ERROR PROPAGATION AND VARIATIONAL DATA ASSIMILATION

F.-X. LE DIMET

MOISE project INRIA , LJK, University of Grenoble, France

With

Y. HUSSAINI *(Florida State University),* ***D.FURBISH*** *(Vanderbilt U.),*

P. NGNEPIEBA *(FAMU)*

I. GEJADZE *(Strathclyde U., Glasgow) ,* ***V. SHUTYAEV*** *(Russian Academy of Sciences, Moscow)*

MAX PLANCK INSTITUTE, DRESDEN

Data Assimilation

- Retrieving the state of a geophysical fluids requires to use all the available information:
  - Mathematical information : model
  - Physical Information : data
  - Statistical information
  - A-priori information
  - Images.

  One source of information is not sufficient to carry out a non trivial prediction (e.g. persistency). Therefore the Problem of Data Assimilation can be stated : how to gather these sources of information for retrieving « at best » the state of the system?

ERRORS

- All sources of information have errors and often (at least in meterorology and oceanography where there are operational prediction) some partial information is known on the statistics of these errors.
- Question : how the errors propagate toward the retrieved state, then to the prediction?

Optimal Control (variational) approach

$$\frac{d\varphi}{dt} = F(\varphi,\lambda) + f$$

$$\varphi(0) = u$$

$$S(u,\lambda,f) = \frac{1}{2}\left(V_b^{-1}(u-u_b),(u-u_b)\right) + \frac{1}{2}\int_0^T V_O^{-1}\left((C\varphi-\varphi_{obs}),(C\varphi-\varphi_{obs})\right)dt +$$

$$\frac{1}{2}\int_0^T V_m^{-1}\left(\left(\lambda-\lambda_g\right),\left(\lambda-\lambda_g\right)\right)dt + \frac{1}{2}\int_0^T V_f\left(\left(f-f_g\right),\left(f-f_g\right)\right)dt$$

Adjoint Model and Gradients: Euler-Lagrange equations

$$\frac{d\eta}{dt} + \left[\frac{\partial F}{\partial \varphi}\right]^T \eta = C^T V_b^{-1}\left(C\varphi - \varphi_{obs}\right)$$

$$\eta(T) = 0$$

$$\nabla_u S = -\eta(0) + V_b^{-1} C(\varphi - \varphi_{obs})$$

$$\nabla_f S = V_f^{-1}\left(f - f_b\right) - f\eta$$

$$\nabla_\lambda S = V_\lambda^{-1}\left(\lambda - \lambda_b\right) + \left[\frac{\partial F}{\partial \lambda}\right]^T .\eta$$

OPTIMALITY SYSTEM

- The **OPTIMALITY SYSTEM** = Model+Adjoint model **contains all the available information**
- Because of this property it has been be considered as a **GENERALIZED MODEL.**
- Data Assimilation is **NOT** just an algorithmic problem. Studying properties of O.S. is of great importance for understanding the properties of the fluids and propagation of information and uncertainties.
- As an example, sensitivity studies (e.g. with respect to observations) must be carried out on the O.S. not on the only model.

APPLICATION 1

- **Question 1**:For a given set of observations, will the « best » model give the best prediction?
- **Question 2** : with some qualitative information on the model error can we improve the prediction?

- ***Published in Tellus 2009 (Furbish, Hussaini, Le Dimet, Ngnepieba)***

Error Control *(see also Arthur Vidard's Ph.D.)*

- Introducing an error in the model

$$\begin{cases} \frac{d\mathbf{u}}{dt} = F(x,\mathbf{u}) + E(x,t), \\ \mathbf{u}(0) = U, \end{cases}$$

- In the cost function

$$\begin{aligned} J(U,E) &= \frac{\alpha}{2}\|U-U_0\|^2 + \frac{1}{2}\int_0^T \|H\mathbf{u} - X_{obs}\|^2 dt \\ &+ \frac{\beta}{2}\int_0^T (E,E)dt. \end{aligned}$$

- Adjoint

$$\begin{cases} \frac{dP}{dt} + \left[\frac{\partial F}{\partial \mathbf{u}}\right]^T \cdot P = H^T(H\mathbf{u} - X_{obs}), \\ P(T) = 0, \end{cases}$$

Gradient

$$
\begin{cases}
\nabla_U J(U,E) = \alpha(U - U_0) - P(0), \\
(\nabla_E J, \delta E) = \int_0^T \int_\Omega \left(- \left[\frac{\partial (F(x, \mathbf{u}(t)) + E)}{\partial E} \right]^T P(t) \right. \\
\qquad \left. + \beta E\,(x,t)\,, \delta E \right) dx dt.
\end{cases}
$$

Discretization of the error

- Discretization in time and space of the error:

$$\begin{cases} \frac{d\mathbf{u}}{dt} = F(\mathbf{u}) + \sum_{i,j} \varepsilon_{ij}\phi_j(t)X_i \\ \mathbf{u}(0) = U, \end{cases}$$

$$\begin{aligned} J(U,\Gamma) &= \frac{\alpha}{2}\|U - U_0\|^2_{\mathcal{H}} + \frac{1}{2}\int_0^T \|H\mathbf{u} - X_{obs}\|^2_{\mathcal{H}_{obs}} dt \\ &+ \frac{\beta}{2}\sum \varepsilon^2_{ij}, \end{aligned}$$

$$\nabla_{\varepsilon_{ij}} J(U,\Gamma) = \beta\varepsilon_{ij} - \int_0^T \phi_j(t)\,(X_i, P)\,dt.$$

$$\nabla_U J(U,\Gamma) = \alpha\,(U - U_0) - P\,(0)$$

Choices of the spaces

- Discretization Error

$$E \approx \sum_{p} \varepsilon(p) \nabla^{p} \mathbf{u}.$$

- With a first order scheme the discretization will depend on second order, e.g. the error could be represented in a basis of eigenvectors of the Laplacian. Without information, the eigenvectors of the covariances matrix can be considered:

$$\begin{cases} \dfrac{d\mathbf{u}}{dt} = F(x, \mathbf{u}) + \sum_{i,j} \varepsilon^{1}_{ij} \phi^{1}_{j}(t) X^{1}_{i} + \sum_{i,j} \varepsilon^{2}_{ij} \phi^{2}_{j}(t) X^{2}_{i} + \cdots \\ \mathbf{u}(0) = U, \end{cases}$$

VDA with Burger's equation

$$\begin{cases} \dfrac{\partial u}{\partial t} + \dfrac{1}{2}\dfrac{\partial (u^2 + au)}{\partial x} - \mu \left[\dfrac{\partial^2 u}{\partial x^2}\right] = f(x), \\ \qquad\qquad (x,t) \in (-1,1) \times (0,\infty), \\ \text{Boundary Condition: } u(-1,t) = u(+1,t) = 0 \ \forall t \in (0,\infty), \\ \text{Initial Condition: } u(x,0) = U, \\ J(U) = \inf_v J(v). \end{cases}$$

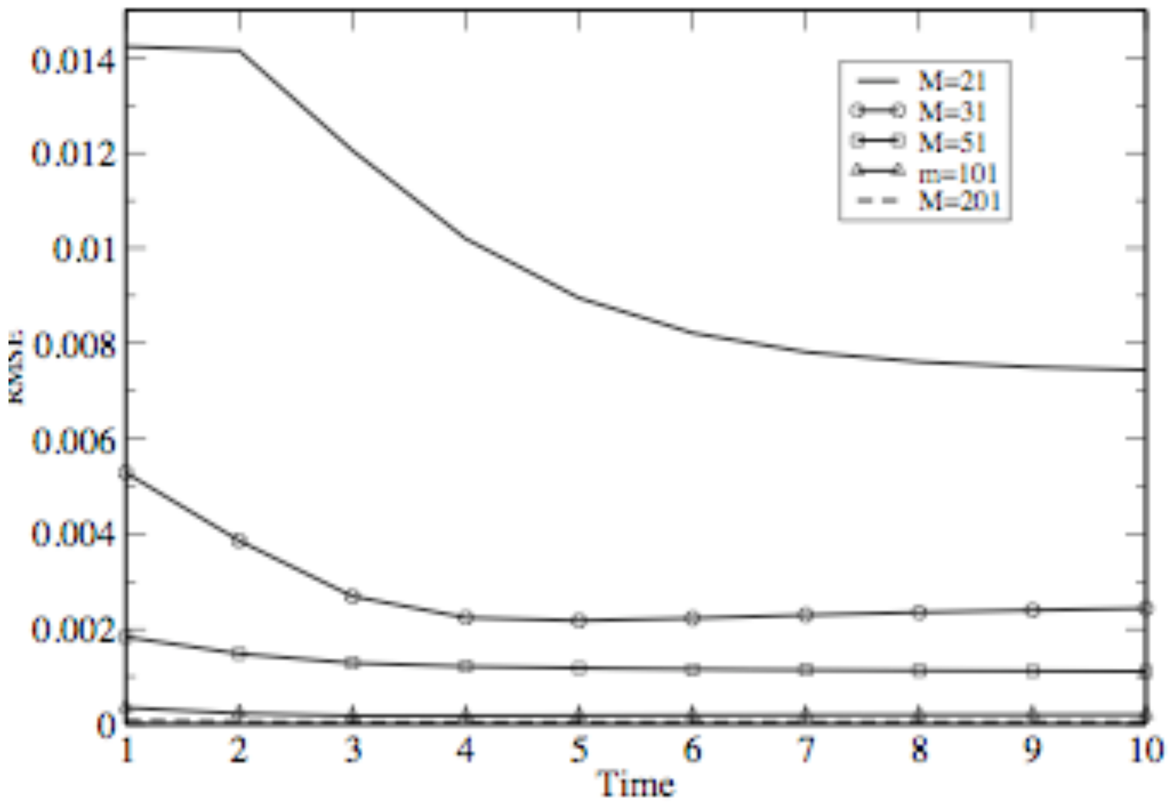
$$J(U) = \frac{1}{2}W_1 \sum_{i=2}^{M-1} (U^i - U_0^i)^2 + \frac{1}{2}W_2 \sum_{s=1}^{2}\sum_{k=1}^{K_s} (H(u(t_s,x)))^k - X_s^k)^2.$$

Burger's (2)

