Combining shadowing and synchronization for data assimilation

Bart de Leeuw

CWI & Univ. Utrecht

May 19, 2017

In collaboration with

Svetlana Dubinkina (Centrum Wiskunde & Informatica)
Jason Frank (Utrecht University)
Andrew Steyer (Sandia National Laboratories)
Xuemin Tu (University of Kansas)
Erik Van Vleck (University of Kansas)

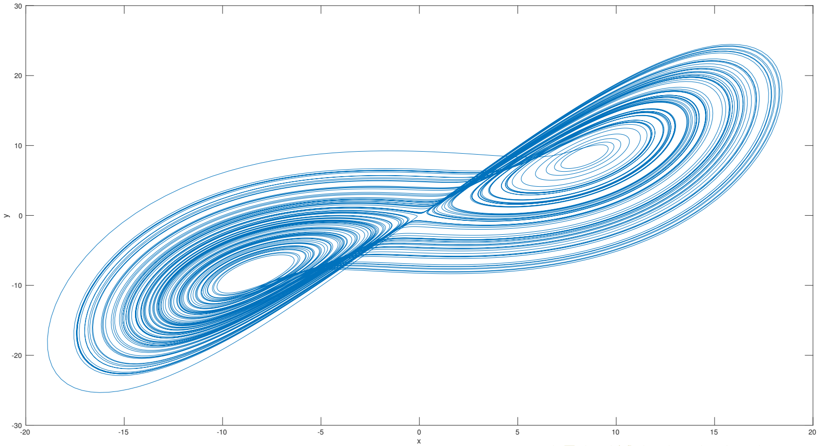
Dynamical systems and flows

$$\dot{x} = f(x)$$

$$x_{n+1} = F(x_n)$$

Lorenz '63 model

$\dot{x} = \sigma(y - z), \qquad \dot{y} = x(\rho - z) - y, \qquad \dot{z} = xy - \beta z.$ (Lorenz, 1963)
Parameters: $\sigma = 10$, $\beta = \frac{8}{3}$ and $\rho = 28$.



What is data assimilation

$$
\begin{aligned}
\text{Model} \qquad & x_{n+1} && = F(x_n), \quad x_n \in \mathbb{R}^d, \\
\text{Observations} \qquad & y_n && = H(X_n) + \eta_n, \quad y_n \in \mathbb{R}^{b \leq d}, \quad n = -M, \ldots, N.
\end{aligned}
$$

What is data assimilation

$$\begin{aligned} &\text{Model} && x_{n+1} &&= F(x_n), \quad x_n \in \mathbb{R}^d, \\ &\text{Observations} && y_n &&= H(X_n) + \eta_n, \quad y_n \in \mathbb{R}^{b \le d}, \quad n = -M, \dots, N. \end{aligned}$$

Data assimilation problem

Find $\boldsymbol{u} = \{u_{-M}, u_{-M+1}, \dots, u_N\}$, $u_n \in \mathbb{R}^d$, with small residual $\|y_n - H(u_n)\|$, $n = -M, \dots, N$ and mismatch $\|G(\boldsymbol{u})\| = \| \begin{pmatrix} G_{-M} & G_{-M+1} & \dots & G_{N-1} \end{pmatrix}^T \|$, with $G_n(\boldsymbol{u}) = u_{n+1} - F(u_n)$.

What is data assimilation

$$\begin{aligned} &\text{Model} && x_{n+1} = F(x_n), \quad x_n \in \mathbb{R}^d, \\ &\text{Observations} && y_n = H(X_n) + \eta_n, \quad y_n \in \mathbb{R}^{b \le d}, \quad n = -M, \dots, N. \end{aligned}$$

Data assimilation problem

Find $\boldsymbol{u} = \{u_{-M}, u_{-M+1}, \dots, u_N\}$, $u_n \in \mathbb{R}^d$, with small residual $\|y_n - H(u_n)\|$, $n = -M, \dots, N$ and mismatch $\|G(\boldsymbol{u})\| = \| \begin{pmatrix} G_{-M} & G_{-M+1} & \dots & G_{N-1} \end{pmatrix}^T \|$, with $G_n(\boldsymbol{u}) = u_{n+1} - F(u_n)$.

- $\boldsymbol{u}$ is the best approximation to the truth $\boldsymbol{X}$ given the available data.
- Example application: weather prediction: relatively accurate chaotic models with uncertain initial condition

How well can we do?

Non-chaotic model: assimilation can become arbitrarily good as the number of observations is increased.

How well can we do?

Non-chaotic model: assimilation can become arbitrarily good as the number of observations is increased.
Chaotic model with observations extending to past infinity: states connected by the unstable manifold are indistinguishable if noise is unbounded (f.e. Gaussian) (Judd & Smith, 2001)
Chaotic model with bi-infinite sequence of observations: no algorithm can reconstruct the true trajectory consistently (technically: there is no measurable function from observation space to $\mathbb{R}^d$ that equals x_0 with probability 1) (Lalley, 1999)

How well can we do?

