Uncertainty Prediction and Reduction in Ocean Fields: Adaptive Dynamically Orthogonal Equations

Pierre F.J. Lermusiaux

T. Sapsis, M. Ueckermann, T. Sondergaard, P. Haley and K. Yilgit

Mechanical Engineering, MIT

Multidisciplinary Simulation, Estimation and Assimilation Systems (MSEAS)
http://mseas.mit.edu/

- **Introduction**
- **Two Grand Challenges in Ocean/Earth-System Sciences & Engineering**
  - **Prognostic Equations for Stochastic Fields of Large-Dimension**
  - **Intelligent Adaptive Sampling: the Science of Autonomy**
  - **Non-Gaussian Data Assimilation with DO eqns and EM-algorithm**
- **Conclusions**

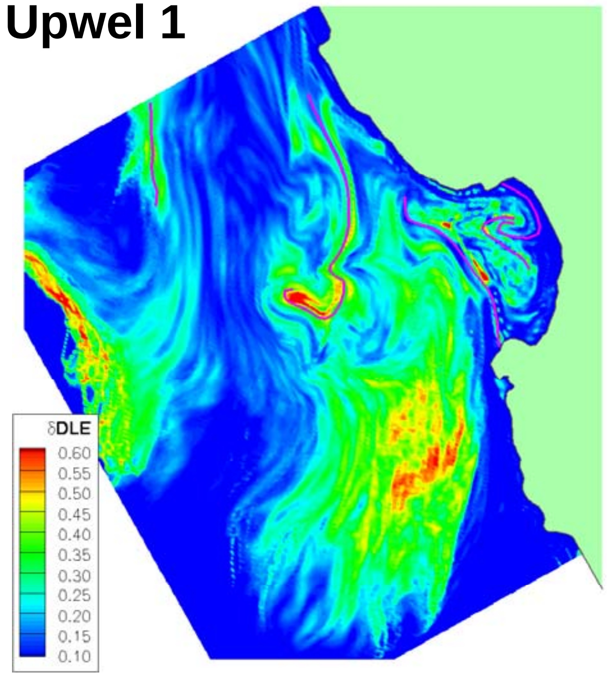
Thanks to MIT, ONR and NSF

Flow Skeletons and Uncertainties:

Mean LCS overlaid on DLE error std estimate for 3 dynamical events

- Two upwellings and one relaxation (**about 1 week apart each**)
- Uncertainty estimates allow to identify most robust LCS (more intense DLE ridges are usually relatively more certain)
- Different oceanic regimes have different LCS uncertainty fields and properties

Upwel 1

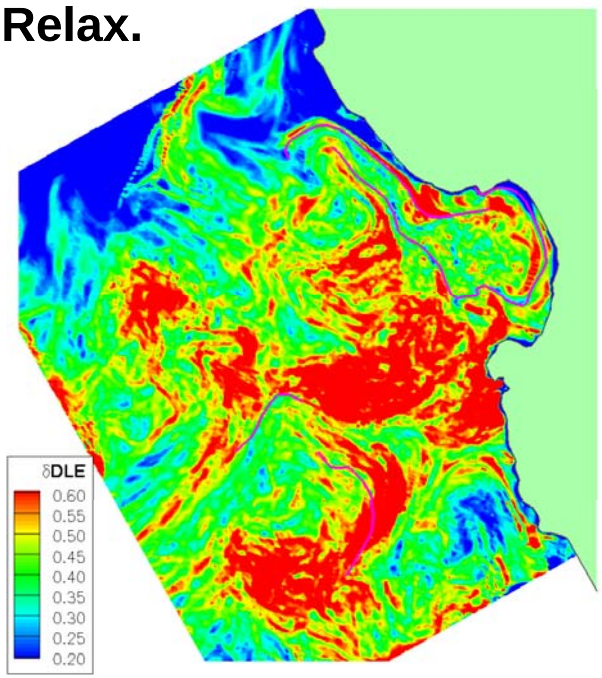


[Lermusiaux and Lekien, 2005. and In Prep, 2010

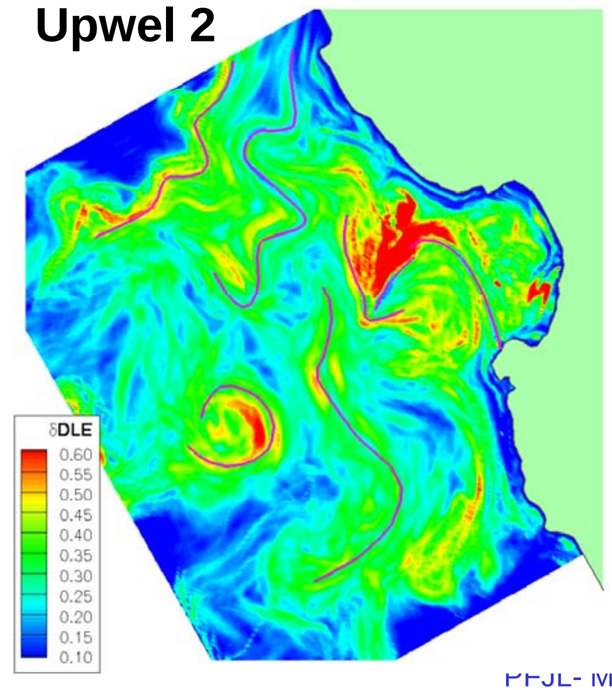
Lermusiaux, JCP-2006

Lermusiaux, Ocean.-2006]

Relax.



Upwel 2



A Grand challenge in Large Nonlinear Systems

Quantitatively estimate the accuracy of predictions

Computational challenges for the deterministic (ocean) problem

- Large dimensionality of the problem, un-stationary statistics
- Wide range of temporal and spatial scales (turbulent to climate)
- Multiple instabilities internal to the system
- Very limited observations

Need for stochastic modeling …

- Approximations in deterministic models including parametric uncertainties
- Initial and Boundary conditions uncertainties
- Measurement models

Need for data assimilation …

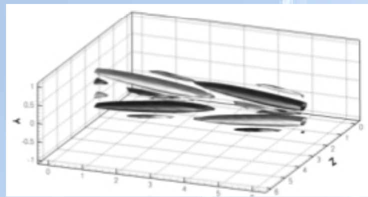
- Evolve the nonlinear, i.e. non-Gaussian, correlation structures
- Nonlinear Bayesian Estimation

Overview of Uncertainty Predictions Schemes

$$\boxed{\boldsymbol{u}(\boldsymbol{x},t;\omega)} = \bar{\boldsymbol{u}}(\boldsymbol{x},t) + \sum_{i=1}^{s} \boxed{Y_i(t;\omega)}\boxed{\boldsymbol{u}_i(\boldsymbol{x},t)}$$

Uncertainty propagation via POD method

[C. Rowley, Oberwolfach, 2008]



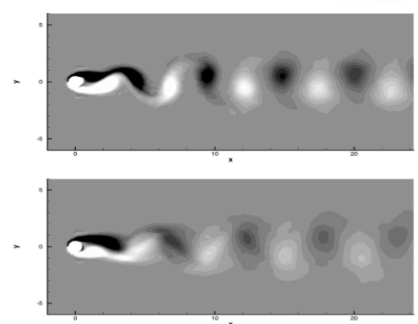
According to Lumley (Stochastic tools in Turbulence, 1971) it was introduced independently by numerous people at different times, including Kosambi (1943), Loeve (1945), Karhunen (1946), Pougachev (1953), Obukhov (1954).

Uncertainty propagation via generalized Polynomial-Chaos Method

Xiu & Karniadakis, J. Comp. Physics, 2002

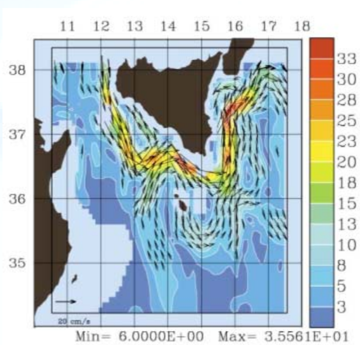
Knio & Le Maitre, Fluid Dyn. Research, 2006

Meecham & Siegel, Phys. Fluids, 1964



[Xiu & Karniadakis, J. Comp. Physics, 2002]

[Lermusiaux & Robinson, Deep Sea Research, 2001]



Uncertainty propagation via Monte Carlo method restricted to an "evolving uncertainty subspace"

(Error Subspace Statistical Estimation - ESSE)

Lermusiaux & Robinson, MWR-1999, Deep Sea Research-2001

Lermusiaux, J. Comp. Phys., 2006

B. Ganapathysubramanian & N. Zabaras, J. Comp. Phys., (under review)

Problem Setup

Statement of the problem: A Stochastic PDE

$$\frac{\partial \boldsymbol{u}(\boldsymbol{x},t;\omega)}{\partial t} = \mathcal{L}\left[\boldsymbol{u}(\boldsymbol{x},t;\omega);\omega\right] \qquad \boldsymbol{x} \in \boldsymbol{D}$$

$$\boldsymbol{u}(\boldsymbol{x},t_0;\omega) = \boldsymbol{u}_0(\boldsymbol{x};\omega) \qquad \boldsymbol{x} \in \boldsymbol{D} \qquad \mathcal{B}\left[\boldsymbol{u}|_{\partial \boldsymbol{D}}\right] = h\left[\partial \boldsymbol{D};\omega\right]$$

$\mathcal{L}[\cdot;\omega]$ Nonlinear differential operator (possibly with stochastic coefficients)

$\boldsymbol{u}_0(\boldsymbol{x};\omega)$ Stochastic initial conditions (given full probabilistic information)

$h[\partial \boldsymbol{D};\omega]$ Stochastic boundary conditions (given full probabilistic information)

Goal: Evolve the full probabilistic information describing $\boldsymbol{u}(\boldsymbol{x},t;\omega)$

An important representation property for the solution: Compactness

$$\boldsymbol{u}(\boldsymbol{x},t;\omega) = \bar{\boldsymbol{u}}(\boldsymbol{x},t) + \sum_{i=1}^{s} Y_i(t;\omega)\,\boldsymbol{u}_i(\boldsymbol{x},t)$$

Advantage: Finite Dimension Evolving Subspace

Disadvantage: Redundancy of representation

Evolving the full representation

Major Challenge : Redundancy

$$\boldsymbol{u}(\boldsymbol{x},t;\omega)=\bar{\boldsymbol{u}}(\boldsymbol{x},t)+\sum_{i=1}^{s}Y_i(t;\omega)\boldsymbol{u}_i(\boldsymbol{x},t)$$

First Step (easy): Separate deterministic from stochastic/error subspace

Commonly used approach: Assume that $\overline{Y_i(t;\omega)}=0$

Second step (tricky): Evolving the finite dimensional subspace $\mathcal{V}_S$

A separation of roles: What can $\dfrac{dY_i(t;\omega)}{dt}$ tell us ?