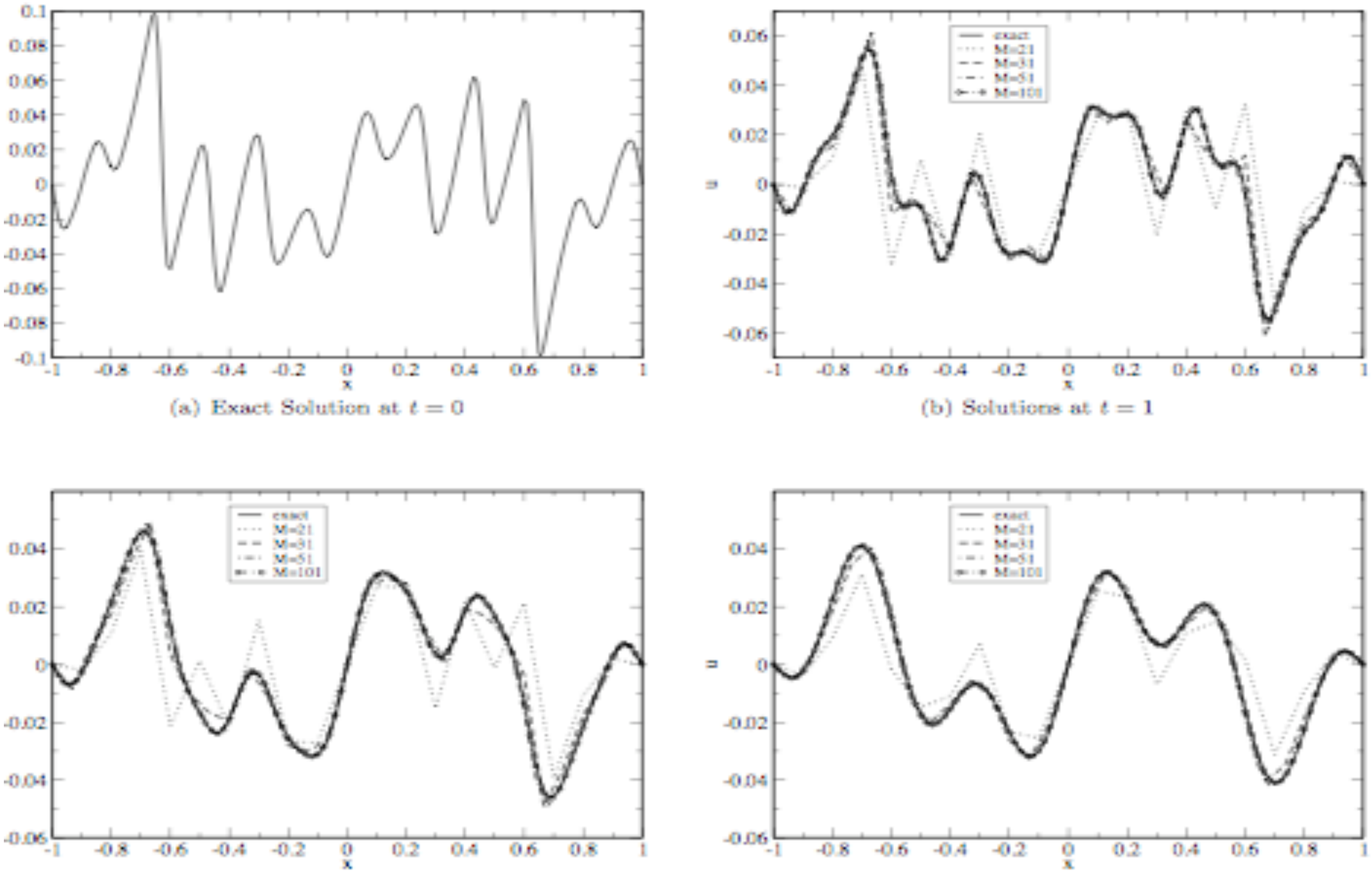
- F is choosen in such a way that the exact solution is known.
- Background term=exact solution+gaussian perturbation.
- Observations generated by the exact solution on a given regular grid.
- H is an interpolation operator.
- Consequently there are only two sources of error : discretization and interpolation.

Discretization Error

$$\text{RMSE} = \sqrt{\frac{1}{M}\sum_{i=1}^{M}(u_{ex}(i \bigtriangleup x, n \bigtriangleup t) - u_i^n)^2},$$



Exact and discretized solutions



(a) Exact Solution at $t = 0$ (b) Solutions at $t = 1$

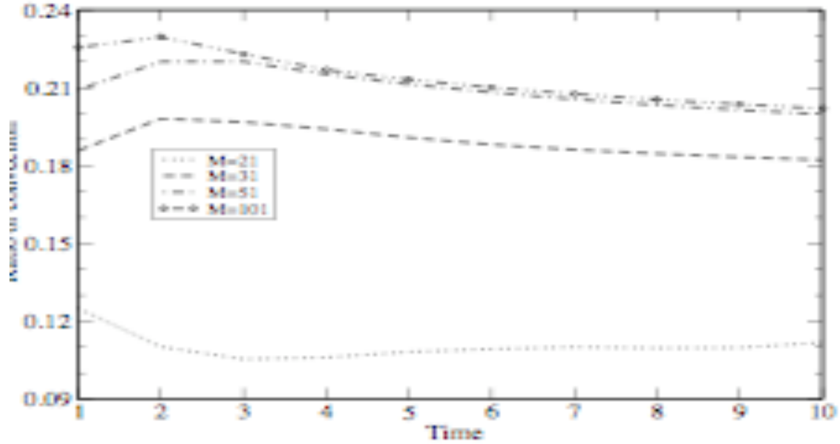
(c) Solutions at $t = 2$ (d) Solutions at $t = 5$

Convergence of D.A.

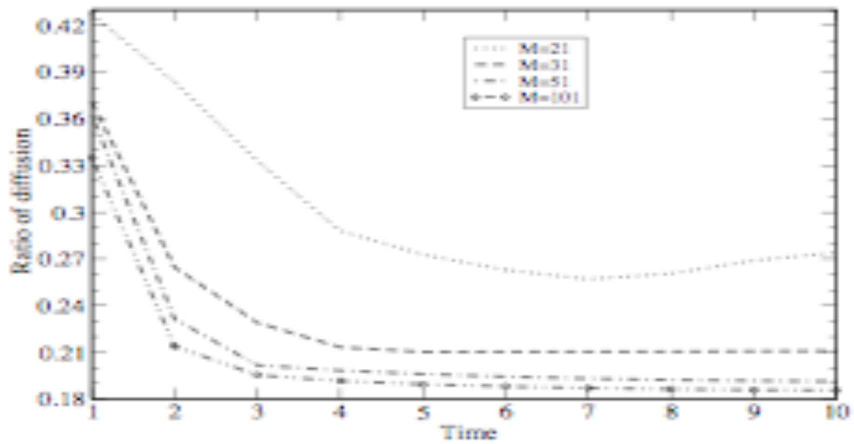
Table 1. Convergence of both the cost function and the L^2-norm of its gradient in the case $M = 101$ and $\sigma = 0.025$.

iteration	J	$\|\nabla J\|$
0	0.35078592×10^{1}	0.17191152×10^{3}
1	0.35058595×10^{1}	0.17180853×10^{3}
2	$0.76635219 \times 10^{-1}$	0.30238631
3	$0.76601067 \times 10^{-1}$	0.30068101
4	$0.76298041 \times 10^{-1}$	0.28555660
5	$0.70462862 \times 10^{-1}$	$0.12666941 \times 10^{-2}$
6	$0.70451991 \times 10^{-1}$	$0.54982245 \times 10^{-3}$
7	$0.70446412 \times 10^{-1}$	$0.47975452 \times 10^{-4}$
8	$0.70446346 \times 10^{-1}$	$0.44660779 \times 10^{-4}$
9	$0.70445638 \times 10^{-1}$	$0.10635374 \times 10^{-4}$
10	$0.70445471 \times 10^{-1}$	$0.20295671 \times 10^{-5}$
11	$0.70445438 \times 10^{-1}$	$0.67729479 \times 10^{-7}$
12	$0.70445438 \times 10^{-1}$	$0.56747292 \times 10^{-8}$
13	$0.70445437 \times 10^{-1}$	$0.23630726 \times 10^{-10}$
14	$0.70445437 \times 10^{-1}$	$0.79286146 \times 10^{-12}$
15	$0.70445437 \times 10^{-1}$	$0.23419518 \times 10^{-13}$

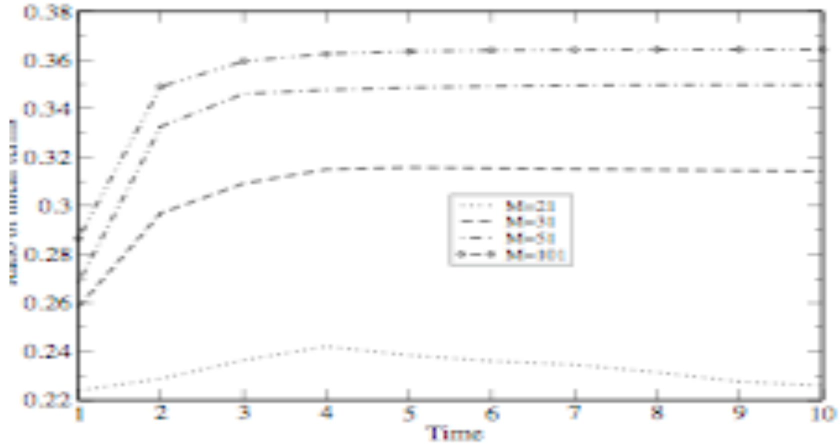
Evolution of various terms according to the discretization:convection,diffusion, linear term and forcing term



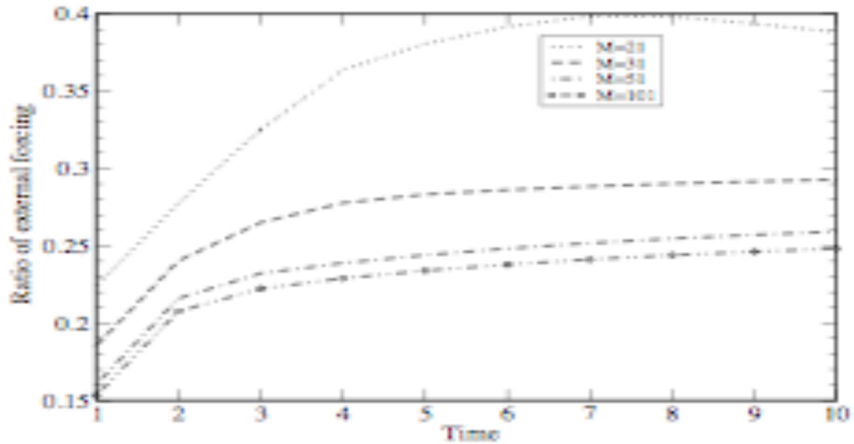
(a) Ratio of convection



(b) Ratio of diffusion



(c) Ratio of linear terms



(d) Ratio of external forcing

Influence of model resolution on D.A.

