A posteriori error covariances in variational data assimilation

Victor Shutyaev, Francois-Xavier Le Dimet, Igor Gejadze

Statement of the problem

$$\begin{cases} \dfrac{\partial \varphi}{\partial t} = F(\varphi, \lambda) + f, \quad t \in (0, T) \\ \varphi\big|_{t=0} = u, \end{cases} \tag{1.1}$$

where $\varphi = \varphi(t)$ is the unknown function belonging for any t to a Hilbert space X, $u \in X$, F is a nonlinear operator mapping $Y \times Y_p$ into Y with $Y = L_2(0, T; X)$, $\| \cdot \|_Y = (\cdot, \cdot)_Y^{1/2}$, Y_p is a Hilbert space (space of control parameters, or control space), $\lambda \in Y_p$, $f \in Y$.

Let us introduce the functional

$$S(\lambda) = \frac{1}{2}(V_1(\lambda - \lambda_b), \lambda - \lambda_b)_{Y_p} + \frac{1}{2}(V_2(C\varphi - \varphi_{obs}), C\varphi - \varphi_{obs})_{Y_{obs}}, \tag{1.2}$$

where $\lambda_b \in Y_p$ is a prior (background) function, $\varphi_{obs} \in Y_{obs}$ is a prescribed function (observational data), Y_{obs} is a Hilbert space (observation space), $C : Y \rightarrow Y_{obs}$ is a linear bounded operator, $V_1 : Y_p \rightarrow Y_p$ and $V_2 : Y_{obs} \rightarrow Y_{obs}$ are symmetric positive definite operators.

Data assimilation problem

Data assimilation problem: find $\lambda \in Y_p$ and $\varphi \in Y$ such that

$$\begin{cases} \dfrac{\partial \varphi}{\partial t} &= F(\varphi, \lambda) + f, \quad t \in (0, T) \\ \varphi\big|_{t=0} &= u, \\ S(\lambda) &= \inf\limits_{v \in Y_p} S(v). \end{cases} \tag{1.3}$$

The necessary optimality condition reduces the problem (1.3) to the following system :

$$\begin{cases} \dfrac{\partial \varphi}{\partial t} &= F(\varphi, \lambda) + f, \quad t \in (0, T), \\ \varphi\big|_{t=0} &= u, \end{cases} \tag{1.4}$$

$$\begin{cases} -\dfrac{\partial \varphi^*}{\partial t} - (F'_\varphi(\varphi, \lambda))^* \varphi^* &= -C^* V_2 (C\varphi - \varphi_{obs}), \quad t \in (0, T) \\ \varphi^*\big|_{t=T} &= 0, \end{cases} \tag{1.5}$$

$$V_1(\lambda - \lambda_b) - (F'_\lambda(\varphi, \lambda))^* \varphi^* = 0. \tag{1.6}$$

Errors

Suppose that $\lambda_b = \bar{\lambda} + \xi_1,\ \varphi_{obs} = C\bar{\varphi} + \xi_2$, where $\xi_1 \in Y_p,\ \xi_2 \in Y_{obs}$, and $\bar{\varphi}$ is the ("true") solution to the problem (1.1) with $\lambda = \bar{\lambda}$:

$$\left\{ \begin{array}{rcl} \dfrac{\partial \bar{\varphi}}{\partial t} & = & F(\bar{\varphi}, \bar{\lambda}) + f, \quad t \in (0, T), \\ \bar{\varphi}\big|_{t=0} & = & u. \end{array} \right. \tag{1.7}$$

The functions ξ_1, ξ_2 represent the errors of the input data λ_b and φ_{obs} (background and observation error, respectively). For V_1 and V_2 in (1.2), we consider

$$V_1 = V_{\xi_1}^{-1}, \quad V_2 = V_{\xi_2}^{-1},$$

where V_{ξ_i} is the covariance operator of the corresponding error ξ_i, i.e.

$$V_{\xi_1} \cdot = E[(\cdot, \xi_1)_{Y_p} \xi_1], \quad V_{\xi_2} \cdot = E[(\cdot, \xi_2)_{Y_{obs}} \xi_2],$$

where E is the expectation. If ξ is a vector, then the covariance matrix is defined by $V_\xi = E[\xi \xi^T]$.