Only *how the stochasticity evolves inside $\mathcal{V}_S$*

source of redundancy

A separation of roles: What can $\dfrac{\partial \boldsymbol{u}_i(\boldsymbol{x},t)}{\partial t}$ tell us ?

How the stochasticity evolves ***both*** *inside and normal to $\mathcal{V}_S$*

Natural constraint to overcome redundancy

Restrict "evolution of $\mathcal{V}_S$" to be "normal to $\mathcal{V}_S$" i.e.

$$\int \frac{\partial \boldsymbol{u}_i(\boldsymbol{x},t)}{\partial t}\boldsymbol{u}_j(\boldsymbol{x},t)\,d\boldsymbol{x}=0 \qquad \text{for all} \quad i=1,\dots,s \quad \text{and} \quad j=1,\dots,s$$

Dynamically Orthogonal Evolution Equations

Theorem 1: For a stochastic field described by the evolution equation

$$\frac{\partial \boldsymbol{u}(\boldsymbol{x},t;\omega)}{\partial t} = \mathcal{L}\left[\boldsymbol{u}(\boldsymbol{x},t;\omega);\omega\right] \quad , \quad \boldsymbol{x}\in \boldsymbol{D}$$

$$\boldsymbol{u}(\boldsymbol{x},t_0;\omega) = \boldsymbol{u}_0(\boldsymbol{x};\omega) \quad , \quad \boldsymbol{x}\in \boldsymbol{D} \qquad \mathcal{B}\left[\boldsymbol{u}(\xi,t;\omega)\right] = h(\xi,t;\omega) \quad , \quad \xi \in \partial \boldsymbol{D}$$

assuming a response of the form $\boldsymbol{u}(\boldsymbol{x},t;\omega) = \bar{\boldsymbol{u}}(\boldsymbol{x},t) + \sum_{i=1}^{s} Y_i(t;\omega)\boldsymbol{u}_i(\boldsymbol{x},t)$

we obtain the following evolution equations

SDE describing evolution of stochasticity inside V_s

$$\frac{dY_j(t;\omega)}{dt} = \int_{\boldsymbol{D}} \left\{\mathcal{L}\left[\boldsymbol{u}(\boldsymbol{y},t;\omega)\right] - E^{\omega}\left[\mathcal{L}\left[\boldsymbol{u}(\boldsymbol{y},t;\omega)\right]\right]\right\} \boldsymbol{u}_j(\boldsymbol{y},t)\,d\boldsymbol{y}$$

Family of PDEs describing evolution of stochastic subspace V_s

$$\frac{\partial \boldsymbol{u}_j(\boldsymbol{x},t)}{\partial t} = E^{\omega}\left[Y_i(t;\omega)\mathcal{L}\left[\boldsymbol{u}(\boldsymbol{x},t;\omega)\right]\right]\boldsymbol{C}^{-1}_{Y_iY_j} - E^{\omega}\left[\int_{\boldsymbol{D}} \boldsymbol{u}_k(\boldsymbol{y},t) Y_i(t;\omega)\mathcal{L}\left[\boldsymbol{u}(\boldsymbol{y},t;\omega)\right] d\boldsymbol{y}\right]\boldsymbol{C}^{-1}_{Y_iY_j}\boldsymbol{u}_k(\boldsymbol{x},t)$$

$$\mathcal{B}\left[\boldsymbol{u}_j(\xi,t;\omega)\right] = E^{\omega}\left[Y_i(t;\omega)h(\xi,t_0;\omega)\right]\boldsymbol{C}^{-1}_{Y_iY_j}, \;\; \xi \in \partial\boldsymbol{D}$$

PDE describing evolution of mean field

$$\frac{\partial \bar{\boldsymbol{u}}(\boldsymbol{x},t)}{\partial t} = E^{\omega}\left[\mathcal{L}\left[\boldsymbol{u}(\boldsymbol{x},t;\omega)\right]\right], \; \boldsymbol{x}\in\boldsymbol{D} \qquad \mathcal{B}\left[\bar{\boldsymbol{u}}(\xi,t;\omega)\right] = E^{\omega}\left[h(\xi,t;\omega)\right], \; \xi \in \partial\boldsymbol{D}$$

Sapsis and Lermusiaux, Physica D (2009)

POD & PC methods from DO equations

Family of PDEs describing evolution of stochastic subspace V_s

$$\frac{\partial \boldsymbol{u}_i(x,t)}{\partial t} = E^{\omega}\left[Y_i(t;\omega)\mathcal{L}\left[\boldsymbol{u}(x,t;\omega)\right]\right]C^{-1}_{Y_iY_j} - E^{\omega}\left[\int_D \boldsymbol{u}_k(y,t)Y_i(t;\omega)\mathcal{L}\left[\boldsymbol{u}(y,t;\omega)\right]dy\right]C^{-1}_{Y_iY_j}\boldsymbol{u}_k(x,t)$$

$$\mathcal{B}\left[\boldsymbol{u}_i(\xi,t;\omega)\right] = E^{\omega}\left[Y_i(t;\omega)h(\xi,t_0;\omega)\right]C^{-1}_{Y_iY_j},\ \xi\in\partial \boldsymbol{D}$$

PDE describing evolution of mean field

$$\frac{\partial \overline{\boldsymbol{u}}(\boldsymbol{x},t)}{\partial t} = E^{\omega}\left[\mathcal{L}\left[\boldsymbol{u}(\boldsymbol{x},t;\omega)\right]\right],\ \boldsymbol{x}\in\boldsymbol{D} \qquad \mathcal{B}\left[\overline{\boldsymbol{u}}(\xi,t;\omega)\right] = E^{\omega}\left[h(\xi,t;\omega)\right],\ \xi\in\partial\boldsymbol{D}$$

Choosing a priori the stochastic subspace V_s using POD methodology we recover POD equations.

Choosing a priori the statistical characteristics of the stochastic coefficients $Y_j(t;\omega)$ we recover the PC equations.

Application I : Navier-Stokes in a cavity

2D viscous flow with stochastic initial conditions and no stochastic excitation

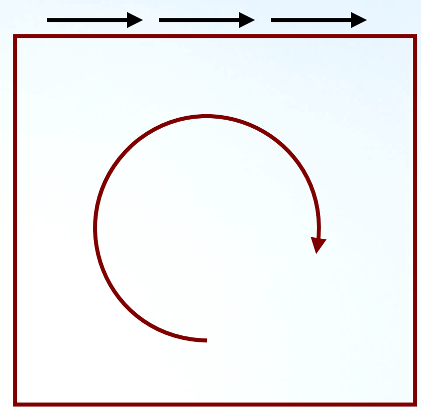
$$\frac{\partial u}{\partial t}+\frac{\partial P}{\partial x}=\frac{1}{\mathrm{Re}}\Delta u-\frac{\partial\left(u^2\right)}{\partial x}-\frac{\partial\left(uv\right)}{\partial y}$$

$$\frac{\partial v}{\partial t}+\frac{\partial P}{\partial y}=\frac{1}{\mathrm{Re}}\Delta v-\frac{\partial\left(uv\right)}{\partial x}-\frac{\partial\left(v^2\right)}{\partial y}$$

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0$$

$u=U, v=0$

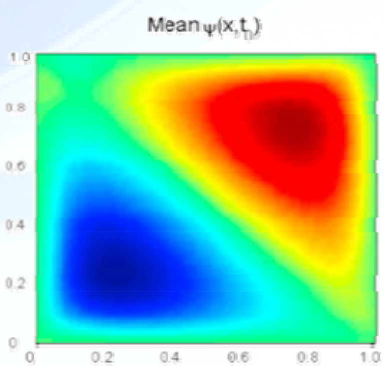
$u=0$ $v=0$



$u=0$ $v=0$

$u=0, v=0$

Initial mean flow



Streamfunction

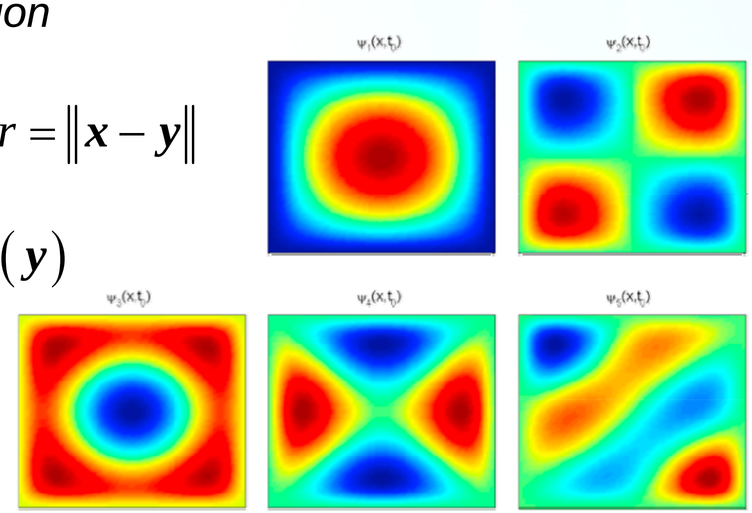
Initial Covariance function

$$C\left(r\right)=\left(1+br+\frac{b^2r^2}{3}\right)e^{-br}\quad r=\left\|\boldsymbol{x}-\boldsymbol{y}\right\|$$

$$\int \boldsymbol{C}\left(\left\|\boldsymbol{x}-\boldsymbol{y}\right\|\right)\hat{\boldsymbol{u}}_i\left(\boldsymbol{x}\right)d\boldsymbol{x}=\lambda_i^2\hat{\boldsymbol{u}}_i\left(\boldsymbol{y}\right)$$

$$\boldsymbol{u}_{0,i}\left(\boldsymbol{x}\right)=\hat{\boldsymbol{u}}_i\left(\boldsymbol{x}\right)$$