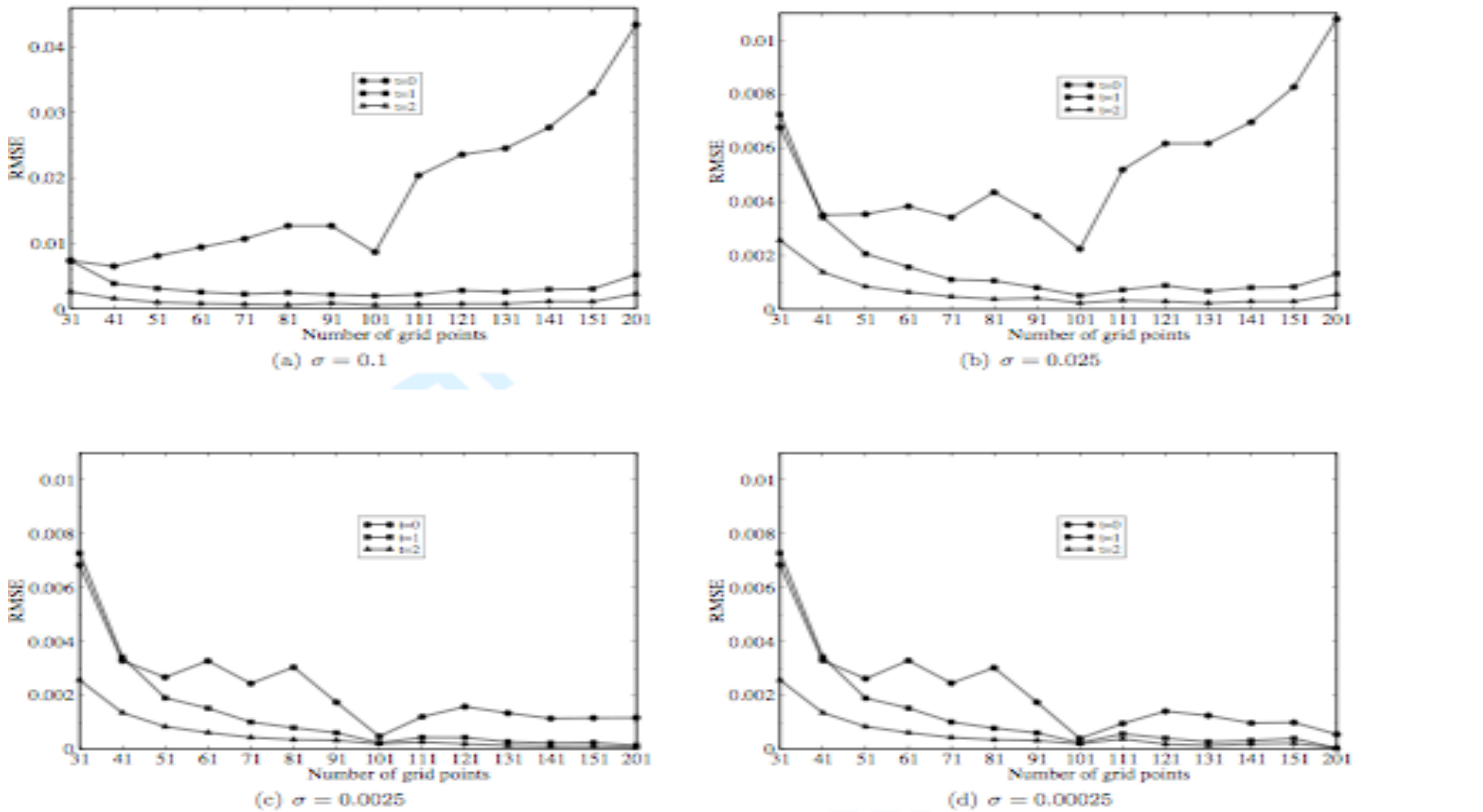


Figure 4. RMSE of numerical solutions initialized with optimal initial condition at times $t = 0$, 1, and 2. The background field is the exact initial condition perturbed by Gaussian noise with mean zero and variance (a) $\sigma = 0.1$, (b) $\sigma = 0.025$, (c) $\sigma = 0.0025$, and (d) $\sigma = 0.00025$.

Influence of Observation Resolution on D.A.(1)

Table 2. RMSE of the numerical solution on 51-point grid.

obs number	49	99	199
$t = 0$	0.1597×10^{-2}	0.2952×10^{-2}	0.3144×10^{-2}
$t = 1$	0.1282×10^{-2}	0.1286×10^{-2}	0.1323×10^{-2}
$t = 0$	0.1277×10^{-2}	0.2835×10^{-2}	0.3110×10^{-2}
$t = 1$	0.1244×10^{-2}	0.1263×10^{-2}	0.1314×10^{-2}

Table 3. RMSE of the numerical solution on 101-point grid.

obs number	49	99	199
$t = 0$	0.4467×10^{-2}	0.2817×10^{-2}	0.1718×10^{-2}
$t = 1$	0.1204×10^{-2}	0.9562×10^{-3}	0.6872×10^{-3}
$t = 0$	0.1956×10^{-2}	0.4035×10^{-3}	0.2411×10^{-3}
$t = 1$	0.2838×10^{-3}	0.2428×10^{-3}	0.2552×10^{-3}

Influence of Observation Operator on D.A.(1)

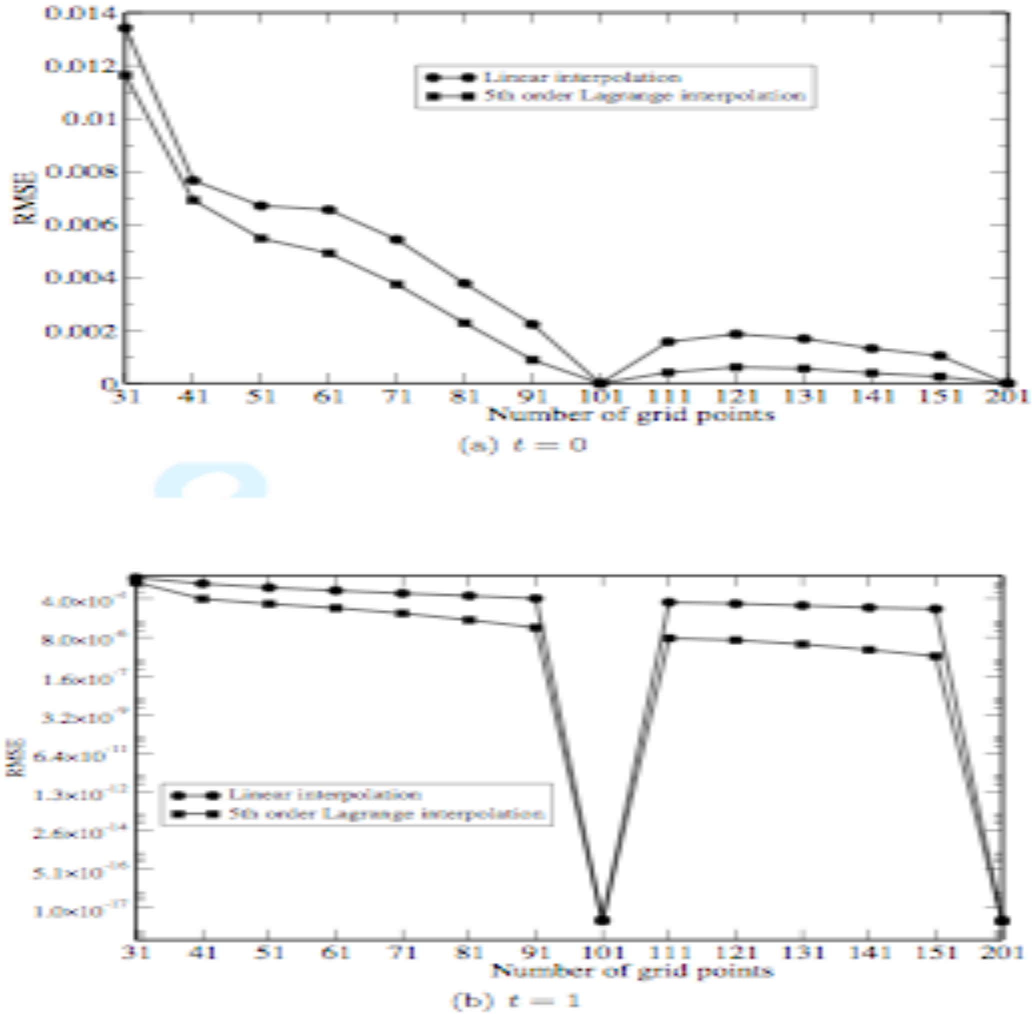
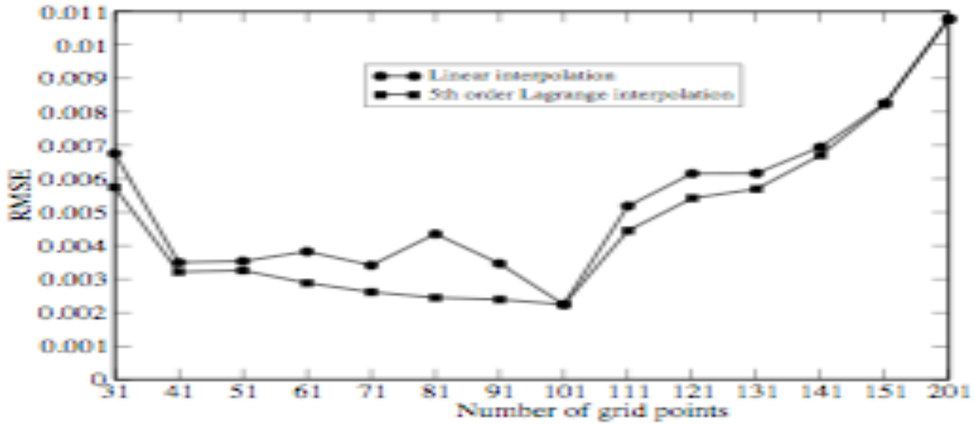
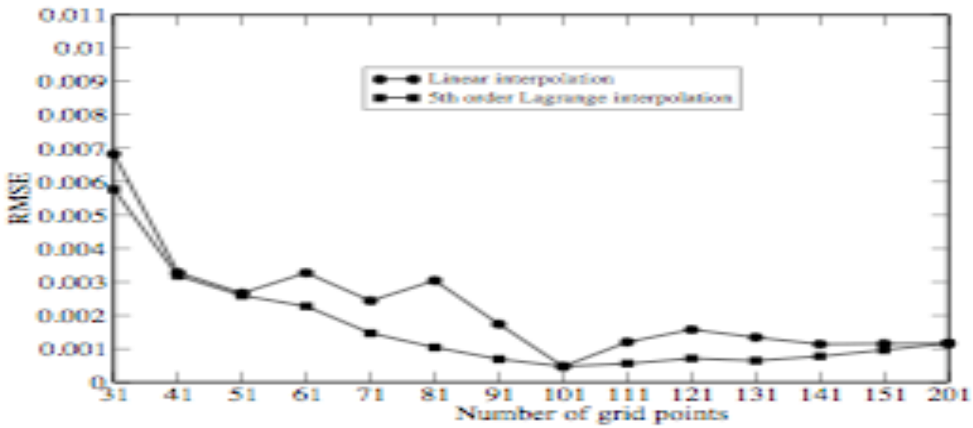


Figure 8. RMSE of the exact solutions interpolated on to the observation sites at (a) $t = 0$ and (b) $t = 1$.

Influence of Observation Operator on D.A.(2)



(a) $\sigma = 0.025$



(b) $\sigma = 0.0025$

e 9. RMSE of the optimal initial condition relative to the exact one with background perturbation variance (a) $\sigma = 0.025$, and $= 0.0025$.