Error analysis via Hessian

Let $\delta\varphi = \varphi - \bar{\varphi},\ \delta\lambda = \lambda - \bar{\lambda}$. Then, from (1.7) and the optimality system (1.4)–(1.6), we obtain

$$\begin{cases} \dfrac{\partial\delta\varphi}{\partial t} - F'_\varphi(\tilde{\varphi},\tilde{\lambda})\delta\varphi = F'_\lambda(\tilde{\varphi},\tilde{\lambda})\delta\lambda, & t \in (0,T), \\ \delta\varphi|_{t=0} = 0, \end{cases} \tag{2.1}$$

$$\begin{cases} -\dfrac{\partial\varphi^*}{\partial t} - (F'_\varphi(\varphi,\lambda))^*\varphi^* = -C^*V_2(C\delta\varphi - \xi_2), & t \in (0,T) \\ \varphi^*\big|_{t=T} = 0, \end{cases} \tag{2.2}$$

$$V_1(\delta\lambda - \xi_1) - (F'_\lambda(\varphi,\lambda))^*\varphi^* = 0, \tag{2.3}$$

where $\tilde{\varphi} = \bar{\varphi} + \tau(\varphi - \bar{\varphi}),\ \tilde{\lambda} = \bar{\lambda} + \tau(\lambda - \bar{\lambda}),\ \tau \in [0,1]$.

System for errors

The system (2.1)–(2.3) may be written in the form:

$$\begin{cases} \dfrac{\partial \delta\varphi}{\partial t} - F'_\varphi(\bar\varphi,\bar\lambda)\delta\varphi = F'_\lambda(\bar\varphi,\bar\lambda)\delta\lambda + \xi_3, \quad t\in(0,T), \\ \delta\varphi\big|_{t=0} = 0, \end{cases} \tag{2.4}$$

$$\begin{cases} -\dfrac{\partial \varphi^*}{\partial t} - (F'_\varphi(\bar\varphi,\bar\lambda))^*\varphi^* = -C^*V_2(C\delta\varphi - \xi_2) + \xi_4, \quad t\in(0,T) \\ \varphi^*\big|_{t=T} = 0, \end{cases} \tag{2.5}$$

$$V_1(\delta\lambda - \xi_1) - (F'_\lambda(\bar\varphi,\bar\lambda))^*\varphi^* = \xi_5, \tag{2.6}$$

where

$$\xi_3 = [F'_\varphi(\tilde\varphi,\tilde\lambda) - F'_\varphi(\bar\varphi,\bar\lambda)]\delta\varphi + [F'_\lambda(\tilde\varphi,\tilde\lambda) - F'_\lambda(\bar\varphi,\bar\lambda)]\delta\lambda,$$

$$\xi_4 = [(F'_\varphi(\varphi,\lambda))^* - (F'_\varphi(\bar\varphi,\bar\lambda))^*]\varphi^*, \quad \xi_5 = [(F'_\lambda(\varphi,\lambda))^* - (F'_\lambda(\bar\varphi,\bar\lambda))^*]\varphi^*.$$

Hessian

$$\begin{cases} \dfrac{\partial \delta\varphi}{\partial t} - F'_\varphi(\bar\varphi,\bar\lambda)\delta\varphi & = & F'_\lambda(\bar\varphi,\bar\lambda)\delta\lambda,\ t\in(0,T), \\ \delta\varphi|_{t=0} & = & 0, \\ S_1(\delta\lambda) & = & \inf\limits_v S_1(v), \end{cases} \tag{2.7}$$

$$S_1(\delta\lambda) = \frac{1}{2}(V_1(\delta\lambda-\xi_1),\delta\lambda-\xi_1)_{Y_p} + \frac{1}{2}(V_2(C\delta\varphi-\xi_2),C\delta\varphi-\xi_2)_{Y_{obs}}. \tag{2.8}$$