Non-chaotic model: assimilation can become arbitrarily good as the number of observations is increased.
Chaotic model with observations extending to past infinity: states connected by the unstable manifold are indistinguishable if noise is unbounded (f.e. Gaussian) (Judd & Smith, 2001)
Chaotic model with bi-infinite sequence of observations: no algorithm can reconstruct the true trajectory consistently (technically: there is no measurable function from observation space to $\mathbb{R}^d$ that equals x_0 with probability 1) (Lalley, 1999)
Conclusion: for chaotic models, we will not be able to reconstruct the truth. A method finding one solution can at best aim for a likely solution. The best we can possibly do is constructing a probability density function over all possible solutions.

Shadowing lemma

If $||G_n(\boldsymbol{u})|| \leq \epsilon$ for all n, then $\boldsymbol{u}$ is called an ϵ-orbit.

Shadowing lemma

In a neighbourhood of a hyperbolic set of the map F, for every $\delta > 0$ there exists $\epsilon > 0$ such that for all ϵ-orbits $\boldsymbol{u}$ there exists an orbit $\boldsymbol{X}$ with $G(\boldsymbol{X}) = 0$ and $||u_n - X_n|| < \delta$ for all n.

- If the is sufficiently good, applying Newton's method starting at $\boldsymbol{u}$ yields $\boldsymbol{X}$.
- With larger errors, convergence to another indistinguishable solution is possible
- Note there is no special point in time: all times being equal stabilizes the method

Pseudo-orbit data assimilation

Shadowing method for data assimilation, based on minimization of $||G_n(\boldsymbol{u})||^2$ by gradient descent.

- In implementation a fixed number of steps is taken and minimization does not converge, hence pseudo-orbits.
- Orbits are obtained by taking the mid-point of the pseudo-orbit and propagating forward in time.
- Can be turned into an importance sampler.
- References: (Judd&Smith, 2001; Bröcker&Parlitz, 2001) etc.

Newton's method for data assimilation

Analogous to shadowing, we apply Newton iterations (starting with observations as initial condition) to solve

$$G(\boldsymbol{u}^{(k)}) \to 0, \qquad \boldsymbol{u}^{(0)} = \boldsymbol{y}.$$

We use the Lorenz '63 model, with noise covariance $E = \sigma^2 I$, $\sigma^2 = 1$ and consider only one window of length 10.

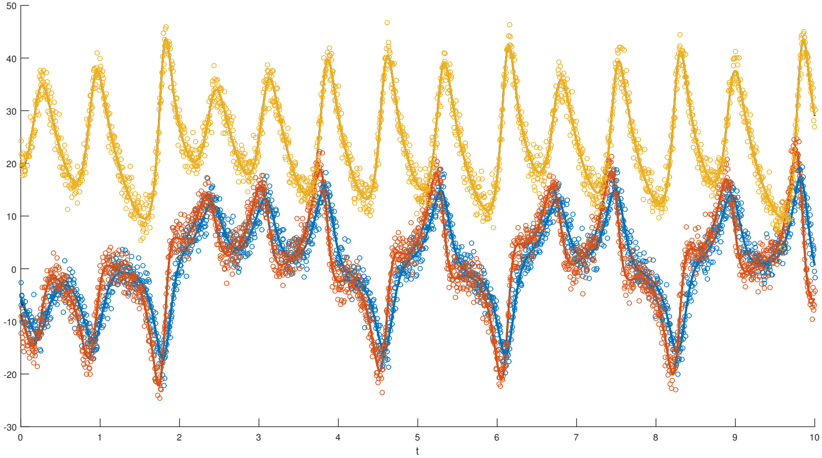


Figure: Different components of the solution are in different colours, with observations as small circles in the corresponding colours. The true trajectory is in black, but the differences between truth and assimilation are so small there is almost complete overlap between truth and assimilation.

Coupling two systems

Driver system	$x_{n+1} = F_n(x_n),$
Receiver system	$z_{n+1} = P_n F_n(x_n) + (I - P_n) F_n(z_n).$

Define $w_n = z_n - x_n$; the transverse dynamics is given by

$$\begin{aligned} w_{n+1} &= P_n F_n(x_n) + (I - P_n) F_n(z_n) - F_n(x_n), \\ &= (I - P_n) \left[F_n(x_n + w_n) - F_n(x_n) \right], \\ &\approx (I - P_n) F_n{}'(x_n) w_n. \end{aligned}$$

Goal: Choose the matrices P_n such that $w_n \to 0$.
Necessary: Conditional Lyapunov exponents of the receiver all nonpositive (Pecora & Carroll, 1991)

Synchronization in chaotic systems

$$\text{(Lorenz, 1963)}: \quad \dot{x} = \sigma(y-z), \quad \dot{y} = x(\rho - z) - y, \quad \dot{z} = xy - \beta z.$$
$$\text{(Pecora\&Caroll, 1990)}: \quad \dot{Y} = x(\rho - Z) - Y, \quad \dot{Z} = xY - \beta Z.$$

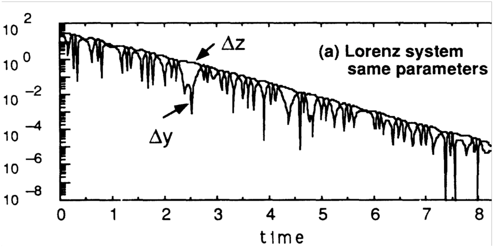


Figure: Synchronization between two copies of the Lorenz '63 model, coupled through the x-variable. The differences $\Delta y := |Y - y|$ and $\Delta z := |Z - z|$ decay exponentially in time. Note that even in this chaotic model, adding more data from the x-coordinate time series helps for convergence.