Influence of Observation Operator on D.A.(3)

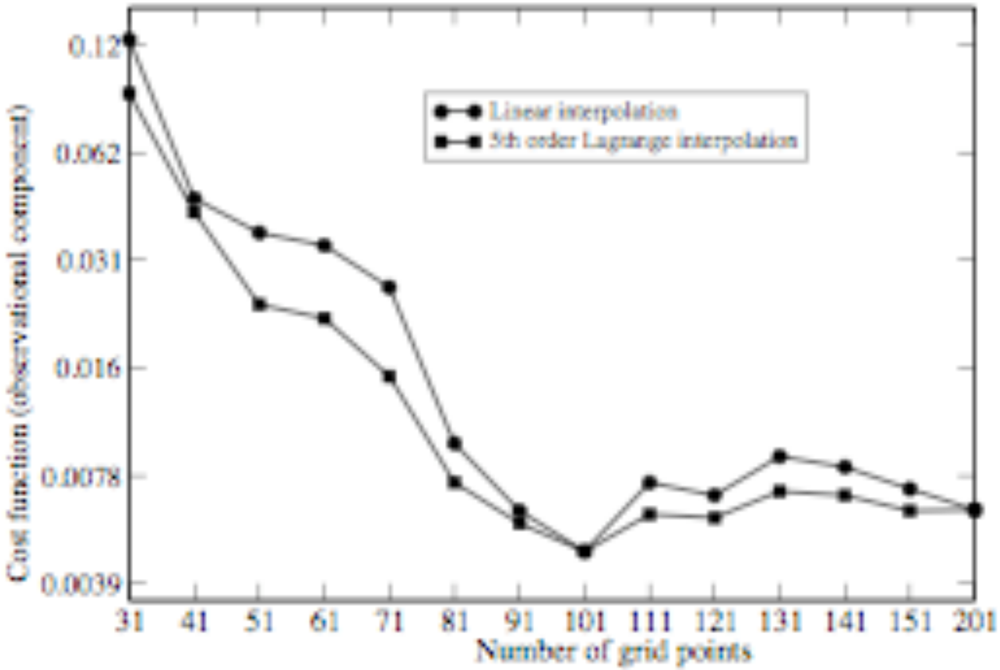
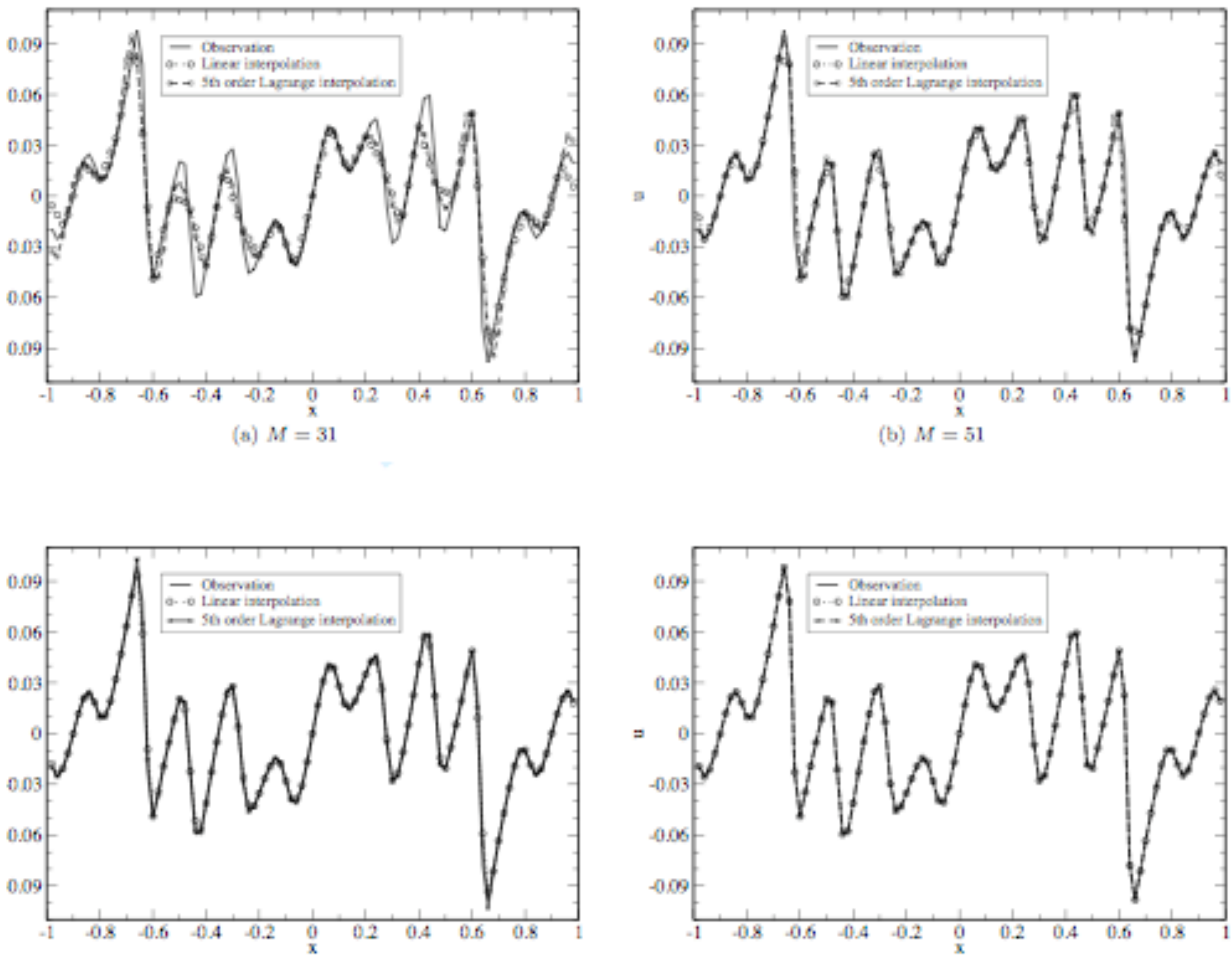


Figure 10. Observational cost function versus model resolution.



(a) $M = 31$

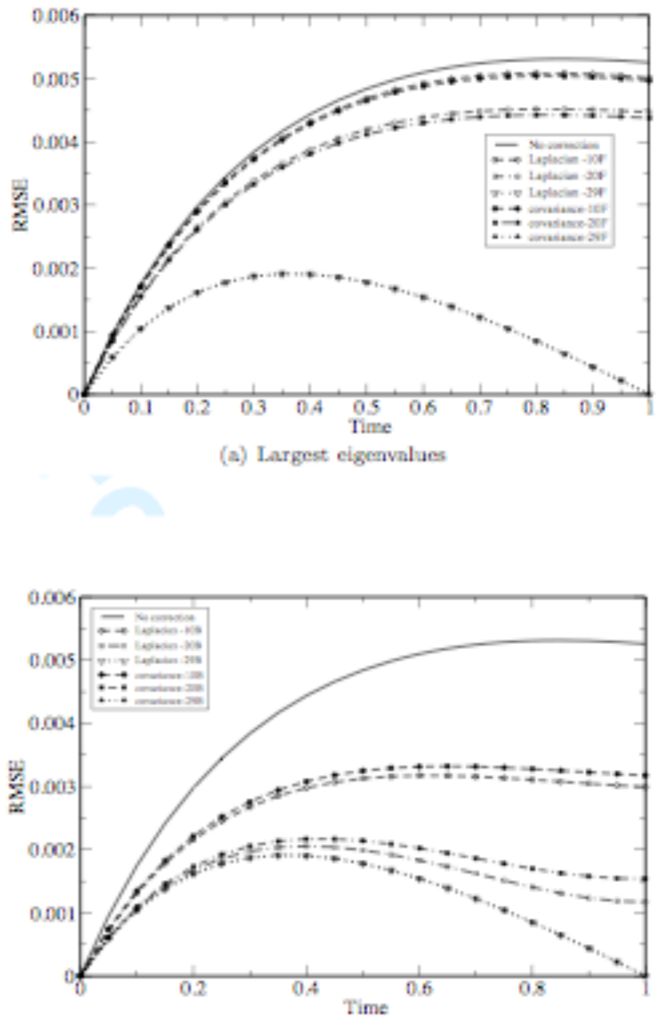
(b) $M = 51$

(c) $M = 81$

(d) $M = 101$

Figure 11. The observations and the exact solutions interpolated on to the observation sites.

Control of Errors (1)

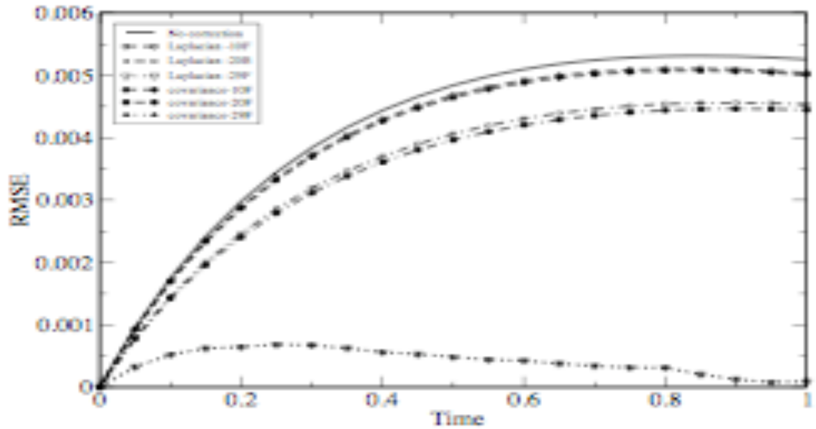


(a) Largest eigenvalues

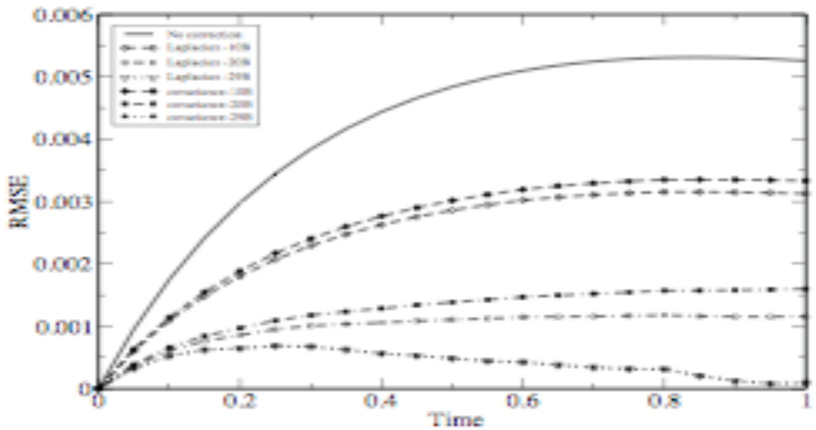
(b) Smallest eigenvalues

Figure 12. Comparison of the evolution of RMSE of the solutions for the modified model based on the eigenvectors of the Laplacian operator and the negative error covariance matrix. The model resolution is $M = 31$.

Control of Errors (2)



(a) Largest eigenvalues



(b) Smallest eigenvalues

Figure 13. Same as figure 12 with 5 piecewise-constant subintervals of the time domain (0, 1).

RESULTS (1)

- With coarse resolution discretization errors are dominant. The importance of background is small
- For an increased resolution, the optimal solution is a balance between background and observation. There exist an optimal resolution of the model.
- When the resolution is increased further beyong that of the observation the background error is dominant.

