SHADOWING APPROACH TO DATA ASSIMILATION

Svetlana Dubinkina

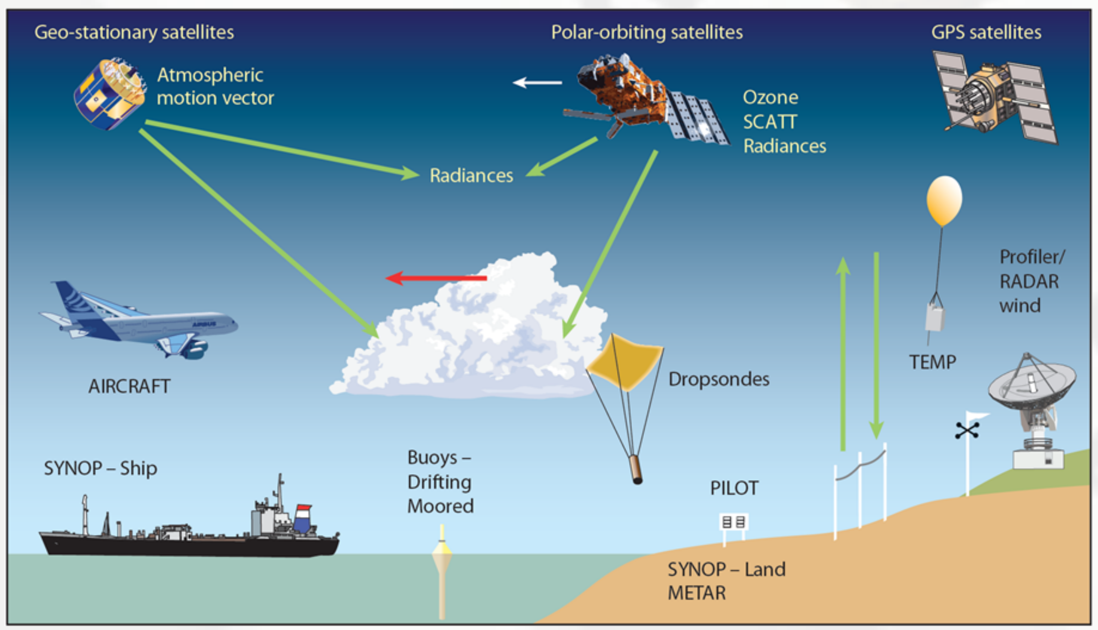
Joint work with Bart de Leeuw (CWI), Jason Frank (Utrecht U.), Andrew Steyer (Sandia National Laboratories), Xuemin Tu (U. of Kansas), Erik Van Vleck (U. of Kansas)

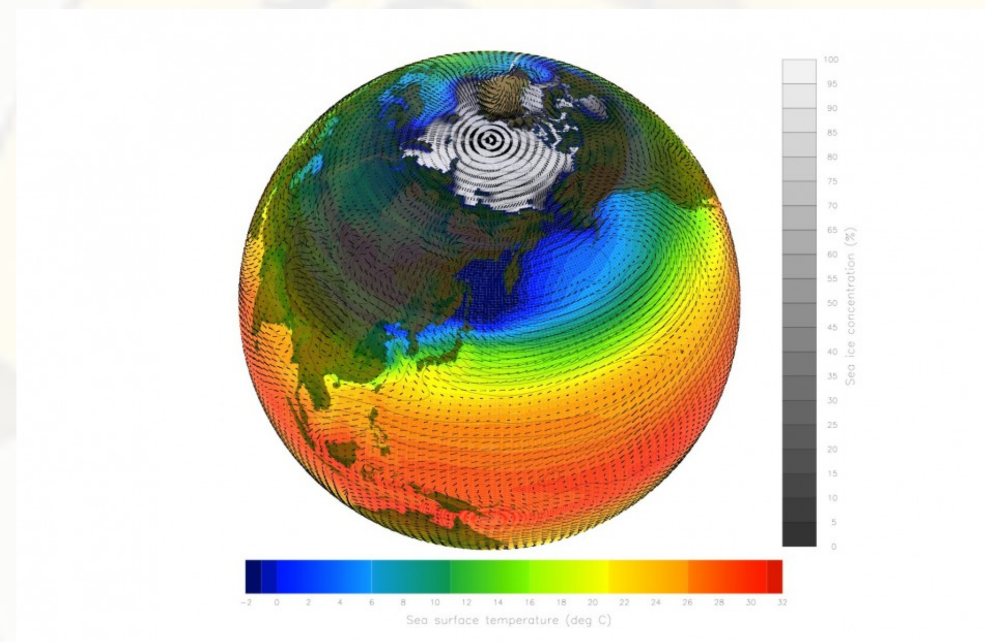
General Mathematics Colloquium

February 3, 2021

DATA ASSIMILATION

Data assimilation combines a computational model and measurements and bring synergy