Consider the Hessian $H : Y_p \longrightarrow Y_p$ of the functional (2.8); it is defined by the successive solutions of the following problems:

$$\begin{cases} \dfrac{\partial \psi}{\partial t} - F'_\varphi(\bar\varphi,\bar\lambda)\psi & = & F'_\lambda(\bar\varphi,\bar\lambda)v,\ t\in(0,T), \\ \psi|_{t=0} & = & 0, \end{cases} \tag{2.9}$$

$$\begin{cases} -\dfrac{\partial \psi^*}{\partial t} - (F'_\varphi(\bar\varphi,\bar\lambda))^*\psi^* & = & -C^*V_2C\psi,\quad t\in(0,T) \\ \psi^*\big|_{t=T} & = & 0, \end{cases} \tag{2.10}$$

$$Hv = V_1v - (F'_\lambda(\bar\varphi,\bar\lambda))^*\psi^*. \tag{2.11}$$

Error equation

Below we introduce four auxiliary operators R_1, R_2, R_3, R_4. Let $R_1 = V_1$. For example, the operator $R_2 : Y_{obs} \longrightarrow Y_p$ is defined on the functions $g \in Y_{obs}$ according to the formula

$$R_2 g = (F'_\lambda(\bar{\varphi}, \bar{\lambda}))^* \theta^*, \tag{2.12}$$

where θ^* is the solution to the adjoint problem

$$\left\{ \begin{array}{rcl} -\dfrac{\partial \theta^*}{\partial t} - (F'_\varphi(\bar{\varphi}, \bar{\lambda}))^* \theta^* & = & C^* V_2 g, \quad t \in (0, T) \\ \theta^* \big|_{t=T} & = & 0. \end{array} \right. \tag{2.13}$$

From (2.9)–(2.11) we conclude that the system (2.4)–(2.6) is equivalent to the single equation for $\delta\lambda$:

$$H\delta\lambda = R_1\xi_1 + R_2\xi_2 + R_3\xi_3 + R_4\xi_4 + \xi_5. \tag{2.14}$$

This is the exact equation for $\delta\lambda$. Under the hypothesis that H is invertible, we get

$$\delta\lambda = T_1\xi_1 + T_2\xi_2 + T_3\xi_3 + T_4\xi_4 + T_5\xi_5, \quad T_i = H^{-1} R_i. \tag{2.15}$$

Approximation

Since $\tilde{\varphi} = \bar{\varphi} + \tau\delta\varphi,\ \varphi = \bar{\varphi} + \delta\varphi$, we assume that $T_3\xi_3 \approx 0,\ \ T_4\xi_4 \approx 0,\ \ T_5\xi_5 \approx 0$. Then

$$\delta\lambda = T_1\xi_1 + T_2\xi_2, \tag{2.16}$$

and (2.4)–(2.6) reduces to the auxiliary DA problem:

$$\begin{cases} \dfrac{\partial\delta\varphi}{\partial t} - F'_\varphi(\bar{\varphi},\bar{\lambda})\delta\varphi &= F'_\lambda(\bar{\varphi},\bar{\lambda})\delta\lambda, \quad t \in (0,T), \\ \delta\varphi|_{t=0} &= 0, \end{cases} \tag{2.17}$$

$$\begin{cases} -\dfrac{\partial\varphi^*}{\partial t} - (F'_\varphi(\bar{\varphi},\bar{\lambda}))^*\varphi^* &= -C^*V_2(C\delta\varphi - \xi_2), \quad t \in (0,T) \\ \varphi^*\big|_{t=T} &= 0, \end{cases} \tag{2.18}$$

$$V_1(\delta\lambda - \xi_1) - (F'_\lambda(\bar{\varphi},\bar{\lambda}))^*\varphi^* = 0. \tag{2.19}$$

The problem (2.17)–(2.19) is a linear data assimilation problem; with the fixed $\bar{\varphi}$ it is the necessary optimality condition to the following minimization problem: find $\delta\lambda$ and $\delta\varphi$ such that (2.7) is satisfied.