RESULTS (2)

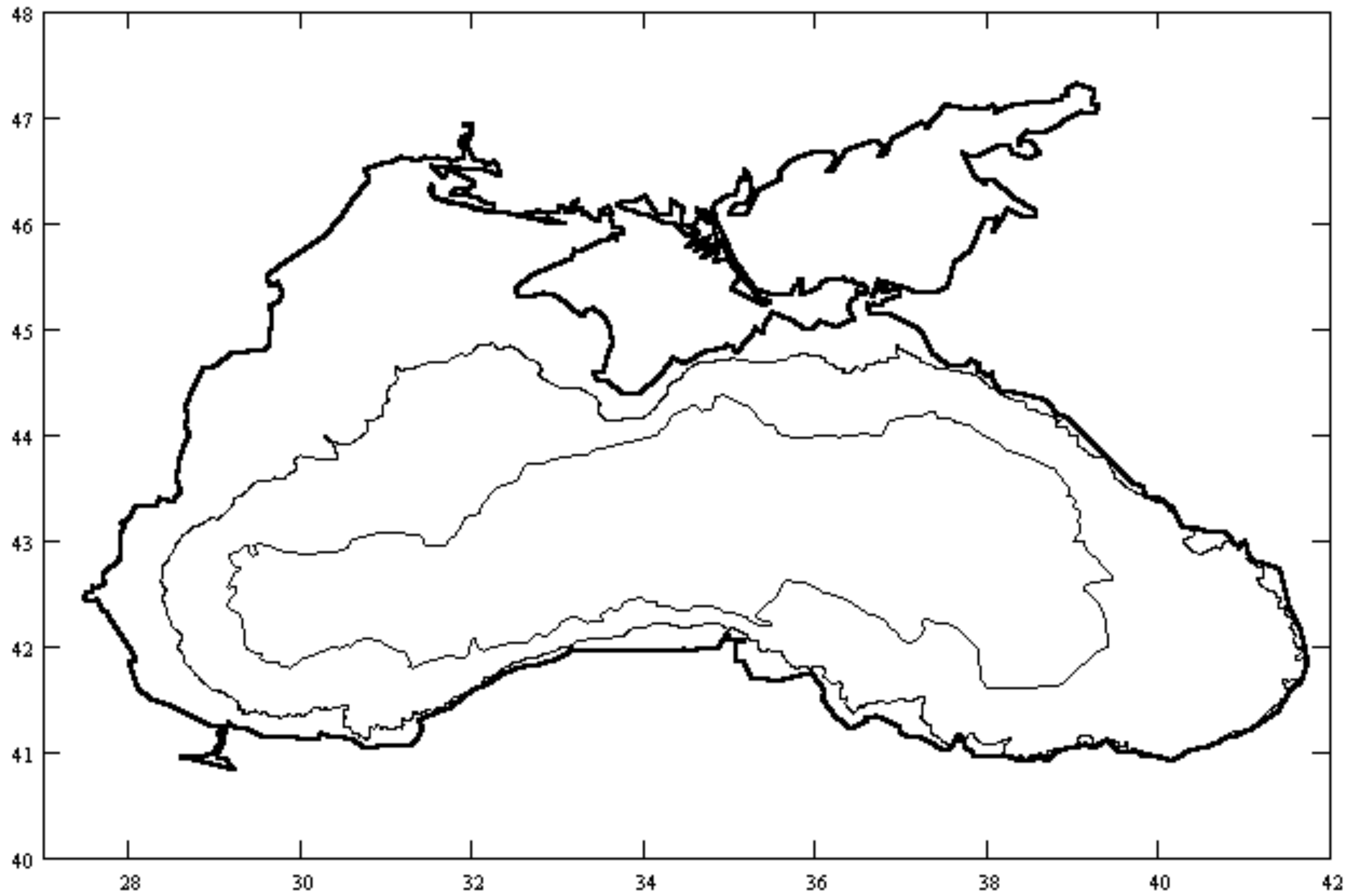
- Prediction can be improved by the control of errors. Without a priori knowledge on the source of errors, the eigenvectors of the covariance matrix are appropriate.
- Increasing resolution does not necessarily translate into improved assimilation. If the model solution is not sufficiently resolved high density of observationdegrades the data assimilation and prediction.
- Extension of the study to error on observations must be carried outon the Optimality System (cf. V. Shutyaev, I. Gejadze and FXLD)

APPLICATION 2

- Evaluation of the covariance of the initial condition (output) , the covariances of the inputs being known.

- *.I. Gejadze, F.-X. Le Dimet, V. Shutyaev published in SIAM J. on Scientific Computing (2009) and J. of. Comp. Phys. (2010)*

A remark on ensemble methods



Computing Covariances : Problem statement

Model of evolution process

$$\begin{cases} \partial\varphi/\partial t = F(\varphi) + f\,, \quad t \in (0,T) \\ \varphi\big|_{t=0} = u \end{cases}$$

Posterior covariance

$$V_{\delta u} = E\left(\delta u\, \delta u^T\right)$$

$\delta u = u - \bar{u}$

$\bar{u} -$ 'true' state

Objective function (for the initial value control)

$$S(u) = \frac{1}{2}\left(V_b^{-1}(u-u_b), u-u_b\right)_X + \frac{1}{2}\left(V_o^{-1}(C\varphi-\varphi_{obs}), C\varphi-\varphi_{obs}\right)_{Y_{obs}}$$

Definition of the Hessian of the auxiliary control problem

$$\begin{cases} \partial\psi/\partial t - F'(\bar{\varphi})\psi = 0, \quad t \in (0,T) \\ \psi\big|_{t=0} = v \end{cases}$$

$$\begin{cases} -\partial\psi^*/\partial t - \left(F'(\bar{\varphi})\right)^*\psi^* = -C^* V_o^{-1} C\psi, \quad t \in (0,T) \\ \psi^*\big|_{t=T} = 0 \end{cases}$$

$\Rightarrow$

$$H(\bar{\varphi})v = V_b^{-1}v - \psi^*\big|_{t=0}$$

Main result: $V_{\delta u} \approx H^{-1}(\bar{\varphi})$

Question: *how far it works ?*

Numerical solution

- *The auxiliary problem is a linear optimal control problem.*
- *It can be solved by a BFGS algorithm.*
- *A subproduct of the BFGS algorithm is the Hessian of the optimization problem.*

Fully nonlinear ensemble method

1. Consider function $\overline{\varphi}$ as the exact solution to the problem

2. Start ensemble loop $l = 1, \ldots, L$

2.1 Generate using Monte-Carlo $\xi_{b,l}$, $\xi_{o,l}$

2.2 Compute $u_b = \overline{u} + \xi_{b,l}$, $\varphi_{obs} = C\overline{\varphi} + \xi_{o,l}$

2.3 Solve the original nonlinear DA problem with perturbed data and find u_l

2.4 Compute $\delta u_l = u_l - \overline{u}$

3. End ensemble loop.

4. Compute the statistics $\hat{V}_{\delta u} = \frac{1}{L}\sum_{l=1}^{L} \delta u_l \delta u_l^T$

The fully nonlinear ensemble method is used to compute benchmark estimates of the posterior covariance matrix, to be compared with the inverse Hessian (see figures below). The sample size can be reduced with the sampling error compensation procedure (presented in a forthcoming slide).
Otherwise, this method is very expensive and can not be used in its original form for large-scale applications.

Example 1: Initialization problem

Model (non-linear convection-diffusion):

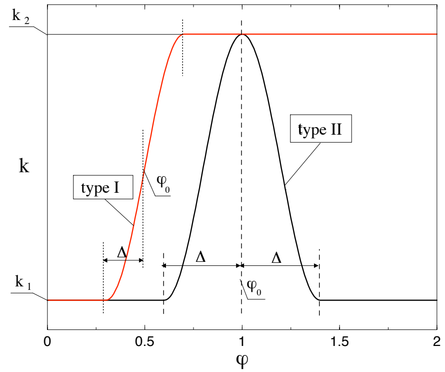
$$\frac{\partial \varphi}{\partial t} + \frac{\partial (w\varphi)}{\partial x} - \frac{\partial}{\partial x}\left(k(\varphi)\frac{\partial \varphi}{\partial x}\right) = Q(\varphi)$$

$$x \in (0,1), \quad t \in (0,T]$$

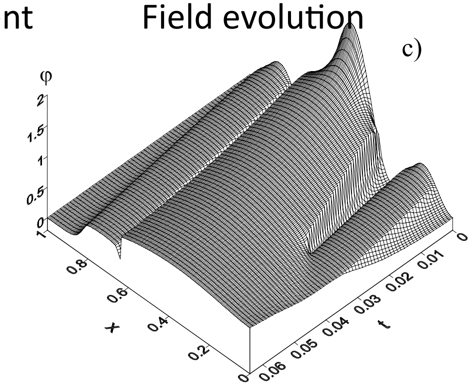
$$\varphi(x,0) = u$$

$$\frac{\partial \varphi(0,t)}{\partial x} = \frac{\partial \varphi(1,t)}{\partial x} = 0$$