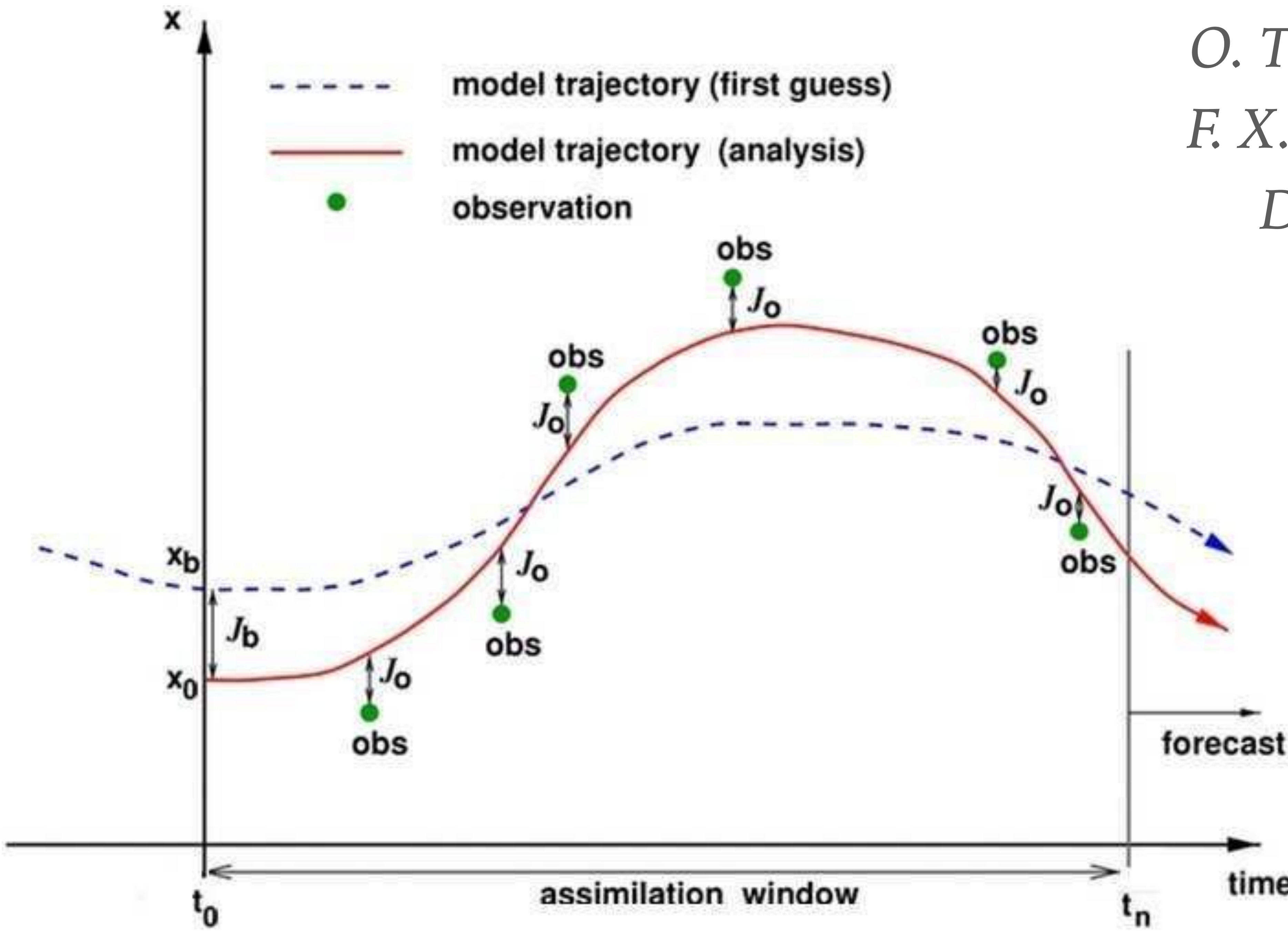


One of the goals of DA is an accurate estimation of initial conditions.

Our goal is accurate initial conditions that are moreover **more accurate** than the current DA methods can provide.

VARIATIONAL DATA ASSIMILATION

The majority of the weather forecasting centres uses a variational data assimilation approach (4DVar)



