum_{n=1}^{N} (v_j^{(n)} - \bar{v}_j) \otimes (v_j^{(n)} - \bar{v}_j).$$

**Perturbed Observations**

$$y_{j+1}^{(n)} = y_{j+1} + \epsilon \xi_j^{(n)}, \quad \xi_j^{(n)} \text{ i.i.d. w/pdf } \rho.$$

High Frequency Data Limit – 3DVAR

\[
\frac{dm}{dt} + Am + B(m, m) + CP^* \Gamma^{-1} \left( P(m - v) + \epsilon \Gamma^{\frac{1}{2}} \frac{dW}{dt} \right) = f
\]

**High Frequency Data Limit – 3DVAR**

\[
\frac{dm}{dt} + Am + B(m, m) + CP^*\Gamma^{-1}\left(P(m - v) + \epsilon \Gamma^{1/2} \frac{dW}{dt}\right) = f
\]

**High Frequency Data Limit – Ensemble Kalman Filter**

\[
\frac{d\mathbf{v}^{(n)}}{dt} + A\mathbf{v}^{(n)} + B(\mathbf{v}^{(n)}, \mathbf{v}^{(n)}) + CP^*\Gamma^{-1}\left(P(\mathbf{v}^{(n)} - \mathbf{v}) + \epsilon \Gamma^{1/2} \frac{dW^{(n)}}{dt}\right) = f,
\]

\[
\bar{v} = \frac{1}{J} \sum_{j=1}^{J} v^{(n)}, \quad C = \frac{1}{J} \sum_{j=1}^{J} (v^{(n)} - \bar{v}) \otimes (v^{(n)} - \bar{v}).
\]
S(P)DE Accuracy see also Azouani, Olson and Titi 2014 [1] and Tong, Majda, Kelly 2015 [16].

Theorem (3DVAR Accurate, with Blömker 2012 et al [2])

Under similar assumptions to the discrete case there is a constant $c > 0$ independent of the noise strength $\epsilon$ such that

$$\limsup_{t \to \infty} \mathbb{E} |v - m|^2 \leq c \epsilon^2$$

Theorem (EnKF Well-Posed, with Kelly et al [9])

Let Assumptions 1 hold and $P=I$. Then there is a constant $c > 0$ independent of the noise strength $\epsilon$ such that

$$\sup_{t \in [0,T]} \sum_{n=1}^{N} \mathbb{E} |v^{(n)}(t)|^2 \leq C(T)(1 + \mathbb{E} |v^{(n)}(0)|^2).$$
SPDE Inaccurate (NSE Torus) (Blömker et al [2])
SPDE Accurate (NSE Torus) (Blömker et al [2])
Summary

- **Chaos** – and resulting **unpredictability** – is the enemy in many scientific and engineering applications.
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.
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.
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.
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

A. Azouani, E. Olson, and E.S. Titi.
Continuous data assimilation using general interpolant observables.

Accuracy and stability of the continuous-time 3DVAR filter for the
Navier-Stokes equation.

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.

A. Carrassi, M. Ghil, A. Trevisan, and F. Uboldi.
Data assimilation as a nonlinear dynamical systems problem: Stability and
cvergence of the prediction-assimilation system.
F. Cérou.
Long time behavior for some dynamical noise free nonlinear filtering problems.

G. Evensen.
Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics.

C. Foias and G. Prodi.
Sur le comportment global des solutions non statiennaires des equations de Navier–Stokes en dimension 2.

K. Hayden, E. Olson, and E.S. Titi.
Discrete Data Assimilation in the Lorenz and 2D Navier–Stokes equations.
References III

Well-posedness and accuracy of the ensemble Kalman filter in discrete and continuous time.

K.Law and A.M. Stuart.
Evaluating data assimilation algorithms.

Analysis of the 3DVAR Filter for the Partially Observed Lorenz ’63 Model.

A.C. Lorenc.
Analysis methods for numerical weather prediction.
References IV

L.M. Pecora and T.L. Carroll.
Synchronization in chaotic systems.

Filter accuracy for chaotic dynamical systems: fixed versus adaptive observation operators.

D. Sanz-Alonso and A.M. Stuart.
Long-time asymptotics of the filtering distribution for partially observed chaotic deterministic dynamical systems.

X.T. Tong, A.J. Majda, and D.T.B. Kelly.
Nonlinear stability and ergodicity of ensemble based Kalman filters.
*NYU preprint 2015*. 
Matlab files and book chapters freely available:  
http://tiny.cc/damat  
http://www2.warwick.ac.uk/fac/sci/maths/people/staff/andrew_stuart/