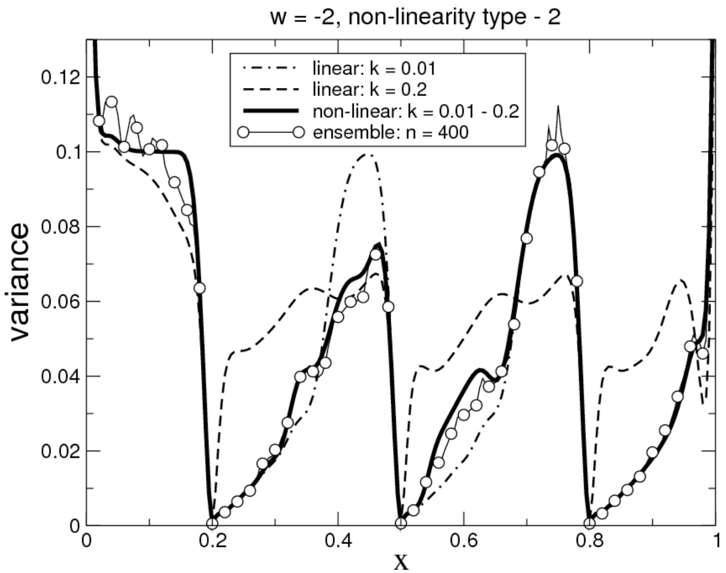
Nonlinear diffusion coefficient



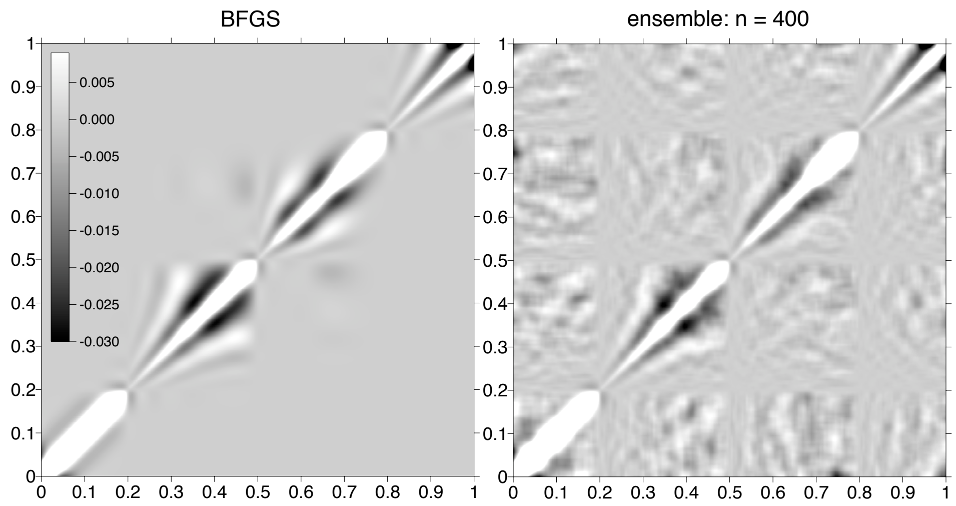
Field evolution



$diag\left(H^{-1}\right)$ and ensemble variance



H^{-1} and ensemble covariance



About the Tangent Linear Hypothesis

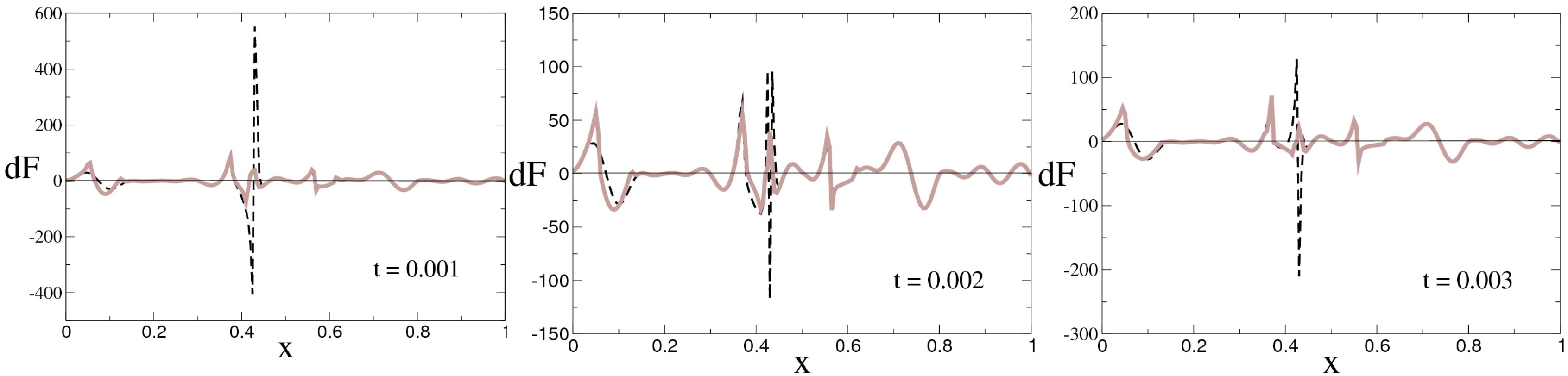
It is said that the inverse Hessian of the objective function is a good approximation of the posterior covariance if the 'tangent linear hypothesis' (TLH) is valid, i.e. the error dynamics is adequately represented by the tangent linear model.

However, this condition is overly restictive and, in many cases, the main result should be valid far beyond the validity of the TLH.

The explanation is that the TLH is a local condition, while the Hessian definition includes both forward and backward time integrations. Therefore, what matters is the reminder of the linearization error after integrations on a set of all possible implementations of random background and observation errors.

Below we show a degree of violation of the TLH relevant to the case presented in the previous slide. Despite that violation, a very good match between the inverse Hessian and the ensemble covariance can be observed.

Solid line: $dF^{(1)} = F(\varphi) - F(\overline{\varphi})$ Dashed line: $dF^{(2)} = F'(\overline{\varphi})(\varphi - \overline{\varphi})$ φ - optimal solution



Example 2: Boundary control problem

Model (non-linear convection-diffusion):

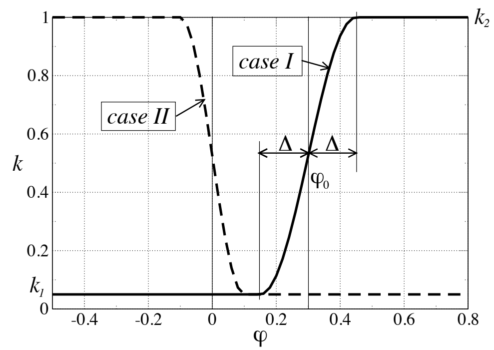
$$\frac{\partial \varphi}{\partial t} + \frac{\partial (w\varphi)}{\partial x} - \frac{\partial}{\partial x}\left(k(\varphi)\frac{\partial \varphi}{\partial x}\right) = Q(\varphi)$$

$$x \in (0,1), \quad t \in (0,T]$$

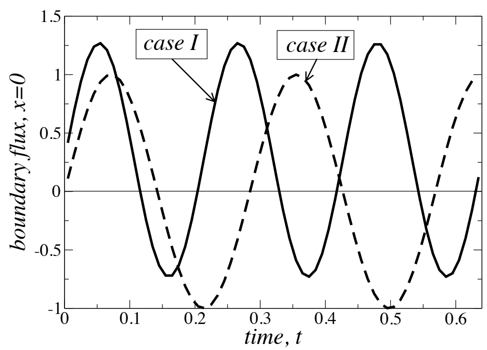
$$\varphi(x,0) = 0$$

$$\frac{\partial \varphi(0,t)}{\partial x} = u_1(t), \quad \frac{\partial \varphi(1,t)}{\partial x} = u_2(t)$$