O. Talagrand, P. Courtier, F. X. Le Dimet, A. Lorenc, D. Dee, V. Penenko, …

LOCAL MINIMA OF THE COST FUNCTION

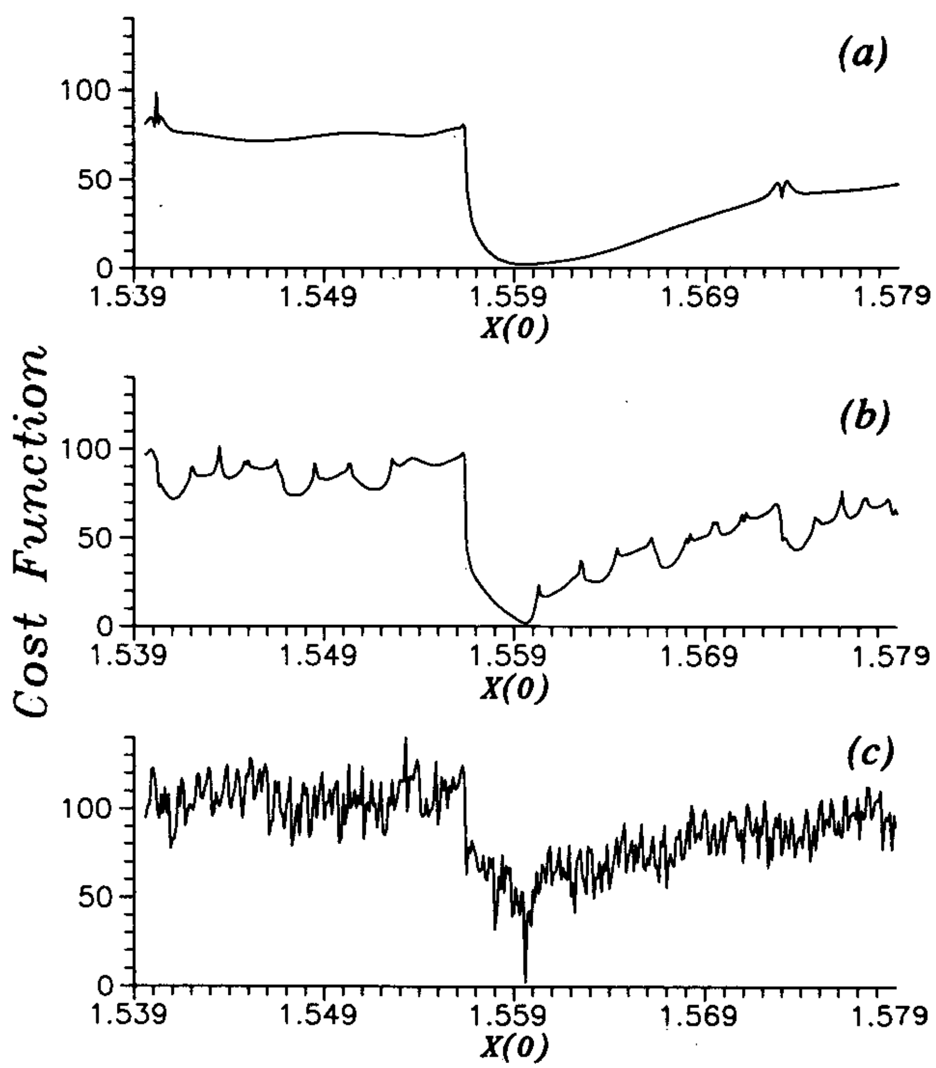


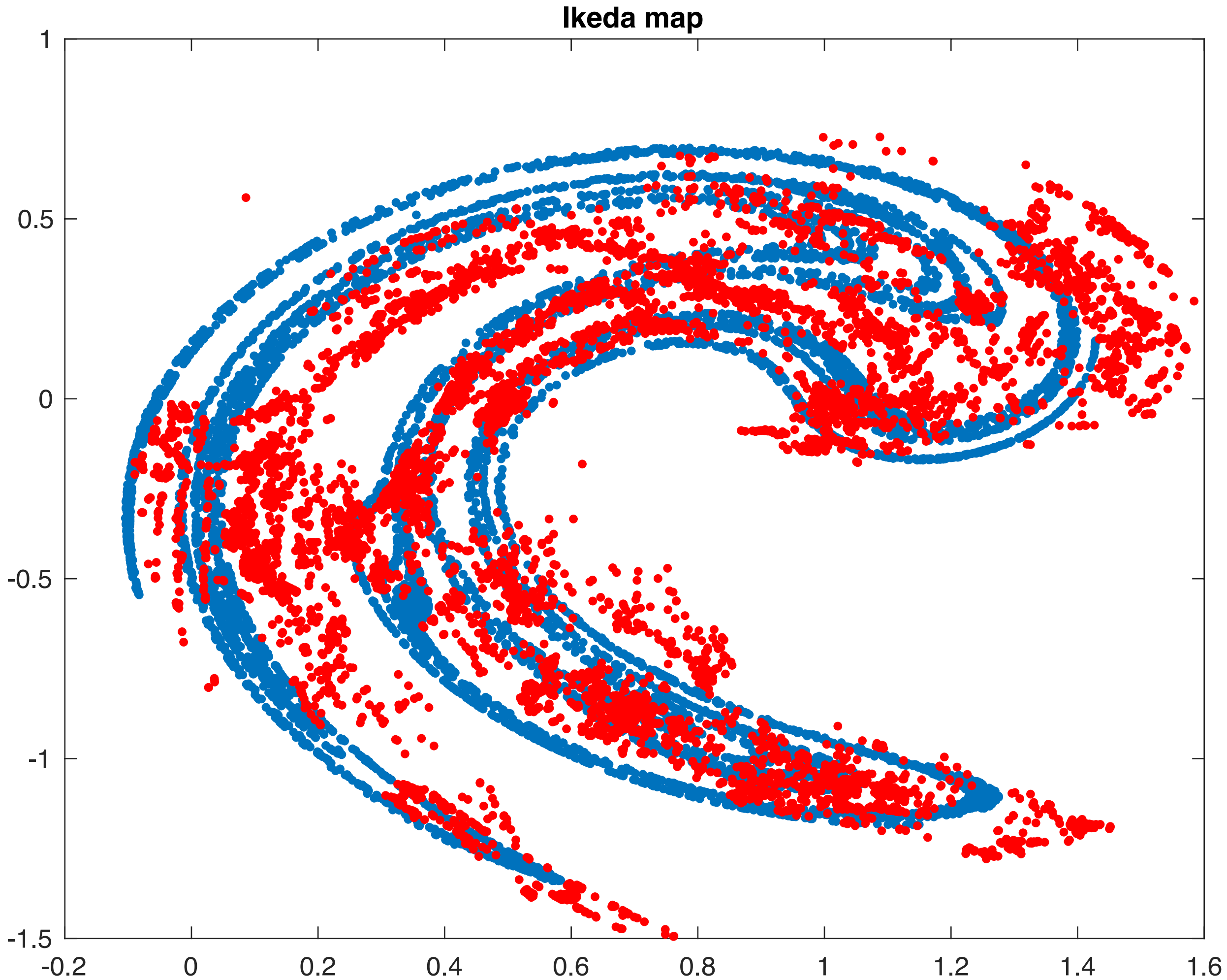
FIG. 6. Values of cost function as a function of initial X, with initial Y and Z held constant, in the neighborhood of the initial values used in calculating the reference solution. (a) Cost function, i.e., mean-square deviation of model solution with given initial data from "observed" values, where observations up to $t = 8$ are considered. (b) As in (a) but for observations up to $t = 10$. (c) As in (a) but for observations up to $t = 15$.

A drastic increase of the number of local minima of the corresponding cost function as the number of measurements increases.

Miller, Ghil, Gauthiez (1994)
J. Atmos. Sci.

SHADOWING APPROACH TO DATA ASSIMILATION

"The principle idea of shadowing-based data assimilation is to take observations of a trajectory (red dots) and to relax these onto a near-by trajectory (blue dots)." *K. Judd and L. Smith (2001)*.



K. Judd and L. Smith (2001);
K. Judd et al. (2008);
T. Stemler and K. Judd (2009)
H. Du and L. Smith (2014);
J. Brocker and U. Parlitz (2001)

SHADOWING LEMMA HAS NOTHING TO DO WITH DA, ORIGINALLY

Definition: $\{u_n\}_{n=0}^{N}$ is called an ε-pseudo-orbit of $\dot{z} = f(z)$ with an associated flow ϕ^{t_n}, if

$$||u_{n+1} - \phi^{t_n}(u_n)|| < \epsilon, \quad \mathit{for} \quad n = 0, \ldots, N-1$$

Shadowing lemma (A. Katok and B. Hasselblatt, 1995):
Let $\{u\}_{n=0}^{N}$ be an ε-pseudo-orbit of $\dot{z} = f(z)$. Then under some conditions there exists $\{u_n^{\text{exact}}\}_{n=0}^{N}$ on the orbit with $u_{n+1}^{\text{exact}} = \phi^{t_n}(u_n^{\text{exact}})$, such that

$$||u_n - u_n^{\text{exact}}|| < \delta, \quad \mathit{for} \quad n = 0, \ldots, N-1$$

The Shadowing lemma guarantees the existence but **not** necessarily the **uniqueness** of a solution u in a δ-neighbourhood of u^{exact}.

SHADOWING IN NUMERICAL DYNAMICAL SYSTEMS

- Shadowing is an important analysis technique for obtaining global error bounds on the numerical approximation to the solution of differential equations exhibiting chaos.

S. Chow, X. Lin, and K. Palmer (1989), S. M. Hammel, J. A. Yorke, and C. Grebogi (1990); E. Van Vleck (1994), …

- Shadowing refinement employs the pseudo-orbit as an initial guess for $u_{n+1} - \phi^{t_n}(u_n)$ and, as opposed to proving the existence of a nearby zero of $u_{n+1} - \phi^{t_n}(u_n)$, iteratively refines the pseudo-orbit to obtain an improved approximation of an exact solution. This is clearly akin to the data assimilation problem.

SHADOWING-BASED DATA ASSIMILATION

Search for zero of the cost operator

$$G(u) = \begin{pmatrix} G_0(u) \\ \vdots \\ G_{N-1}(u) \end{pmatrix}, \quad G_n(u) = u_{n+1} - \phi^{t_n}(u_n), \ \ for \ \ n = 0, \ldots, N-1,$$

using a contractive iteration starting from noisy observations

$$y_k = u_k^{\text{true}} + \xi_k, \quad for \quad 0 \le k \le N-1, \quad where \quad \xi_k \sim \mathcal{N}(0, \ R)$$

- Numerical shadowing refinement seeks a pseudo-orbit over the whole time interval at ones. This makes it applicable over **long time intervals**
- But it is **computationally more demeaning** than 4DVar
- Initialisation needs to be done using **full observations**

ASSIMILATION IN THE UNSTABLE SUBSPACE

- Recent efforts to improve speed and reliability of data assimilation specifically address the partitioning of the tangent space into stable, neutral, and unstable subspaces corresponding to Lyapunov vectors associated with negative, zero, and positive Lyapunov exponents, respectively: 4DVAR-AUS, projected ensemble Kalman filter

A. Trevisan, M. D'Isidoro, and O. Talagrand (2010);
L. Palatella, A. Carrassi, and A. Trevisan (2013);
C. Gonzalez-Tokman and B. R. Hunt (2013);
K. J. H. Law, D. Sanz-Alonso, A. Shukla and A. M. Stuart (2016)

- A dimension of the unstable subspace is smaller than a dimension of the model: 24 vs 14724 for a QG model (*R. Rotunno and J.-W. Bao 1996*)