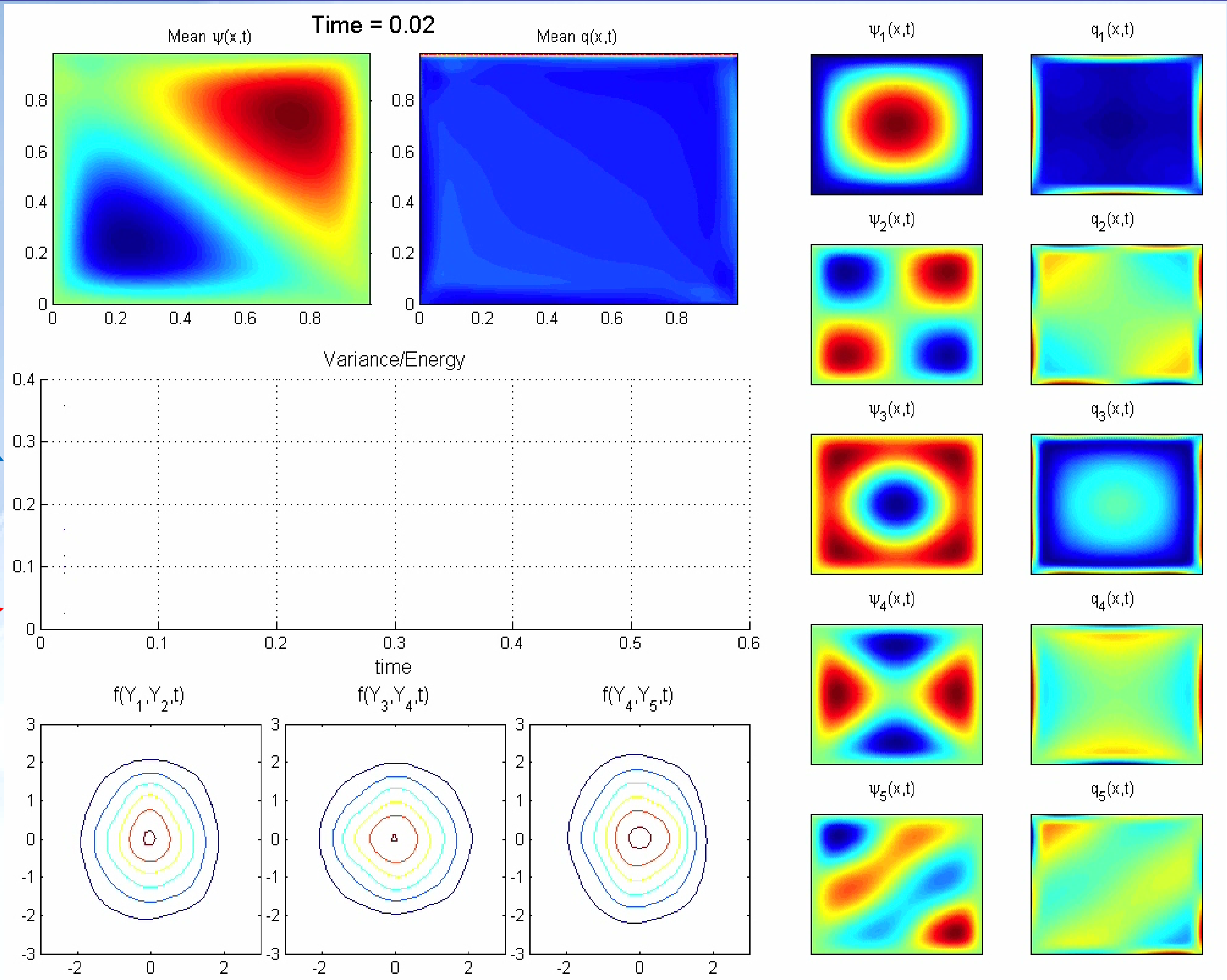
$$Y_i\left(t_0;\omega\right)\sim\mathcal{N}\left(0,\lambda_i\right)$$



PDE Numerics: C-grid, upwind [M. Griebel et al., 1998] SDE Numerics: here, s-dimensional Monte-Carlo