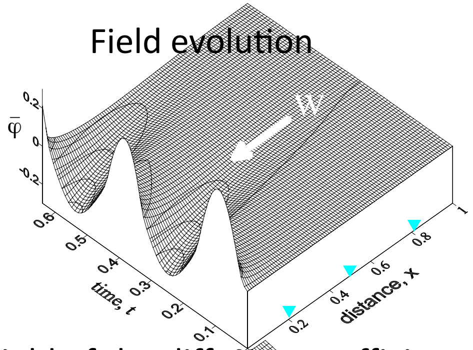
Nonlinear diffusion coefficient



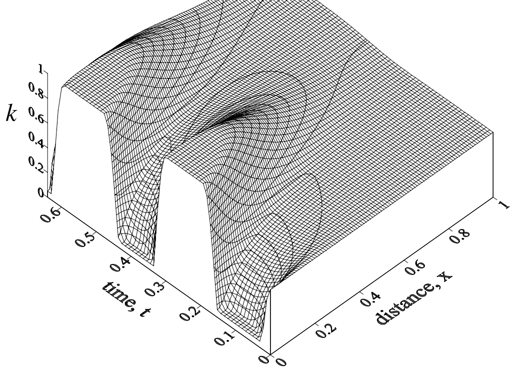
'True' boundary condition



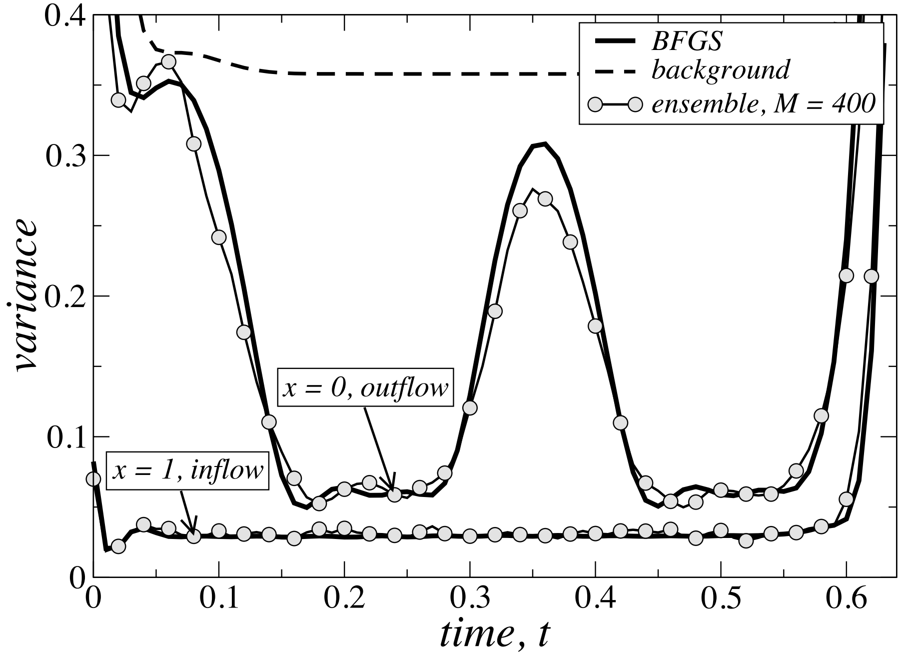
Field evolution



Field of the diffusion coefficient

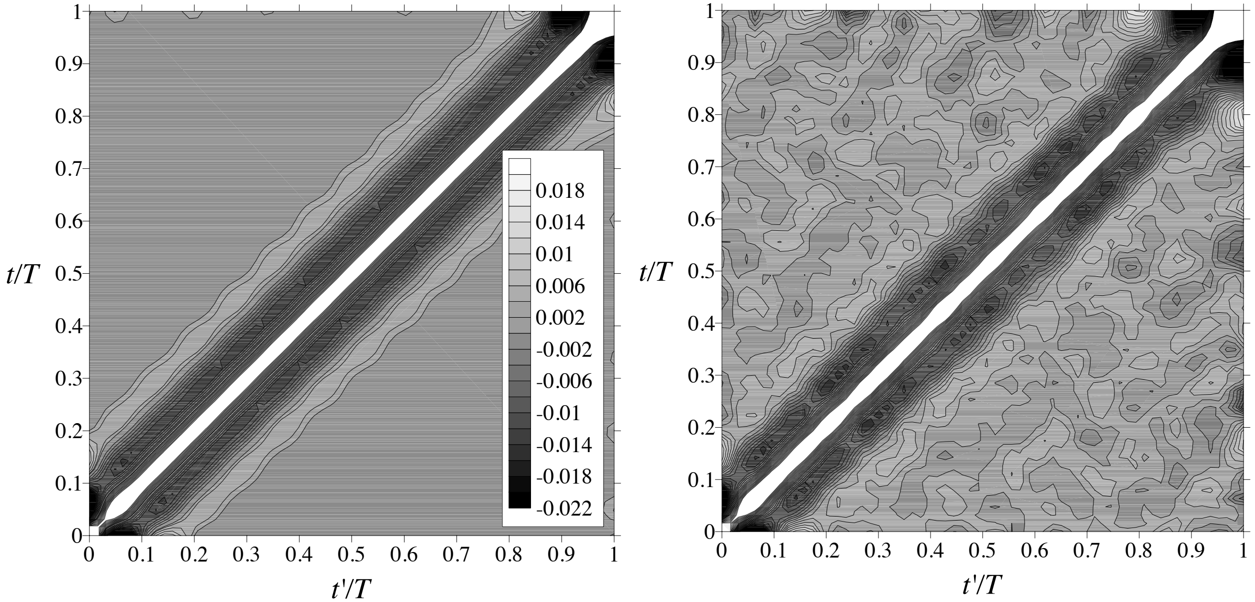


$diag(H^{-1})$ and ensemble variance

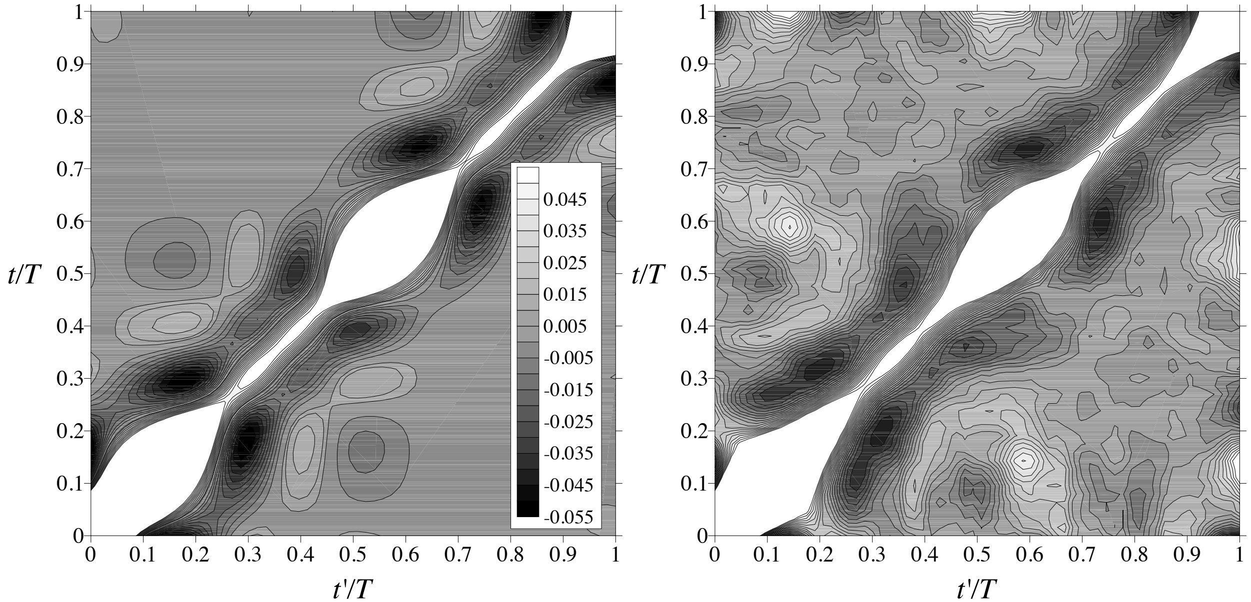


Example 2: Boundary control problem: H^{-1} and ensemble covariance

Inflow boundary
$x = 1$



Outflow boundary
$x = 0$



Example 3: Distributed coefficient estimation problem

Model (linear convection-diffusion):

$$\frac{\partial \varphi}{\partial t} + \frac{\partial (w\varphi)}{\partial x} - \frac{\partial}{\partial x}\left(k(x) \frac{\partial \varphi}{\partial x}\right) = Q(\varphi)$$

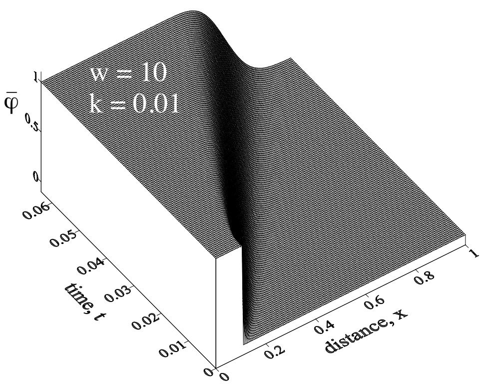
$$x \in (0,1), \quad t \in (0,T]$$

$$\varphi(x,0) = u$$

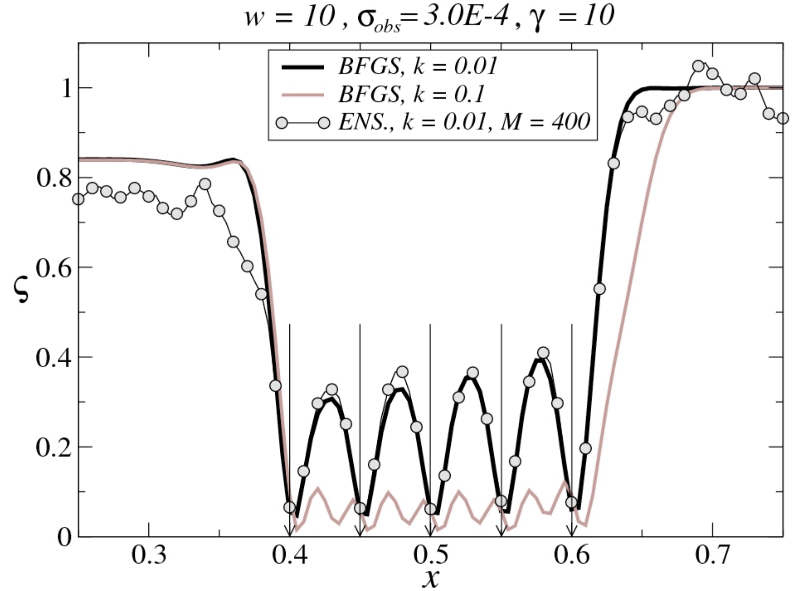
$$\frac{\partial \varphi(0,t)}{\partial x} = 0, \quad \frac{\partial \varphi(1,t)}{\partial x} = 0$$

Parameter estimation problem is nonlinear even for linear dynamics !

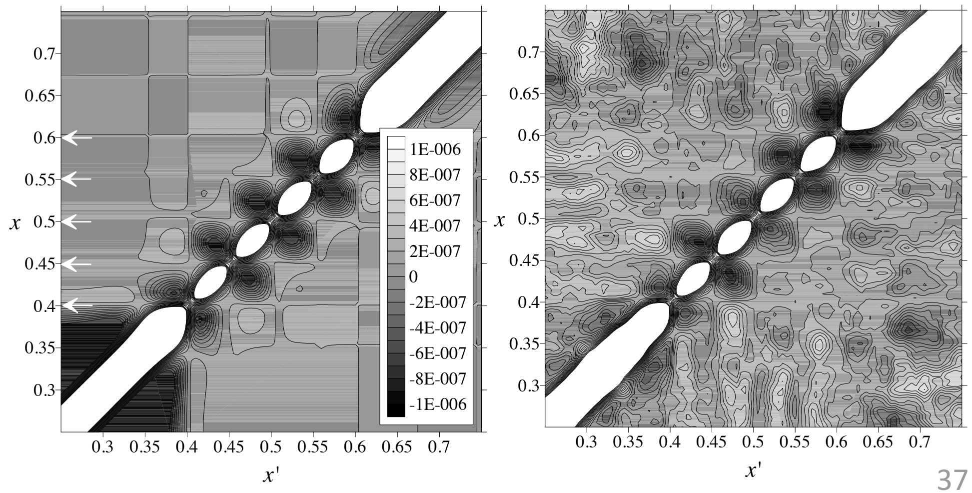
Field evolution



$diag(H^{-1})$ and ensemble variance



H^{-1} and ensemble covariance



If the inverse Hessian doesn't approximate covariance ?

$$V_{\delta u} \neq H^{-1}(\overline{\varphi})$$

In general case one may not expect the inverse Hessian to be a satisfactory approximation to the posterior covariance (see below).

Model: 1D Burgers with strongly nonlinear dissipation term

$$\frac{\partial \varphi}{\partial t} + \frac{1}{2}\frac{\partial(\varphi^2)}{\partial x} - \frac{\partial}{\partial x}\left(k\left(\varphi, \frac{\partial \varphi}{\partial x}\right)\frac{\partial \varphi}{\partial x}\right) = 0$$

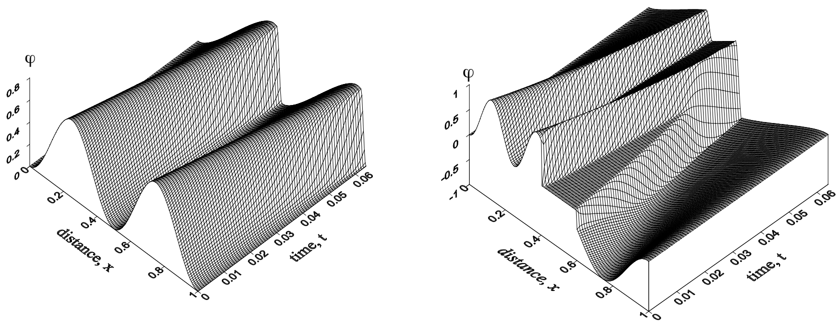
$$x \in (0,1), \quad t \in (0,T]$$

$$\varphi(x,0) = u \; ?$$

$$\partial\varphi(0,t)/\partial x = 0, \quad \partial\varphi(1,t)/\partial x = 0$$

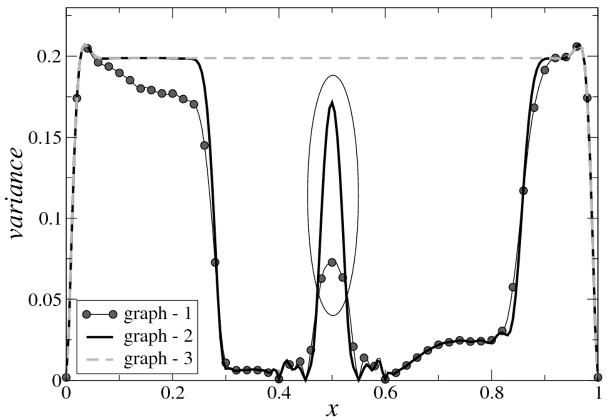
$$k = k_0 + k_1(\partial\varphi/\partial x)^2$$

Field evolution: case A and case B

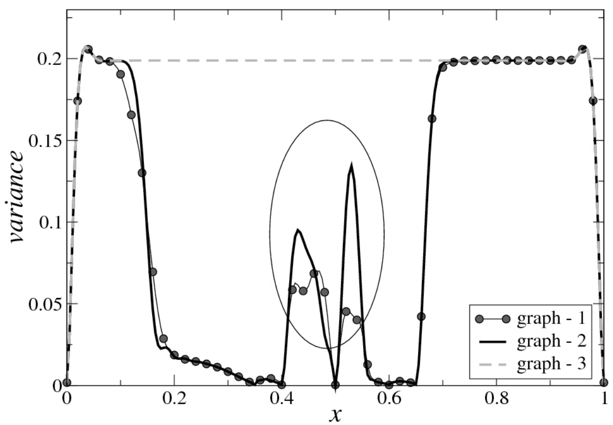


$diag(H^{-1})$ and ensemble variance for initialization problem

Case A: sensors at $x_k = (0.4, \quad 0.45, \quad 0.6, \quad 0,65)$



Case B: sensors at $x_k = (0.4, \quad 0.5, \quad 0.6)$



In Figures: inverse Hessian – solid line, ensemble estimate – dotted line

Compensation of the sampling error

If the problem is linearized around a 'true' state, the following error equation is valid:

$$\delta u_l^1 = H^{-1}(V_b^{-1}\xi_{b,l} + R\xi_{o,l})$$

Then, for any integer L (l is the sample element index), the sampling error is:

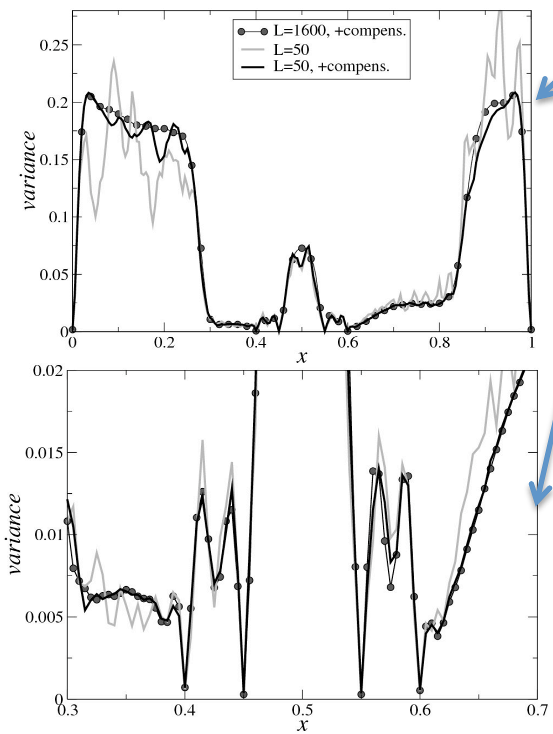
$$\delta V = \frac{1}{L}\sum_{l=1}^{L}\delta u_l^1 \delta u_l^{1,T} - H^{-1}$$

Assuming the same estimation of the sampling error is valid in the nonlinear case, one can compute the following approximation of the posterior covariance:

$$V_{\delta u} = H^{-1} + \frac{1}{L}\sum_{l=1}^{L}\left(\delta u_l \delta u_l^T - \delta u_l^1 \delta u_l^{1,T}\right)$$

where δu_l is the optimal solution, δu_l^1 is the result of the first Gauss-Newton iteration.