PROJECTED SHADOWING-BASED DA METHOD

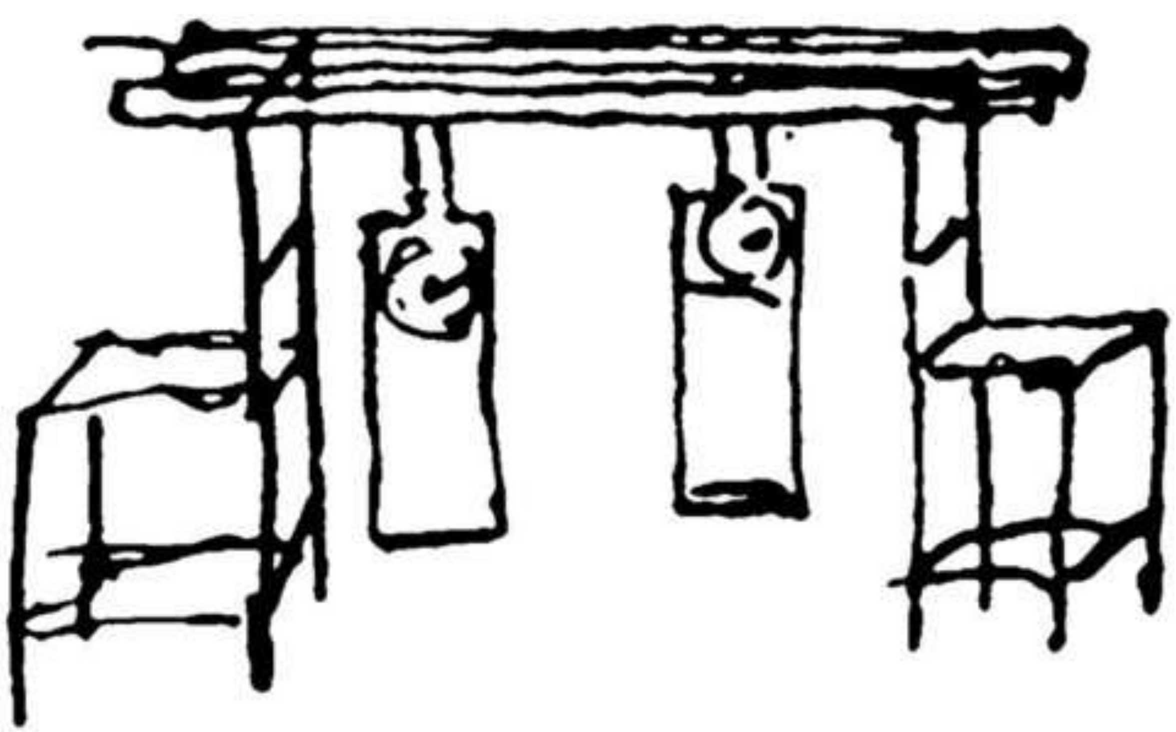
Motivated by these works, we propose a new method for shadowing-based data assimilation that utilises distinct treatments of the dynamics in the stable and nonstable (neutral and unstable) directions *(B. de Leeuw et al, 2018)*.

Novel projected shadowing-based DA method:

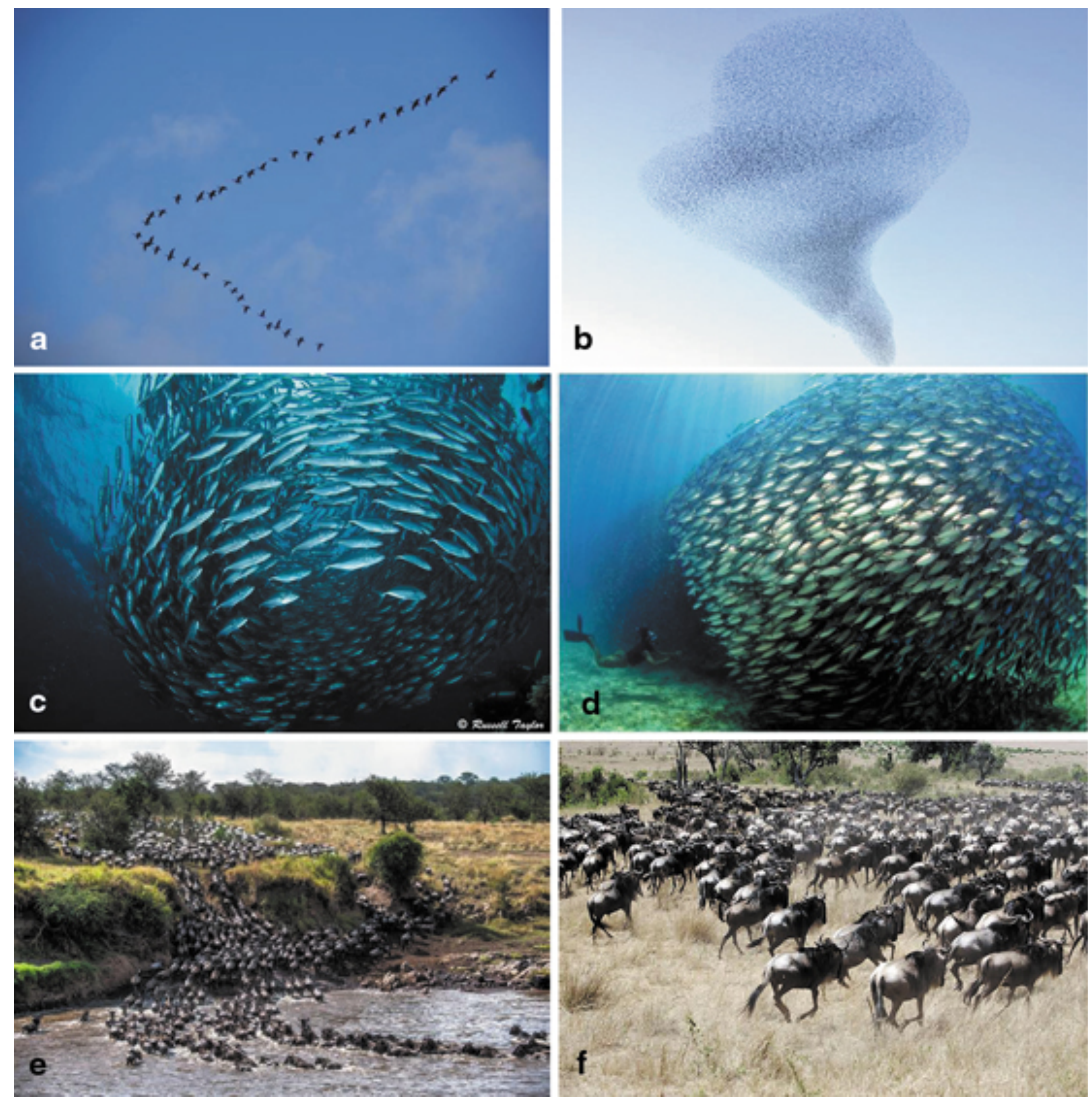
- We construct projection operators onto the stable and nonstable subspaces.
- In the nonstable subspace, we perform (expensive) shadowing-based DA that gives us a very accurate estimate.
- In the stable subspace, we decrease error by means of synchronisation to that accurate estimate.