Application I : Navier-Stokes in a cavity

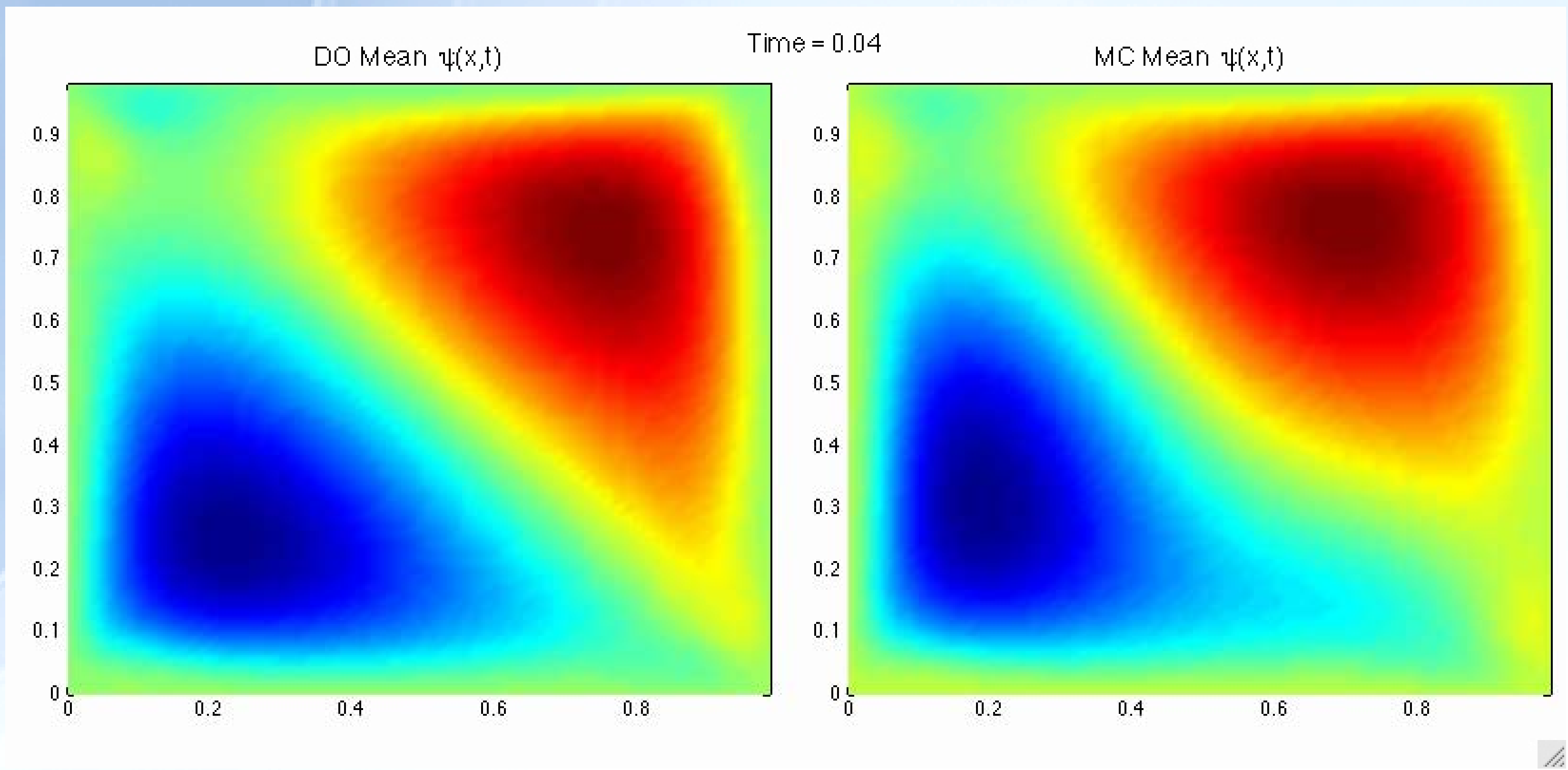
Re = 1000



Variances of each mode

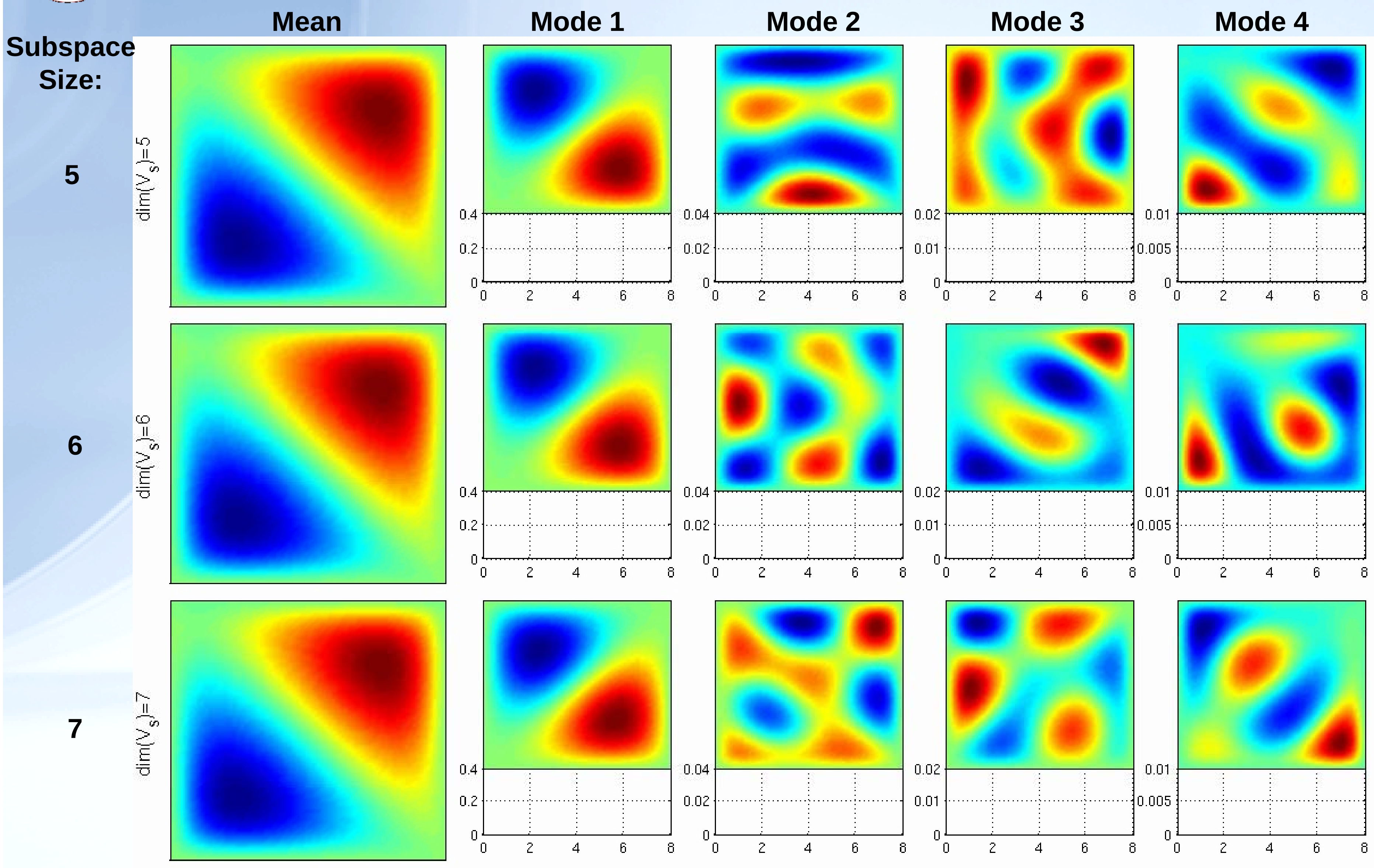
Energy of mean flow

Comparison with Monte-Carlo

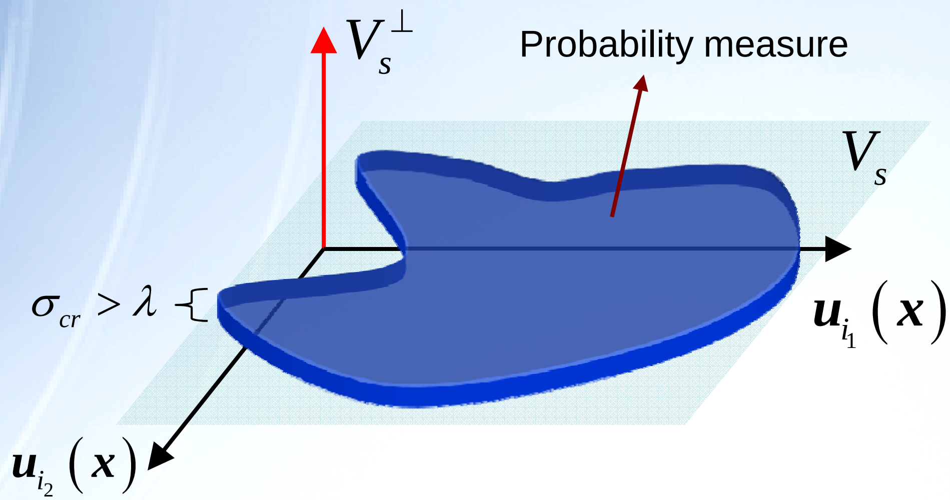


Comp. time: 11min (4000 samples or analytical Y_i) 12,3h (300 samples)

Example of Convergence Study



Adapt the stochastic subspace dimension



- In the context of DO equations so far the size of the stochastic subspace V_s remained invariant.
- For intermittent or transient phenomena the dimension of the stochastic subspace may vary significantly with time. This is accounted for by ESSE.

We need criteria to evolve the dimensionality of the stochastic subspace

This is a particularly important issue for stochastic systems with deterministic initial conditions

Criteria for dimension reduction / increase

Dimension Reduction

Comparison of the minimum eigenvalue of the correlation matrix $\boldsymbol{C}_{Y_iY_j}$.

$$\lambda_{\min}\left[\boldsymbol{C}_{Y_iY_j}\right] < \sigma_{cr} \longrightarrow \text{pre-defined value}$$

Removal of the corresponding direction from the stochastic subspace.

Dimension Increase

Comparison of the minimum eigenvalue of the correlation matrix $\boldsymbol{C}_{Y_iY_j}$.

$$\lambda_{\min}\left[\boldsymbol{C}_{Y_iY_j}\right] > \Sigma_{cr} \longrightarrow \text{pre-defined value}$$

Addition of a new direction $\boldsymbol{u}_i\left(\boldsymbol{x},t\right)$ in the stochastic subspace V_s.

How do we choose this new direction ?

Same problem when we start with deterministic initial condition
(dimension of stochastic subspace is zero)

Analytical criteria for selection of new directions

Theorem 2: For a stochastic field described by the evolution equation

$$\frac{\partial \boldsymbol{u}(\boldsymbol{x},t;\omega)}{\partial t} = \mathcal{L}\left[\boldsymbol{u}(\boldsymbol{x},t;\omega);\omega\right] \quad , \quad \boldsymbol{x} \in \boldsymbol{D}$$

and with current state at $t = t_c$ described by

the maximum variance growth rate of a stochastic perturbation in $V_s^{\perp}$ will be given by

Frechet Derivative

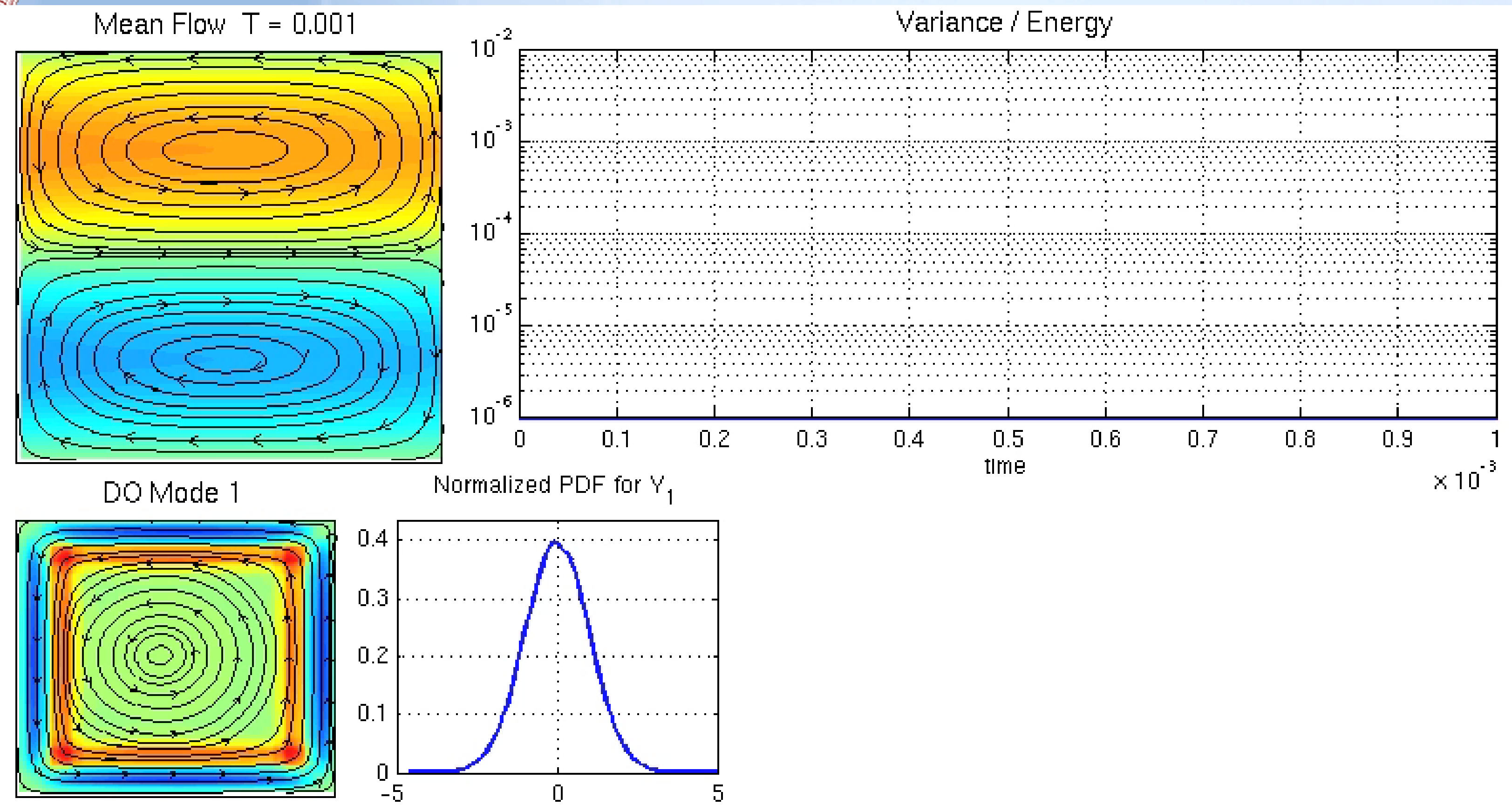
$$\rho\left[t_c;\boldsymbol{u}(\boldsymbol{x},t;\omega)\right] = \max_k \lambda_k \left[\frac{A_{ij}+A_{ji}}{2}\right], \qquad A_{ij} \equiv \int_{\boldsymbol{D}} \vartheta_j(\boldsymbol{y},t) \frac{\delta \mathcal{L}\left[\boldsymbol{u}(\cdot,t;\omega)\right]}{\delta \boldsymbol{u}}\left[\vartheta_i(\boldsymbol{y},t)\right] d\boldsymbol{y}$$

where $\vartheta_i(\boldsymbol{y},t), \quad i = 1,\dots,m$ is a finite basis that approximates $V_s^{\perp}$.

The corresponding direction of maximum growth is given by

where $v_i, \quad i = 1,\dots,m$ is the eigenvector associated with ρ.

Example: Double Gyre, Re=10,000



Grand challenge II in Large Nonlinear Systems

Optimally sense the (ocean) system with large numbers of heterogeneous and autonomous vehicles that are smart

Smart Sensing Vehicle Swarms

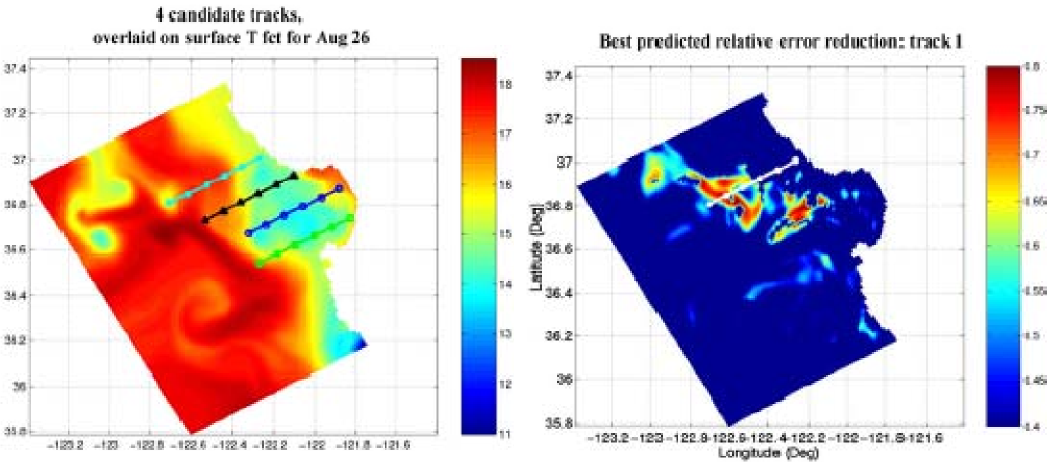
- Knowledgeable about the predicted (ocean) system and its uncertainties
- Knowledgeable about the predicted effects of their sensing on future estimates