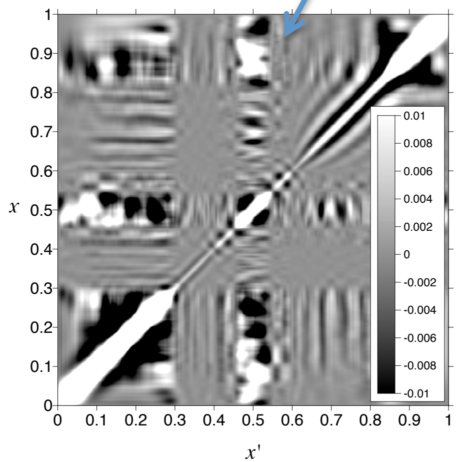
<u>This simple approach allows us to reduce significantly the sample size.</u>

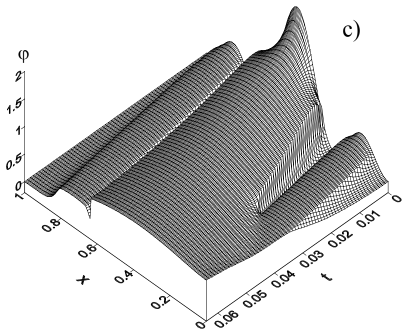


Posterior variance

Posterior variance, central part of the domain magnified

Posterior covariance by fully nonlinear ensemble method without and with compensation for $L=50$





‘Effective’ inverse Hessian approach

Instead of computing $V_{\delta u}$ via δu_l it is possible to compute estimation of the posterior covariance as the average of inverse Hessians defined on optimal solutions u_l :

$$V_{\delta u} = \frac{1}{L}\sum_{l=1}^{L} H^{-1}(u_l)$$

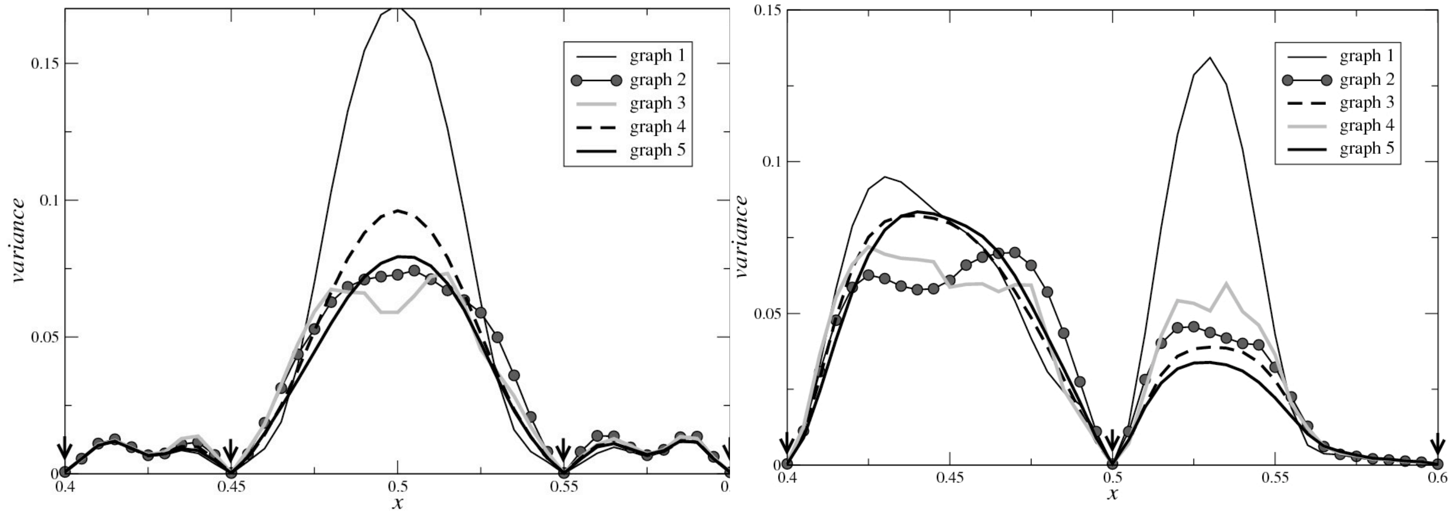
Moreover, instead of optimal solutions it is possible to use a sample of functions $\tilde{u}_l$, such that

$$E\left[(\tilde{u}_l - \bar{u})(\tilde{u}_l - \bar{u})^T\right] = V_0$$

where $V_0 \approx V_{\delta u}$. For example, we use $V_0 = H^{-1}(\bar{u})$.

This approach allows us to avoid the main difficulty associated to the fully nonlinear ensemble method: computing sample of optimal solutions!

Variance for initialization problem



In Figures: graph 1 - inverse Hessian on exact solution, graph 2 – estimate by the fully nonlinear ensemble method (L=1600), graph 3 (pale line) – estimate by the fully nonlinear ensemble method (L=50) with compensation, graph 4 (dashed line) – ‘effective’ inverse Hessian method using sample of optimal solutions (L=50), graph 5 (bold line) - effective’ inverse Hessian method using sample of randomly generated functions (L=50).

'Effective' inverse Hessian approach

Covariance for initialization problem

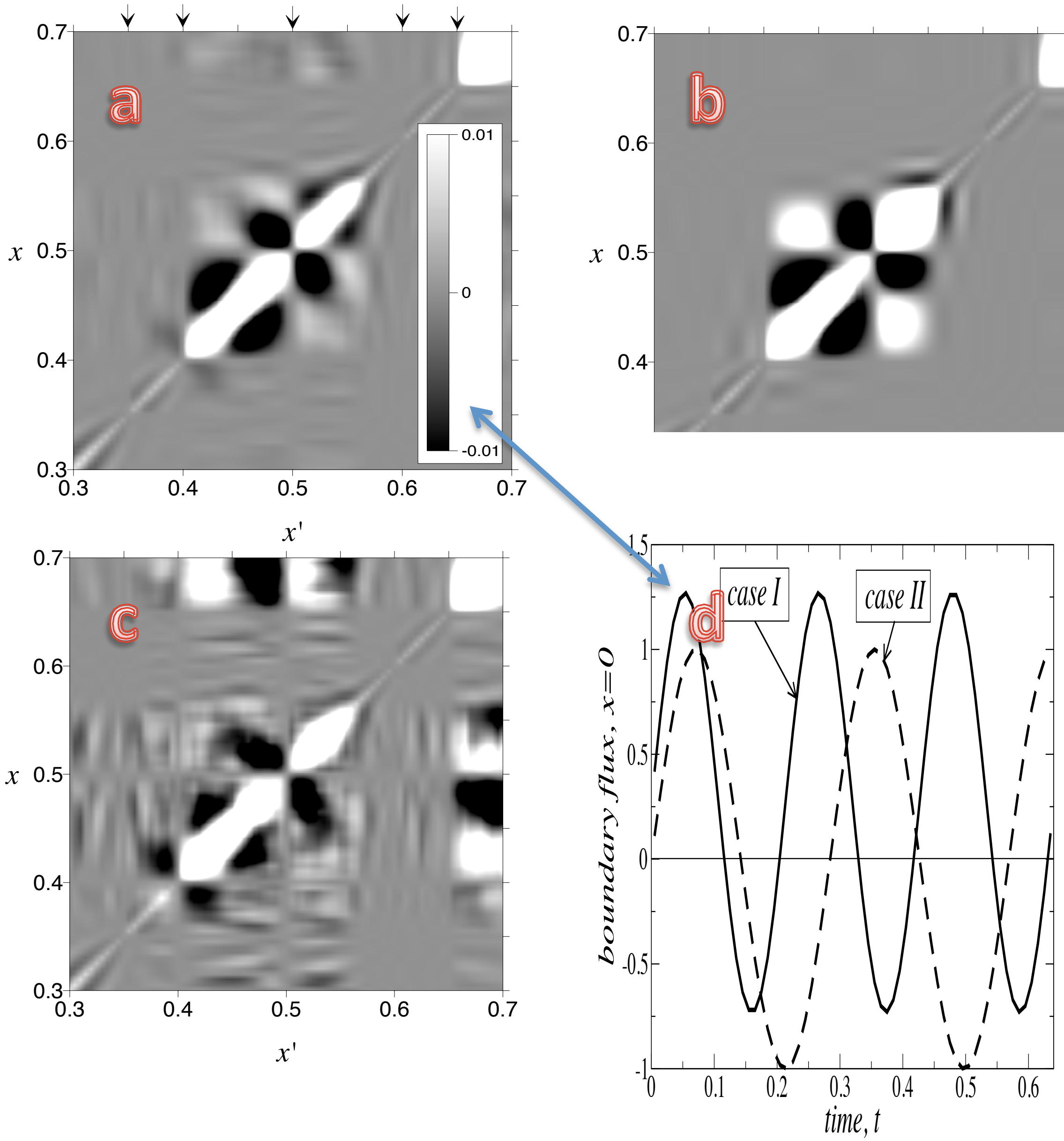


Fig. a - estimate by the fully nonlinear ensemble method ($L=1600$) with compensation;
Fig. b - estimate by the inverse Hessian on the 'exact' solution;
Fig. c - estimate by the fully nonlinear ensemble method ($L=50$) with compensation;
Fig. d – 'effective' inverse Hessian method using sample of randomly generated functions ($L=50$).

Conclusions

- In the linear case the posterior covariance is equal to the inverse Hessian.
- In the nonlinear case the posterior covariance can be well aproximated by the inverse Hessian if the *tangent linear hypothesis* is valid.
- the posterior covariance can still be well aproximated by the inverse Hessian if the *tangent linear hypothesis* is **not** valid (to some extent). It depends on the structure of the linearization error.
- If the nonlinear DA (estimation) problem exhibits a '*close-to-linear*' statistical behaviour, then the posterior covariance can be approximated by the '**effective**' inverse Hessian.
- Computation of the '**effective**' inverse Hessian might be feasible for large-scale applications, in the case when the target areas of the covariance matrix (for example its diagonal) are sought.
- The correction to the inverse Hessian which takes into account the nonlinearity can be evaluated by means of reduced-order modeling.
- If the nonlinear DA (estimation) problem does not exhibit '*close-to-linear*' statistical behaviour, the posterior covariance **cannot** characterize the probability distribution function.