SYNCHRONISATION

- Huygens synchronisation of two clocks



- Synchronisation in nature



SYNCHRONISATION IN DATA ASSIMILATION

Research on synchronisation of chaos indicates that

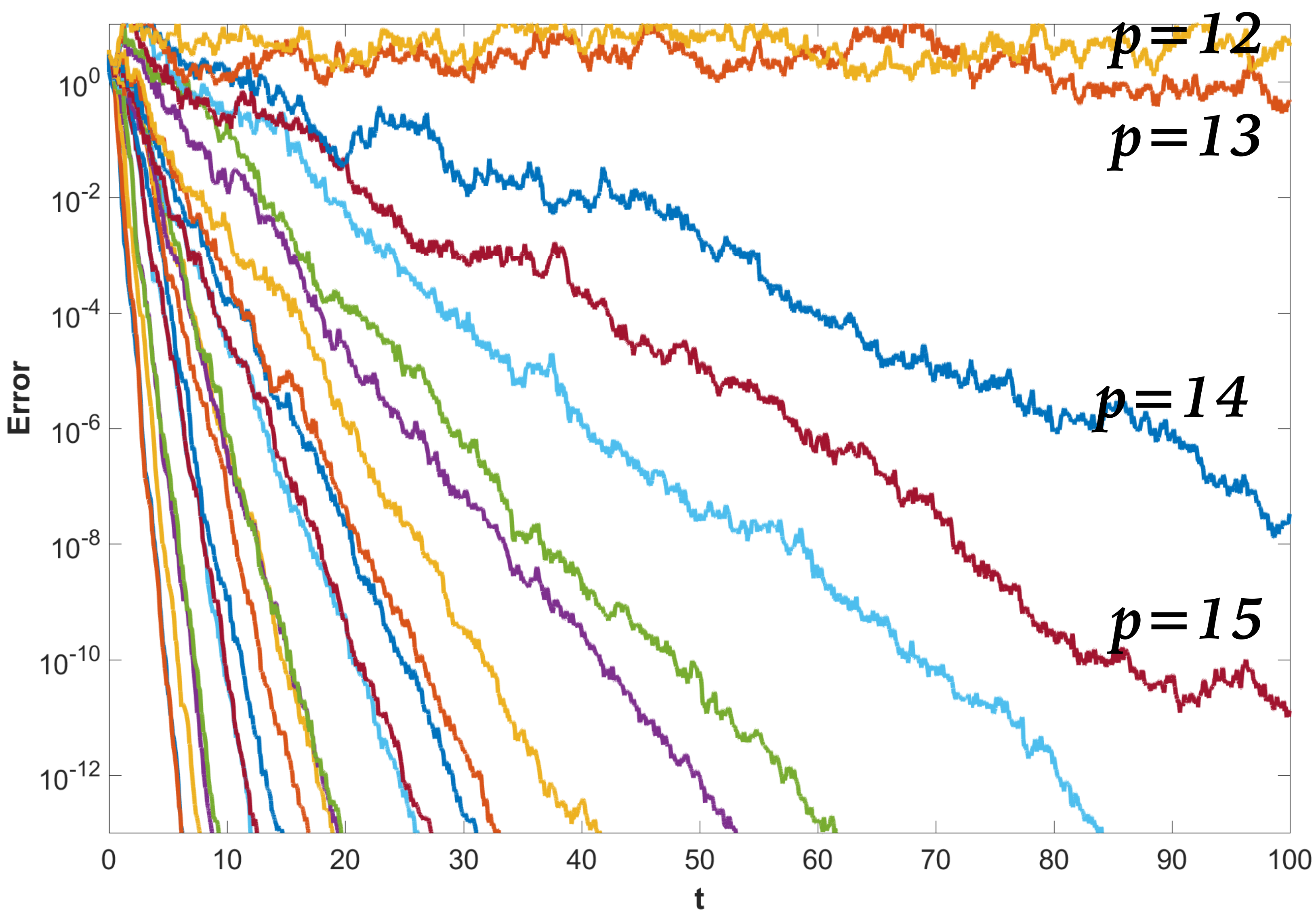
- when partial observations are sufficient to constrain the unstable subspace,
- an orbit of a chaotic dynamical system can be made to converge exponentially in time to a different, driving orbit.

(provided exponential dichotomy)

Pecora and Carroll (1990);
Pecora et al. (1997);
Boccaletti et al. (2002)

SYNCHRONISATION OF THE LORENZ 96 MODEL

- We consider the Lorenz 96 model (36 variables). It has 13 positive Lyapunov exponents.
- The true solution is partially observed (noise free): we have access to the true solution projected onto the non-strongly stable subspace of dimension p.
- Note that the dimension of the nonstable subspace is 14.