Our collaborative experience …

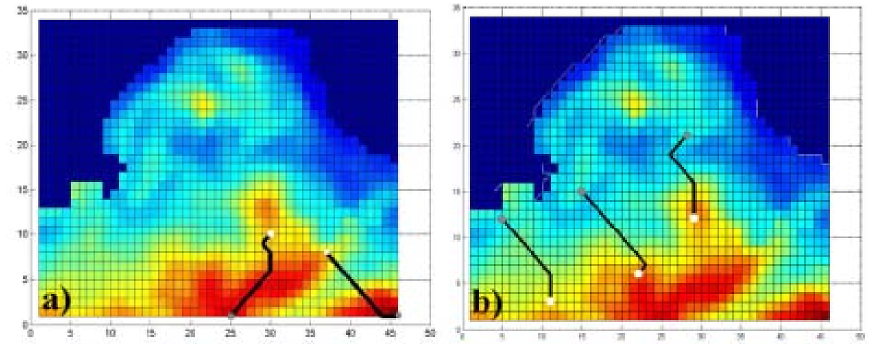
- Adaptive sampling via ESSE with non-linear predictions of error reductions
- Mixed Integer Linear Programming (MILP) for path planning
- Nonlinear path planning using Genetic algorithms
- Dynamic programming and onboard routing for path planning
- Command and control of vehicles over the Web, directly from model instructions

Adaptive Sampling Methodologies for Smart Robotic Swarms

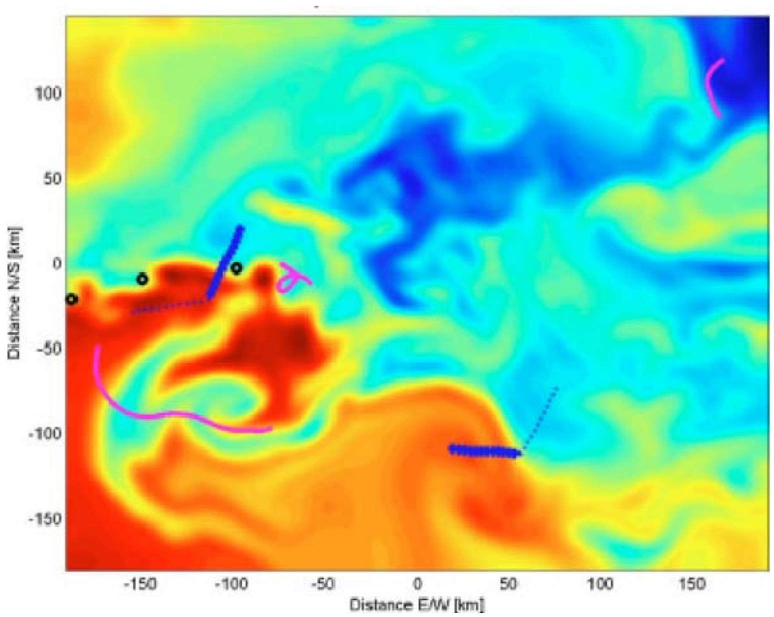
Lermusiaux et al



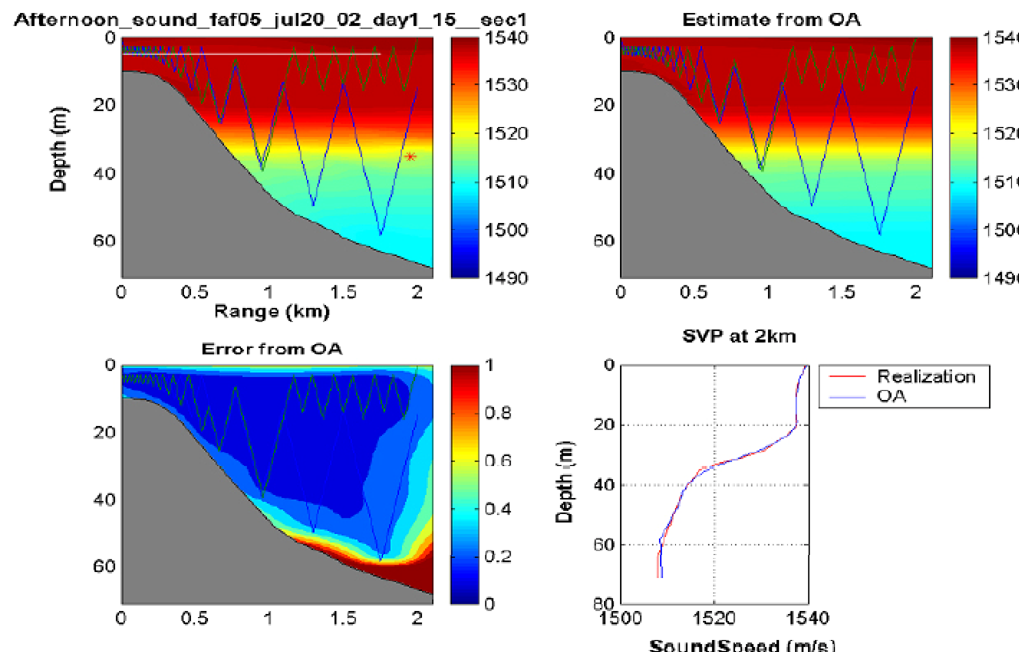
Adaptive Sampling via ESSE
[Lermusiaux, Phys.D-2007]



Optimal Path generation w/ 'fixed' objective field
[Yilmaz et al, OE-2008; Lermusiaux, et al 2007]



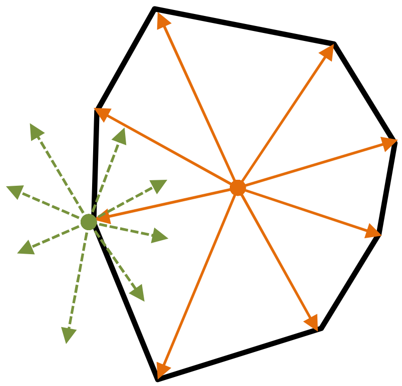
Genetic Algorithm
[Heaney et al., JFR-2007, OSSE OM-2010]



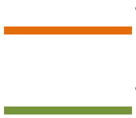
Path Planning based on Dynamic Programming
[Wang et al., Proc. IEEE-2006, JMS-2009]

Level Set Representation for Optimal Path Planning for Swarms in (strong) Currents

Advance many vehicles in many directions

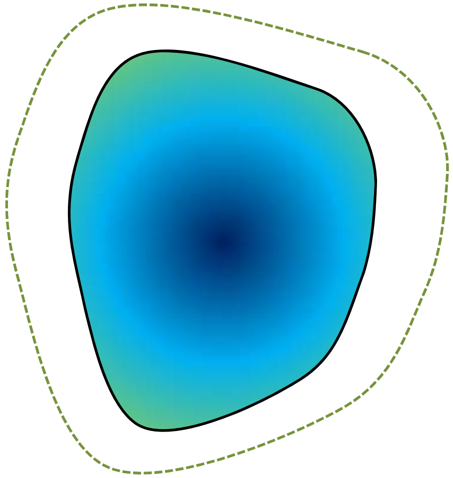


- Can lead to poorly defined curves
    - Only have to solve for 1D curve
- What to do with multiple vehicles?
- Exponential increase in cost



OR

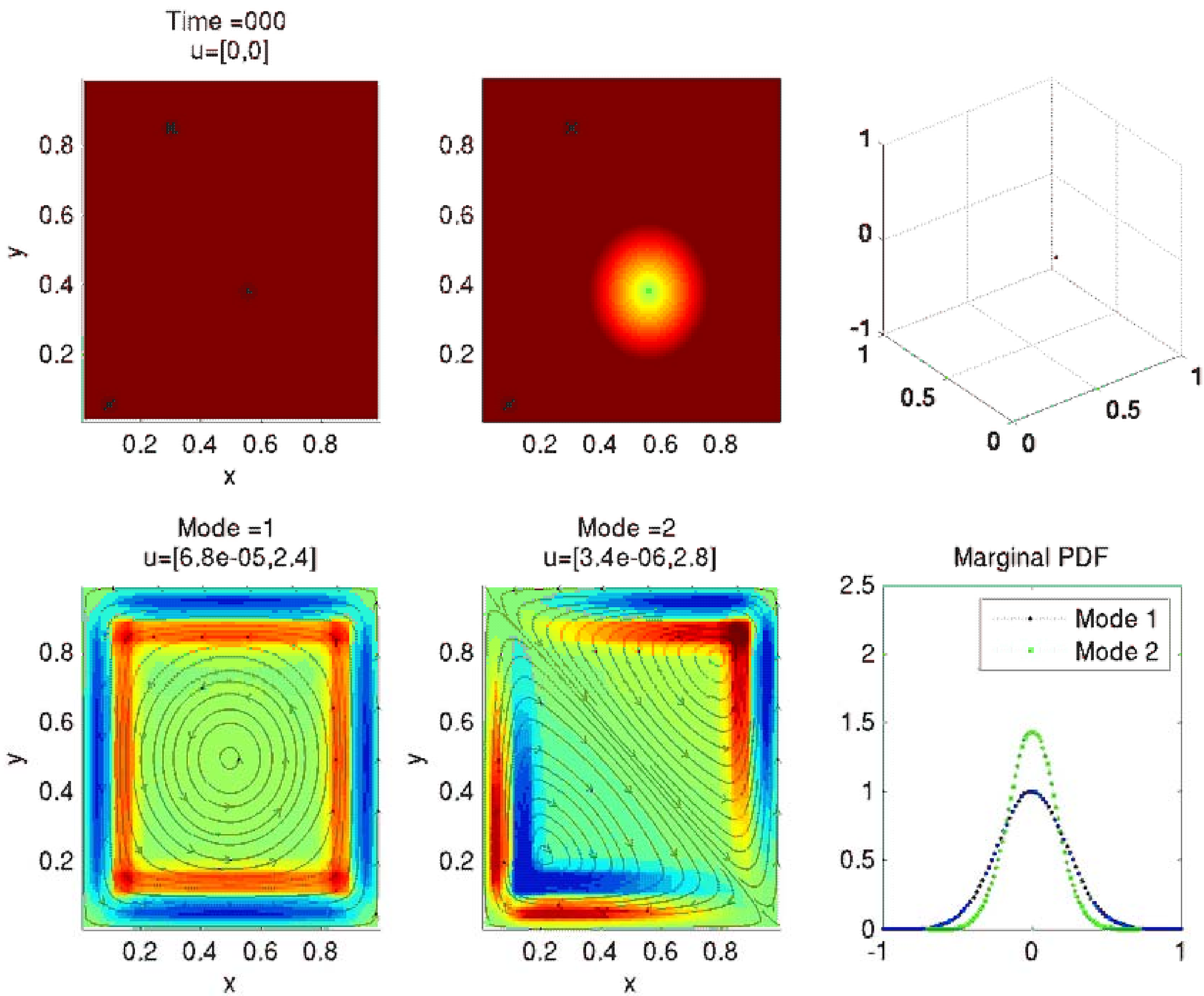
Represent vehicles 'front' with a level set