Projection

- There exist data assimilation algorithms that focus on using observations to control error in only the unstable direction, e.g. (Trevisan, D'Isodoro & Talagrand, 2010).
- Newton's method may be too expensive.

Summary of the algorithm

Solve $G(\boldsymbol{u}^{(k)}) = 0$ with starting guess $\boldsymbol{u}^{(0)} = \boldsymbol{y}$:

Iterate to convergence

- Compute the projections on the unstable direction $P_n^{(k)}$ along $\boldsymbol{u}^{(k)}$.
- Do a projected Newton step to compute $\hat{\boldsymbol{u}}^{(k)} = \boldsymbol{u}^{(k)} + P\boldsymbol{\delta}^{(k)}$.
- Synchronize in the stable directions by a forward integration: $u_{n+1}^{(k+1)} = P_{n+1}^{(k)} \hat{u}_{n+1}^{(k)} + (I - P_{n+1}^{(k)}) F_n(u_n^{(k+1)})$.

Lorenz '63

Model: $\dot{x}_1 = \sigma(x_2 - x_1)$, $\dot{x}_2 = x_1(\rho - x_3) - x_2$, $\dot{x}_3 = x_1 x_2 - \beta x_3$.
We choose: $\sigma = 10$, $\beta = \frac{8}{3}$ and $\rho = 28$ and add noise with coveriance $E = \sigma^2 I$, with $\sigma = 2$. We take $\ell^* = 2$.

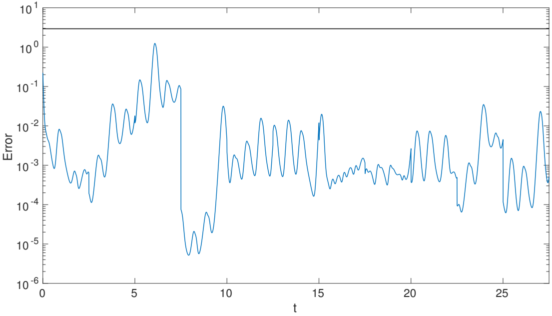
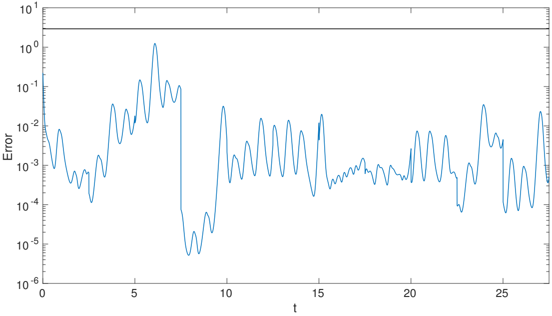


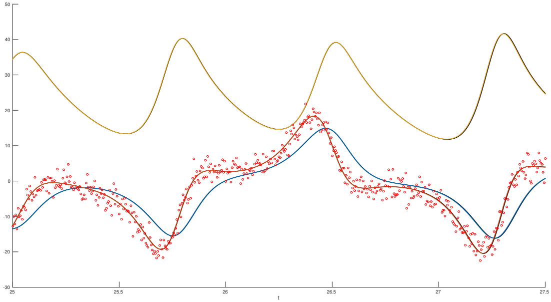
Figure: Applying Newton's method to the positive and neutral directions yields good results. Observational error level is in black, reduced error is in blue.

Lorenz '63



Property	Value
Average number of iterations	6.6
Average squared initial error	3.9
Average squared final error	0.09

Incomplete observations (only y)



Property	Value
Average number of iterations	6.2
Average squared initial error in y	4.0
Average squared final error in y	0.11
Average squared final error	0.08

Conclusions

- We have developped a new data assimilation algorithm based on numerical shadowing ideas.
- With this method, it is possible to treat error components in strongly stable directions different from errors in non-strongly stable directions.
- The splitting of the root-finding problem is independent from the methods chosen to solve the splitted equations.
- The use of synchronization decreases the cost of the algorithm.
- The projected Newton algorithm succeeds in finding piecewise model trajectories.

Outlook

- Incomplete observations
- Parameter estimation
- Zero Lyapunov exponents
- Forecasting
- Applying the method to other models
- More rigourous analysis of the method
- Uncertainty quantification

Thank you for listening

Questions?