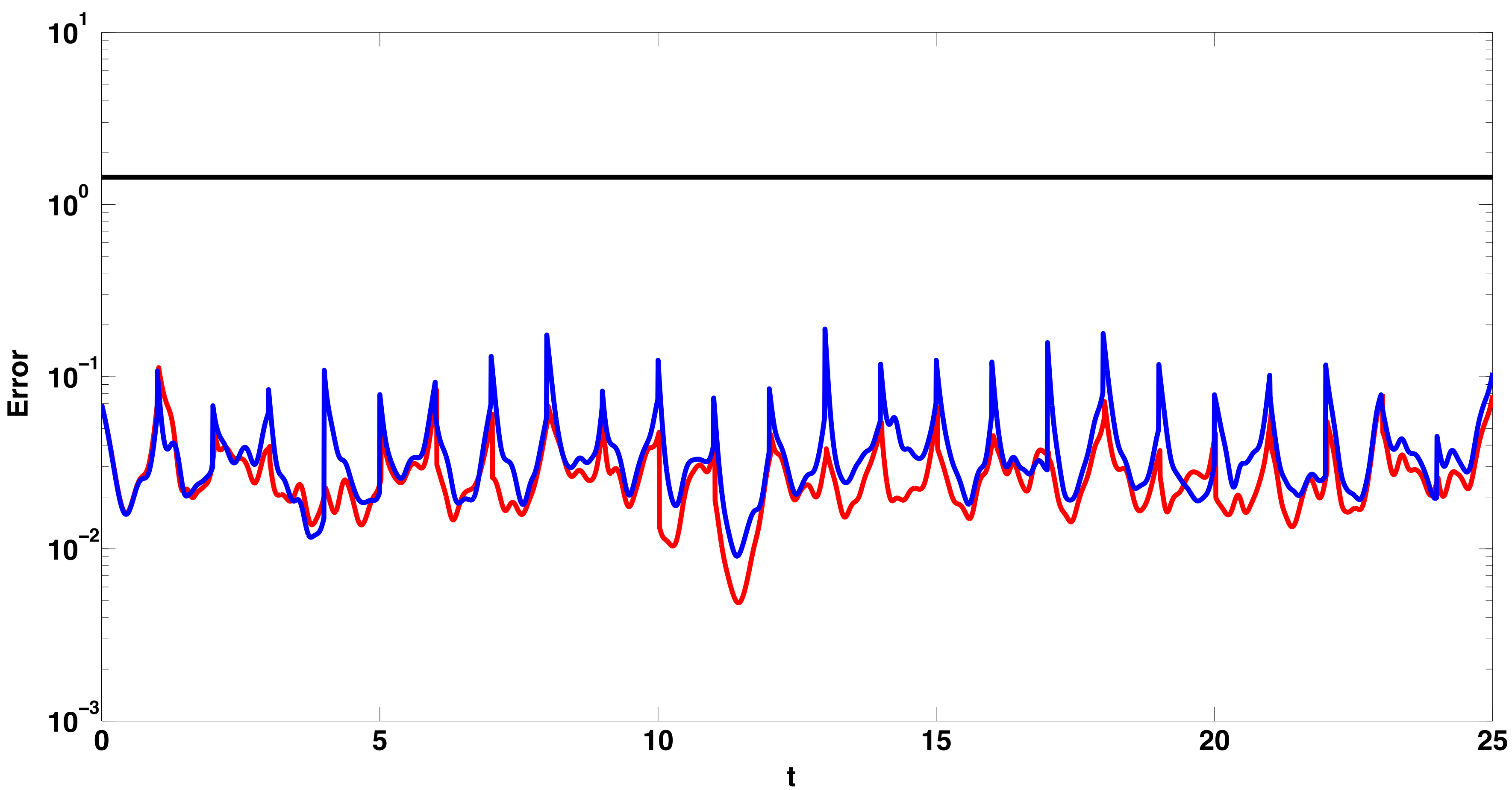
We plot

- *the difference between the true solution and the synchronisation approximation in the infinity norm*
- *as a function of time*
- *for different p*

THE PROJECTED SHADOWING-BASED DA METHOD: NUMERICAL EXPERIMENT

- Fully observed L96 model in space (36 variables).
- Observation noise is 1 unit. We observe every 3 hours, integration time step is 30 minutes.
- We project on 25 nonstable directions.

Error w.r.t. the truth in $\infty-$norm as a function of time



Black is for observations

Red is for shadowing

Blue is for 4DVar

GREAT LIMITATION OF FULL OBSERVATIONS

A shadowing-based approach to data assimilation has a very strict assumption of full observations in space.

Until now, this assumption was weakened by proposing to have a preprocessing before shadowing-based data assimilation can be performed.

A preprocessing consists of applying another data assimilation method (4DVar) with partial observations to have a proxy of the whole trajectory.

We propose to lift up the requirement of full observations and of a preprocessing.

B. de Leeuw and S.D. (2020)

SHADOWING-BASED DA FOR PARTIAL OBSERVATIONS

The initial guess now consists of the observations and a background trajectory—a solution of the model with an arbitrary initial condition.

Initiating at this initial guess, we use a regularized Gauss-Newton method to find a pseudo-orbit of G

$$G_n := u_{n+1} - \phi^{t_n}(u_n)$$

- What can we say about convergence of the shadowing-based DA method with partial observations?
- What can we say about closeness of an estimate to the true orbit, from which observations are generated?

LOCAL CONVERGENCE AND TRUST REGION

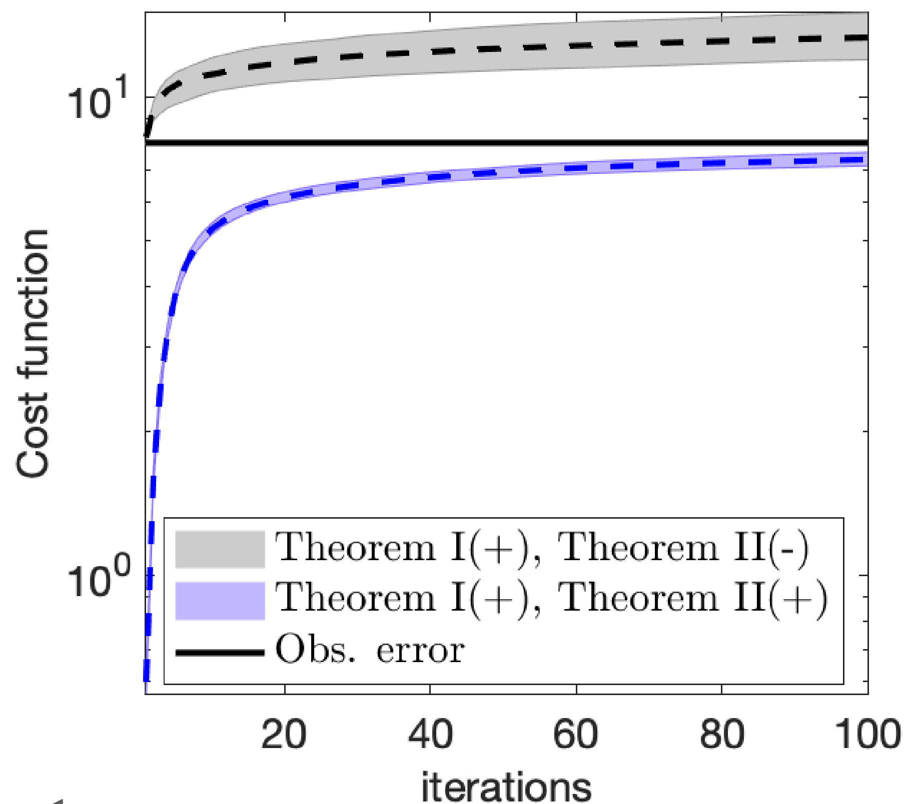
- Theorem I: Under some conditions on the initial guess and a regularization parameter, the shadowing-based DA method converges locally to the solution manifold $G(u) = 0$.

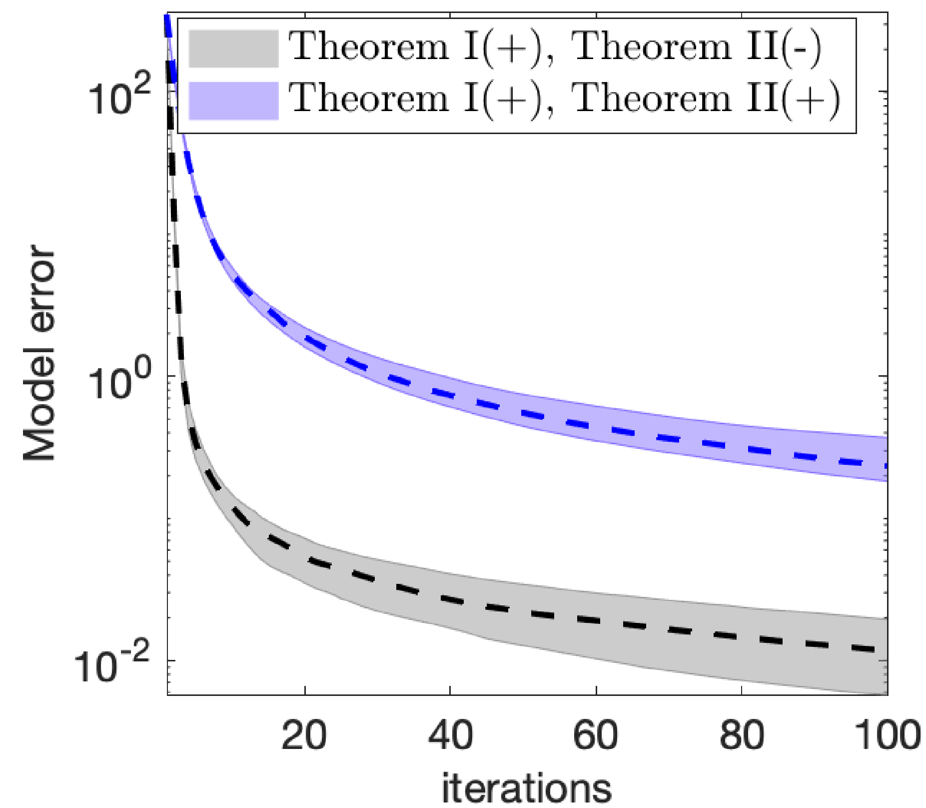
- Theorem II: Under some conditions, a shadowing-based estimate projected on the observation space remains in a ball centred at the observations and radius of the observation error.