- Continuous representation
    - Need so solve 2D field
- Easily deals with multiple vehicles
- Front propagates normal to itself

Level Set Path Planning in currents : Uncertainty - DA

1) Utilize Uncertainty on Level Sets and/or 2) Uncertain level sets



Combine Partially Observable MDPs (POMDPs) with DO/ESSE equations for Adaptive Sampling

Key Idea: Steer UAVs using hierarchical "Partially Observable Markov Decision Processes"

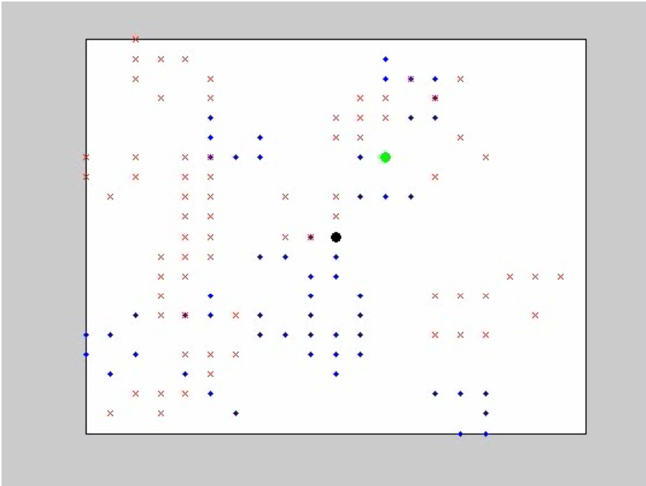
Examples of *global* goals may be:

- Track region / ocean feature
- Mimic swarming scheme
- Investigate region of large predicted uncertainties
- Combinations of the above

Goal: Maximize utility (Bellman Optimality Equation): $U_\infty(b) = \gamma \max_a \left(r(b,a) + \int U_\infty(b') p(b'|b,a) db' \right)$

Initial Simple Test Case: Game of Life.

- One AUV (**black circle**).
- One *global* goal (**green circle**).
- Many *local* goals:
  - "Good" cells (**blue dots**).
  - "Bad" cells (**red x's**).
- Multiple uncertainties:
  - observations and actions.



M. Gardner, *The fantastic combinations of John Conway's new solitaire game "life"*, Scientific American 223 (October 1970): 120-123.

DO equations and ESSE data assimilation

Data Assimilation (by Kalman update, combining DO uncertainty predictions with ESSE):

- Generate realizations:

$$\left\{u_r\left(x,t_o\right)\right\}_1^N = \bar{u}\left(x,t_o\right) + \left\{Y_i\left(t_o;\omega\right)\right\}_1^N u_i\left(x,t_o\right)$$

- Calculate Kalman Gain:

$$K = BH^T\left(R + HBH^T\right)^{-1}$$

- Perform Kalman update:

$$\left\{u_r^+\left(x,t_o\right)\right\}_1^N = \left\{u_r\left(x,t_o\right)\right\}_1^N + K\left(y\left(x,t_o\right) - H\bar{u}\left(x,t_o\right)\right)$$

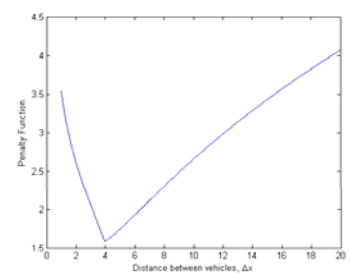
- Project back into D.O. framework:

$$u_i\left(x,t_o\right) = \phi_i\left(x,t_o\right); \quad Y_i = \left\langle \left\{u_r^+\left(x,t_o\right)\right\}_1^N, u_i\left(x,t_o\right)\right\rangle$$

Inter-vehicle communication/potential:

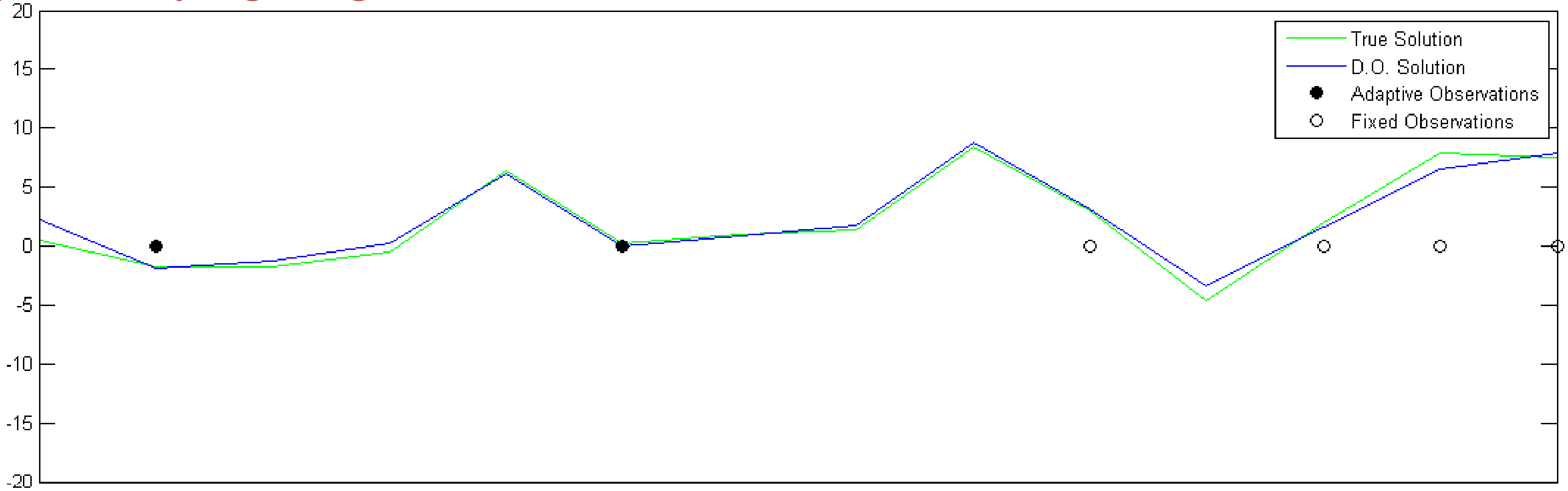
- Add penalty term to POMDP reward function:

$$Penalty = k\sqrt{\left|\Delta x - x_0\right| + 10/\left|\Delta x\right|}$$



Example: Lorenz-95

Adaptive Sampling using POMDP-like scheme:



Lorenz-95 Equation (with added diffusion):

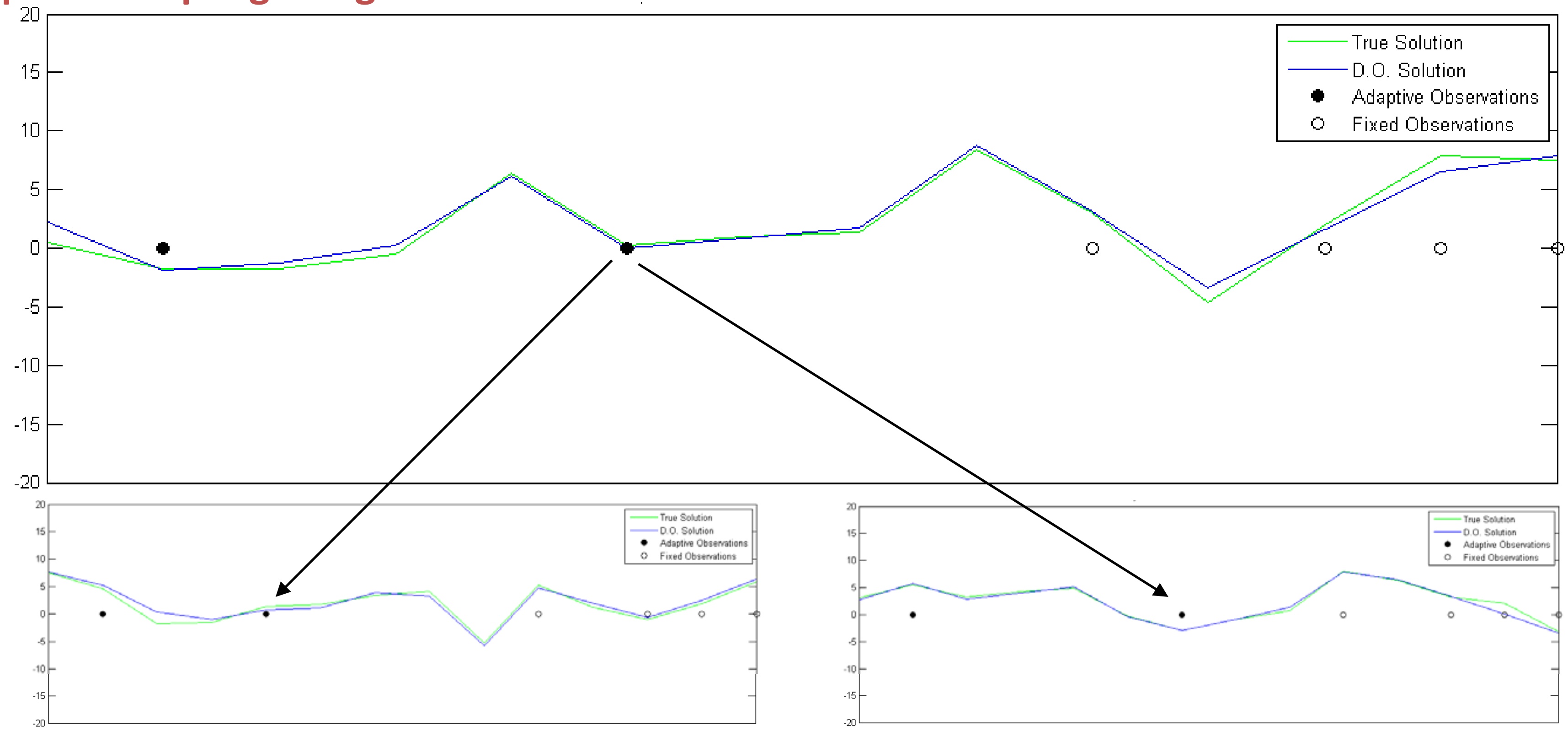
$$\frac{du_i}{dt} = \underbrace{u_{i-1}(u_{i+1} - u_{i-2})}_{advection} + \underbrace{k(u_{i+1} - 2u_i + u_{i-1})}_{diffusion} - \overbrace{u_i}^{dissipation} + \underbrace{f}_{forcing}$$

"40 ODEs: represent an atmospheric quantity at 40 sites spaced equally about a latitude circle.."