  In practice: in order to fulfil the conditions of Theorem II, we need to choose a specific preconditioning for the Gauss-Newton method.

THE SHADOWING-BASED DA METHOD WITH PARTIAL OBSERVATIONS: NUMERICAL EXPERIMENT

We observe every 2nd variable of the Lorenz 96 model every 6 hours over 25 days. Variance of the observation error is 8.





$$\frac{1}{\#K}\sum_{k\in K}(Hu_k - y_k)^T(Hu_k - y_k)$$

H is the observation operator that projects an estimate onto the observation phase space

$$\frac{1}{N}\sum_{n=0}^{N-1} G_n^T G_n$$

$$G_n := u_{n+1} - \phi^{t_n}(u_n)$$

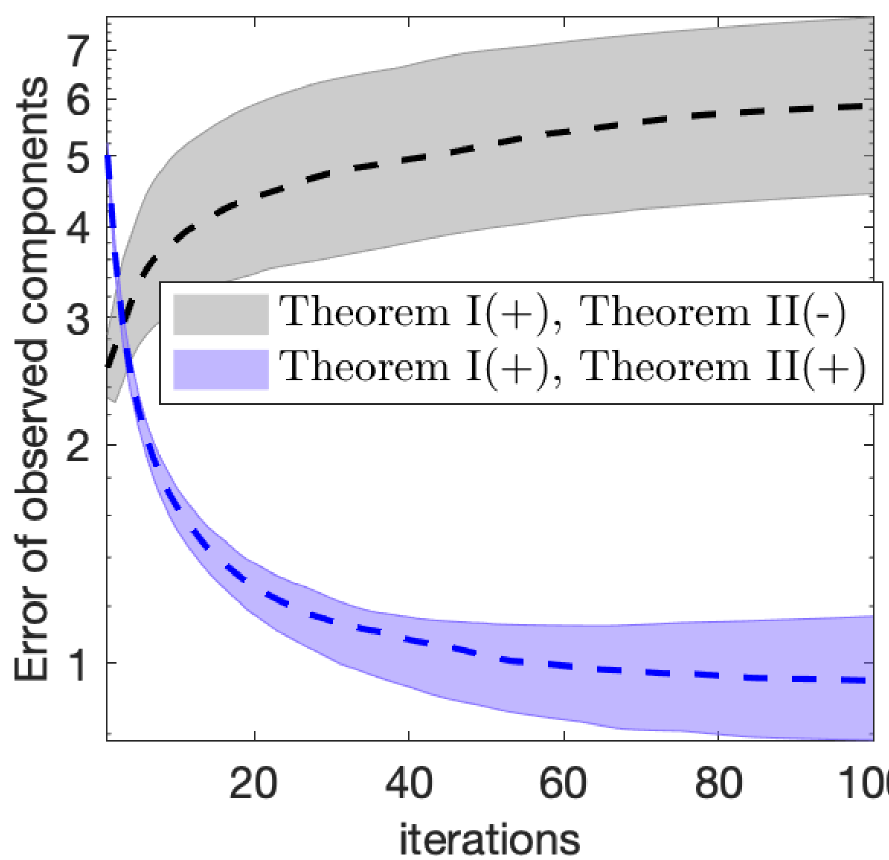
ERROR WITH RESPECT TO THE TRUE SOLUTION

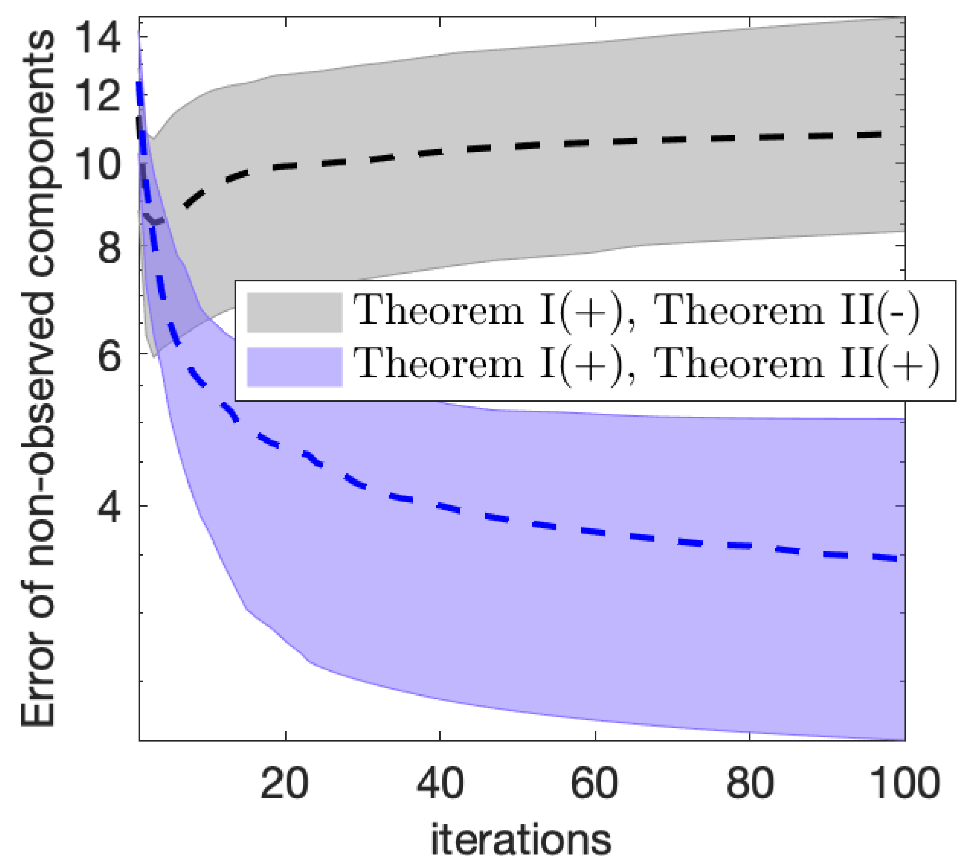
$$\frac{1}{N}\sum_{n=0}^{N-1}(Hu_n - Hu_n^{\text{true}})^T(Hu_n - Hu_n^{\text{true}})$$