Where to make supplementary observations?

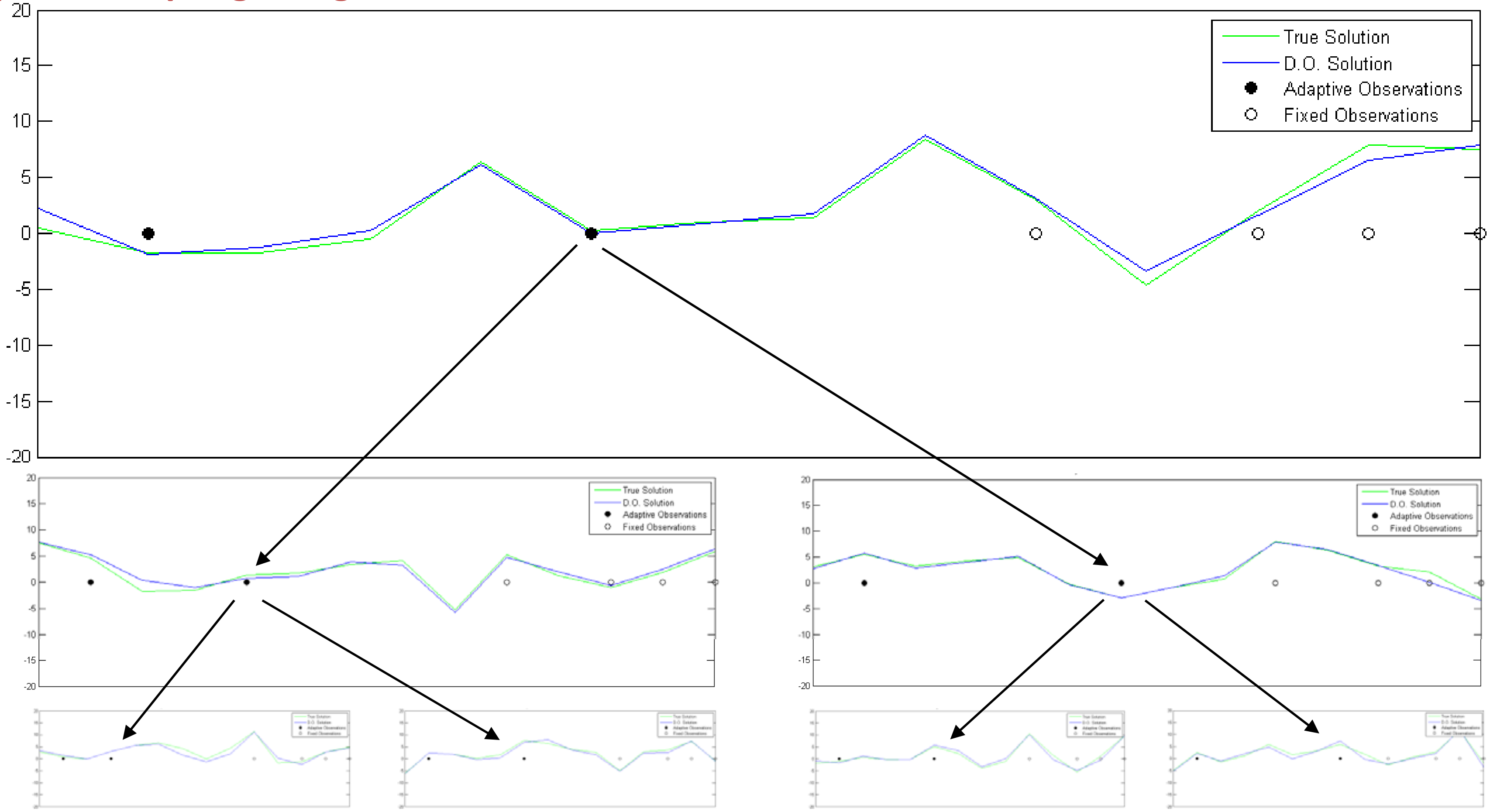
Example: Lorenz-95

Adaptive Sampling using POMDP-like scheme:



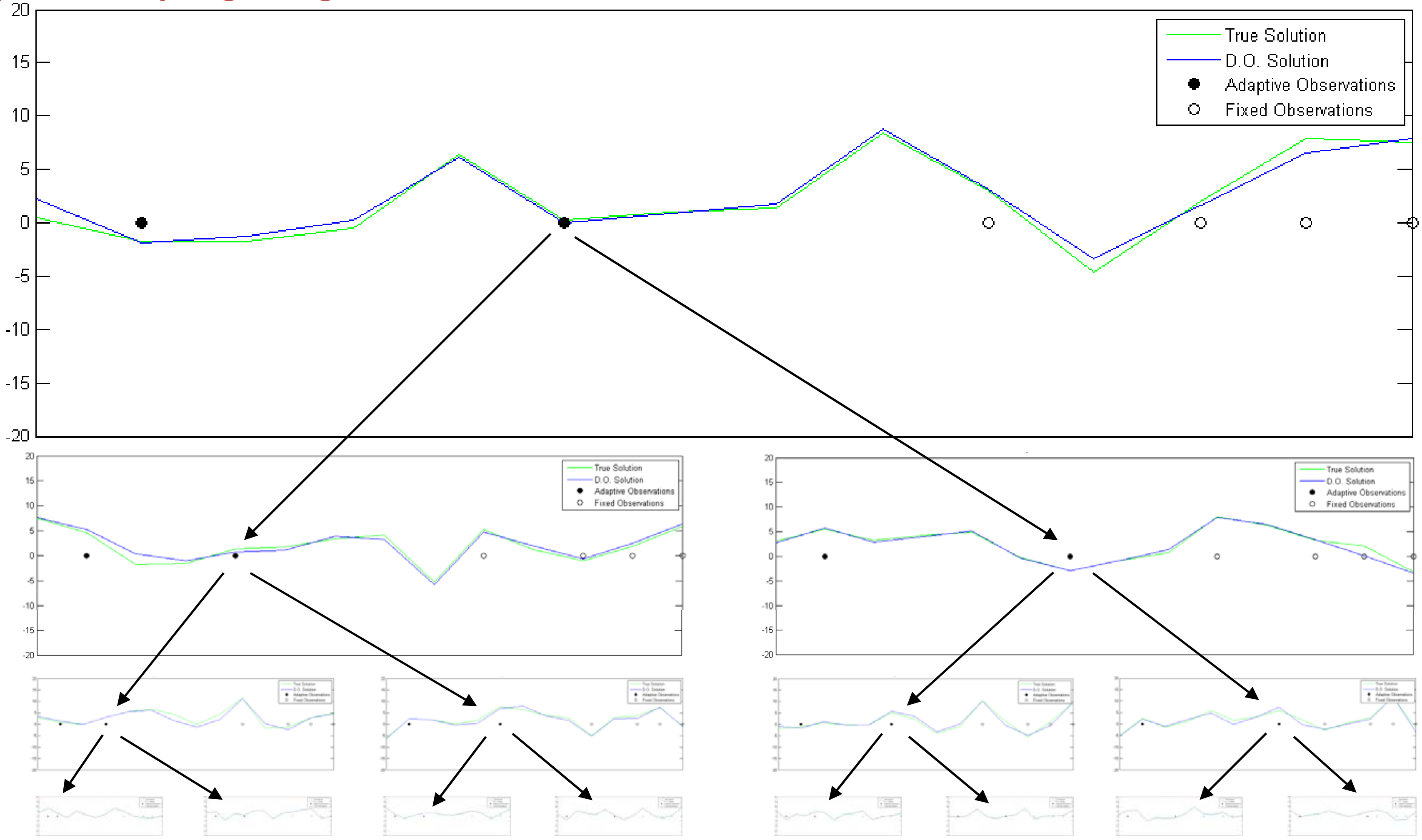
Example: Lorenz-95

Adaptive Sampling using POMDP-like scheme:



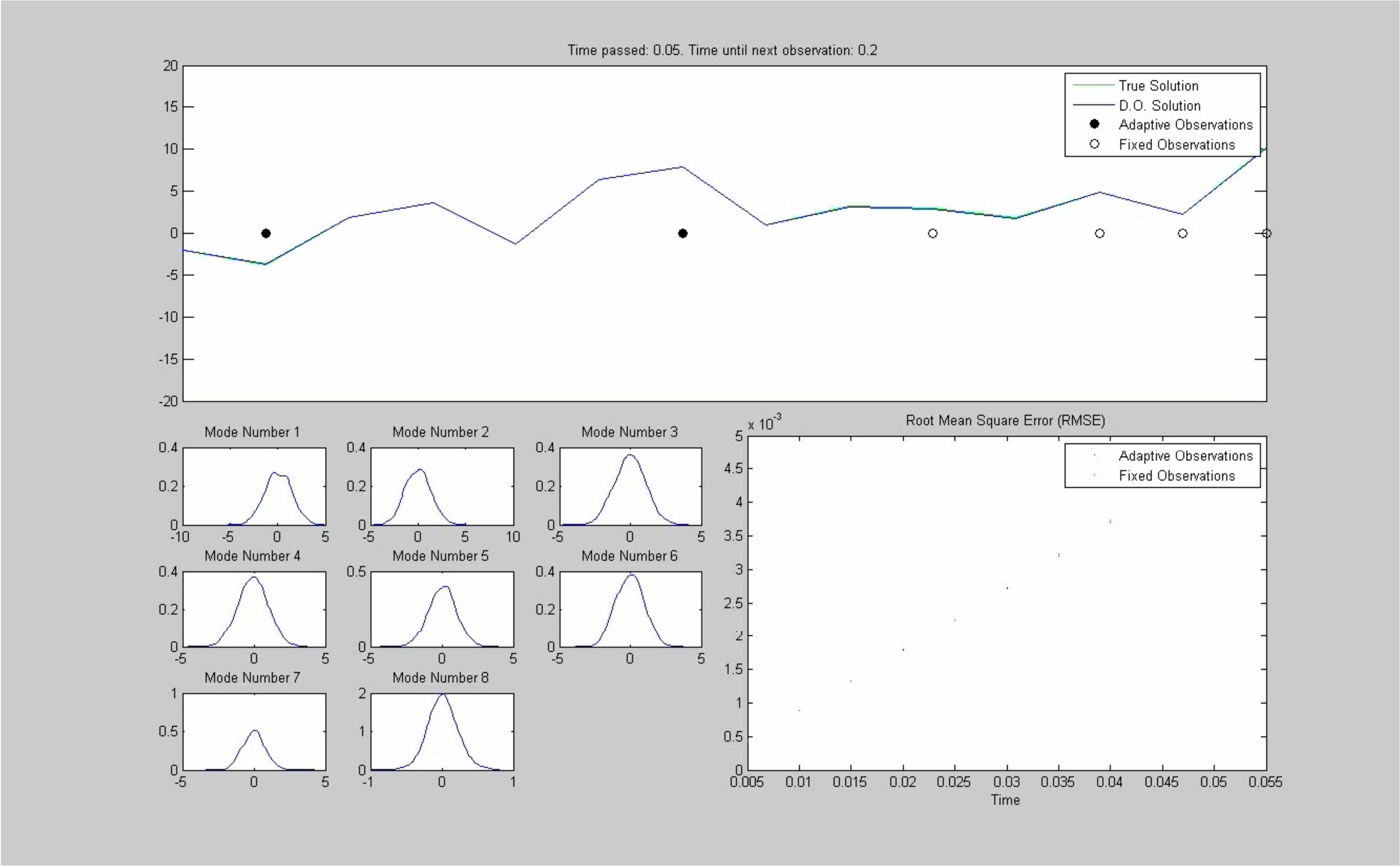
Example: Lorenz-95

Adaptive Sampling using POMDP-like scheme:



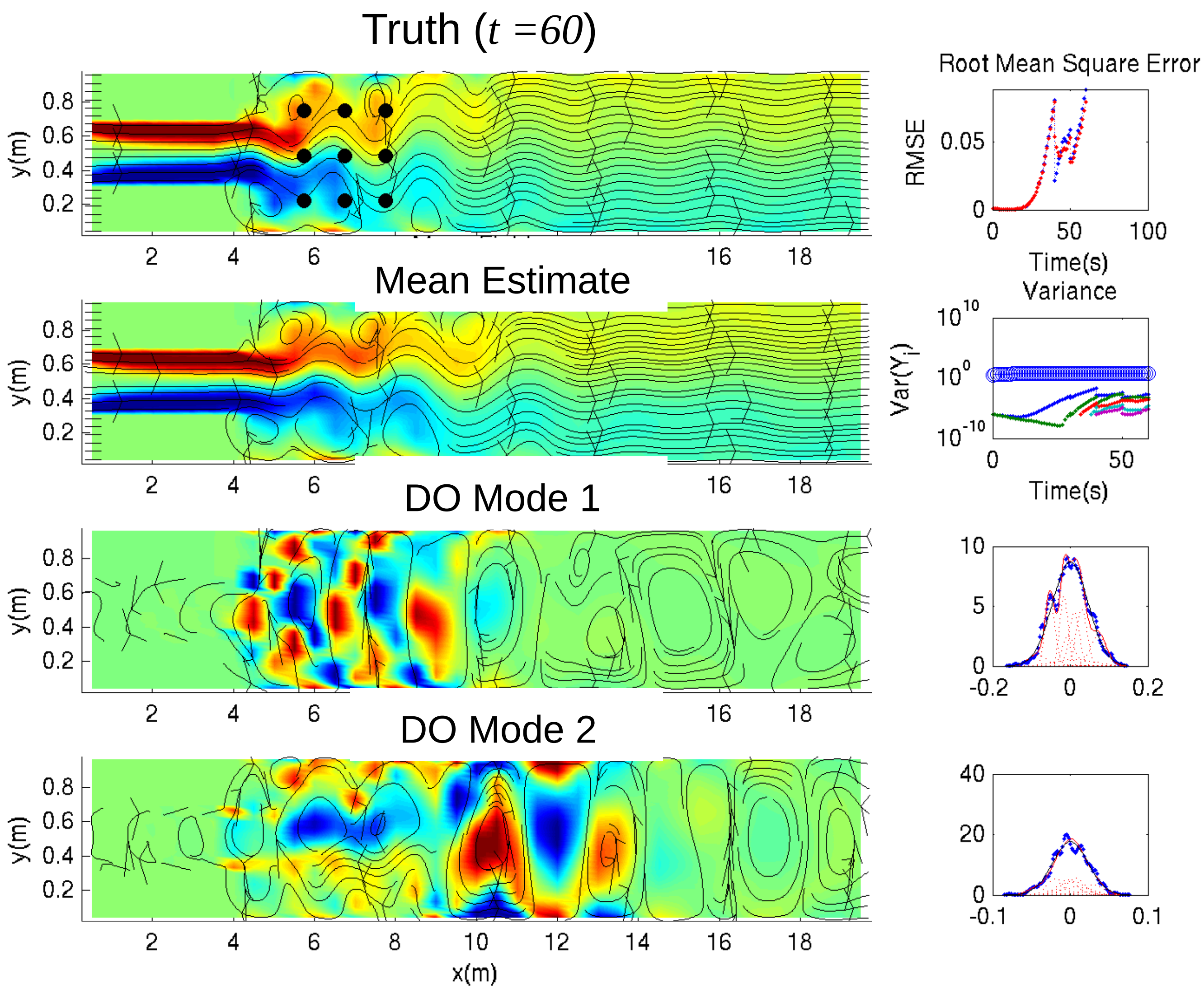
Example: Lorenz-95

- Adaptive Observations (overlaid on Fixed observations)-

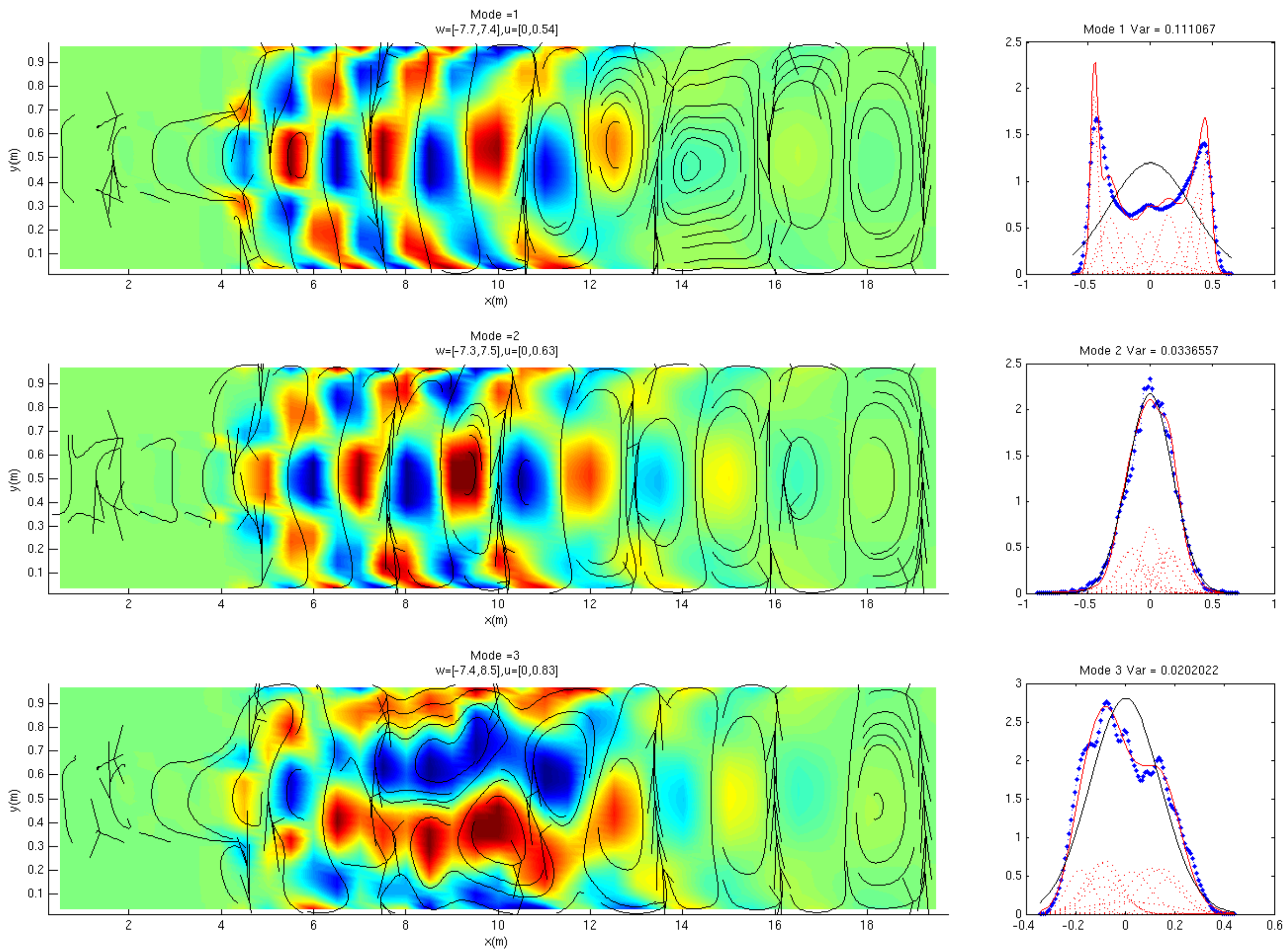


Non-Gaussian Data Assimilation: Sudden Expansion in a Pipe

- Fit Gaussian mixtures to coefficients Y_i
- EM-algorithm to obtain maximum likelihood fit
- Bayesian Information Criterion to select number of Gaussians



Non-Gaussian Data Assimilation: Sudden Expansion in a Pipe



CONCLUSIONS – DO equations

- **New Prognostic Equations for Stochastic Fields**
  - Derived new closed DO field equations (applied to several 2D NS/cases)
  - Adapted the size of the subspace (as in ESSE)
  - Multiscale information/energy flows between mean and stochastic-modes

- **DO Data Assimilation (Bayesian + Information Theory)**

- **Intelligent Adaptive Sampling: the Science of Autonomy**
  - Developed and utilized varied Adaptive Sampling schemes
  - Path Planning for Sensing Swarms using Level Set Methods
  - Merging adaptive DO equations, ESSE-Data-Assimilation and POMDPs for Smart Adaptive Sampling in 1D

CONCLUSIONS – DO equations

❖ Ongoing Research:

- Idealized Climate – MOC uncertainty
- DO and non-Gaussian DA with more realistic ocean fields
- Continue combination of “ESSE+DO + Level Set + Information theory” for collaborative Sampling Swarms
- Evolve the subspace based on data (learning as in ESSE)
- “DO Numerics” manuscripts
  - Direct cost of DO eqns:

$$O(s^2) \times (\text{Cost of Determ. PDE})$$

  - With projection methods, cost reduces to:

$$O(s) \times (\text{Cost of Determ. PDE})$$

Impose all modes to be incompressible and solve for pseudo-pressures that contain all cross-pressure terms

Note: looking into non-intrusive methods too