H is the observation operator that projects an estimate onto the observation phase space

$$\frac{1}{N}\sum_{n=0}^{N-1}(H^{\perp}u_n - H^{\perp}u_n^{\text{true}})^T(H^{\perp}u_n - H^{\perp}u_n^{\text{true}})$$

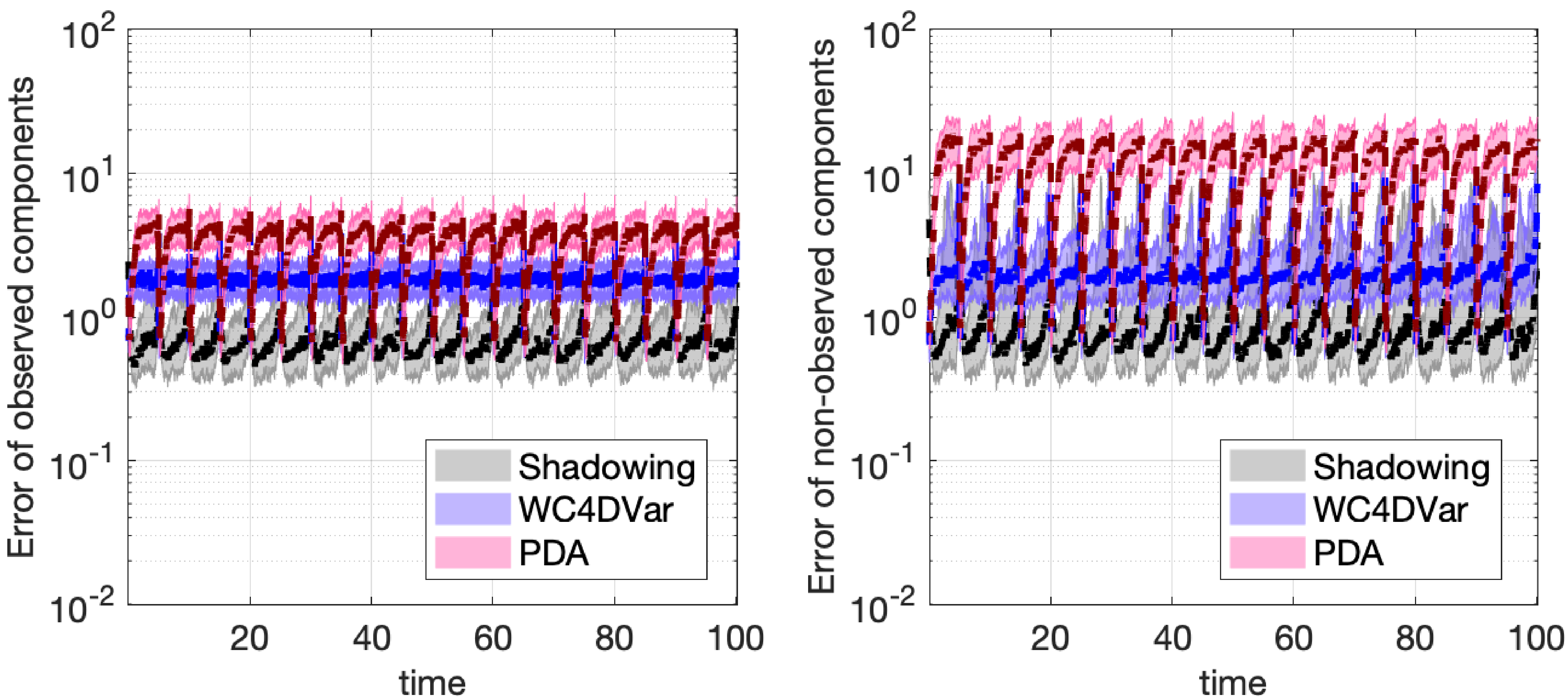
$H^{\perp}$ is an operator that projects an estimate onto the "non-observed" phase space





COMPARISON TO OTHER DA METHODS

We compare the shadowing-based DA method to a weak constraint 4DVar and to a Pseudo-orbit DA method. WC4DVar is a weak-constraint variational method, while PDA is a shadowing-based DA method as well. We plot error with respect to the true solution over time.



CONCLUSIONS

- We are developing a shadowing-based DA approach.
- We showed that the computation costs can be decreased by using a tangent splitting to the stable and nonstable subspaces.
- We extended the shadowing-based DA method to partial observations.
- We proved that the method converges to the solution manifold.
- Furthermore, we proved that the solution projected onto the observation space is within a ball centred at the observation with radius of the observation error.
- We showed numerically that the shadowing-based DA method provides a very accurate estimation of the true solution, which is more accurate than an estimation of a variational data assimilation method.

FUTURE WORK

- Tangent splitting in the shadowing-based approach with partial observations?
- Bound on the error with respect to the true solution? - Nazanin Abedini has already first results!
- Which variables to observe to control the system?
- Ensemble approximation?
- Model error?

References to our work:

- B. de Leeuw, S. Dubinkina, J. Frank, A. Steyer, X. Tu and E. Van Vleck, "Projected Shadowing-based Data Assimilation", SIAM J. Appl. Dyn. Syst., 17(4) (2018)
- B. de Leeuw and S. Dubinkina, "Shadowing-based data assimilation method for partially observed models" arXiv:1810.07064

Thank you

for

your attention!

FAIR WARNING, OR NOT A WARNING AT ALL?

Which systems have the shadowing property?

- Hyperbolic systems (with non-zero Lyapunov exponents) can be shadowed for infinite times (*D. Anosov 1967*)
- Non-uniformly hyperbolic systems (with zero Lyapunov exponents) can be shadowed for finite but nontrivial times (*S. M. Hammel, J. A. Yorke, and C. Grebogi 1990; E. Van Vleck 1995*)
- If the dimension of the unstable subspace, which is related to the positive Lyapunov exponents, is not constant, then shadowing of numerical trajectories for relatively long time is impossible (*S.P. Dawson 1996*)

Shadowing time

Infinite / *Finite but long* / *Relatively not long*

EXISTING SHADOWING-BASED DA METHODS

$$u^{(j+1)} = u^{(j)} + \Delta^{(j)}, \qquad D := G'(u^{(j)})$$

1) $\Delta^{(j)} = -\gamma D^T G(u^{(j)})$ *Judd and Smith (2001)*

2) $\Delta^{(j)} = -D^T \Lambda^{-1} G(u^{(j)})$ *Brocker and Parlitz (2001)*

3) $\Delta^{(j)} = -D^T (DD^T)^{-1} G(u^{(j)})$ *de Leeuw et al. (2018)*

4) $\Delta^{(j)} = -\Sigma D^T (D\Sigma D^T + \alpha Q)^{-1} G(u^{(j)})$ *de Leeuw and S.D. (2020)*

Under Theorem I all these methods (1)—(4) converge to the solution manifold. The aim of the method (4) is to provide a good estimation of the true solution when using partial observations (Theorem II).