Sensitivity coeffi cients

Since $\delta\lambda = T_1\xi_1 + T_2\xi_2$ for $T_i = H^{-1}R_i$, the *sensitivity coefficients* are defined by

$$r_i = \sqrt{\|T_i^* T_i\|}.$$

For $V_1 = \alpha E,\ V_2 = E$ we have for r_1:

$$r_1 = \sqrt{\|T_1^* T_1\|} = \frac{\alpha}{\mu_{\min}}. \tag{2.20}$$

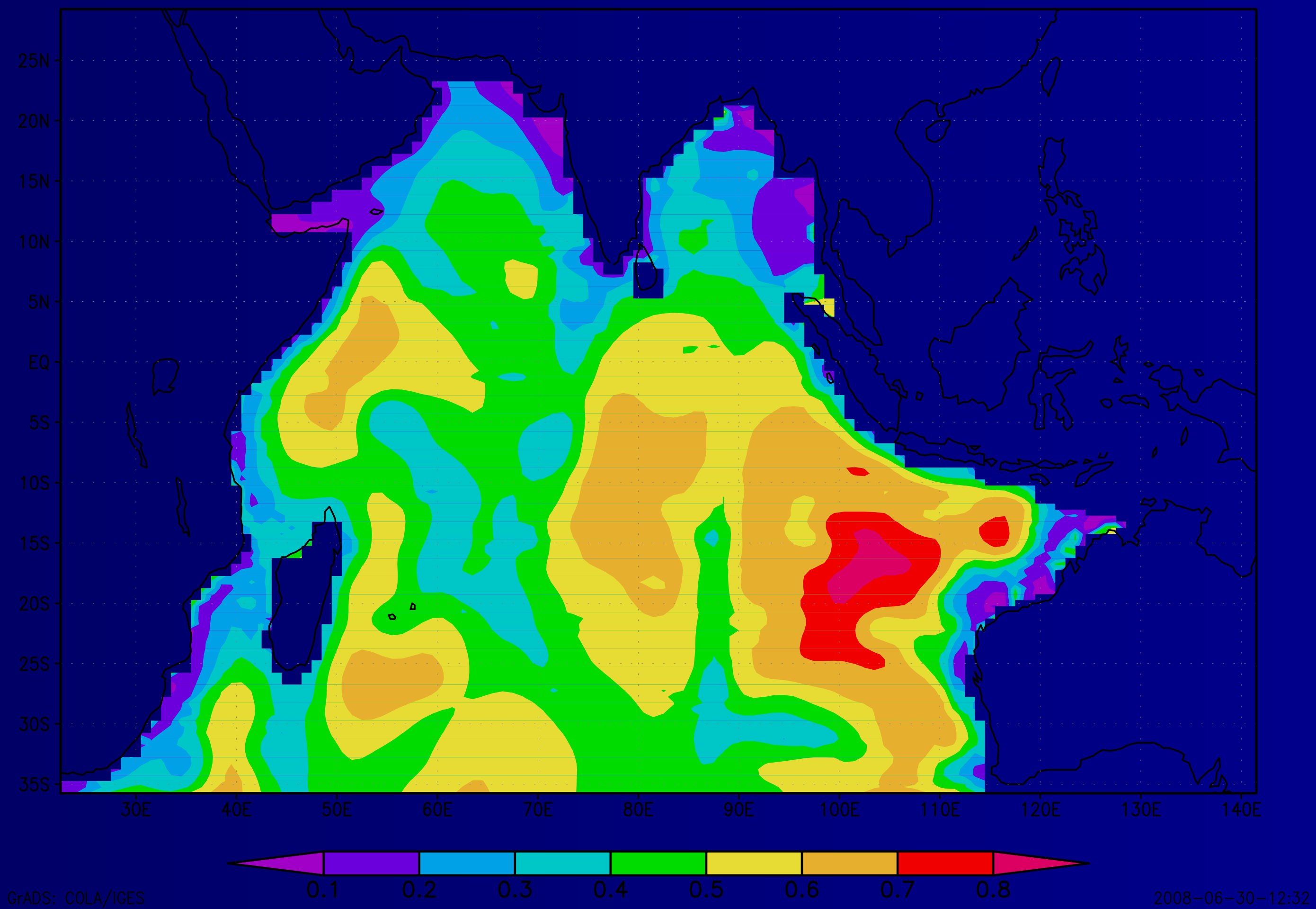
The singular values σ_k^2 and the corresponding orthonormal (right) singular vectors $w_k \in Y_{obs}$ of the operator T_2 are defined by the formulas (Le Dimet, Shutyaev, 2005):

$$\sigma_k^2 = \frac{\mu_k - \alpha}{\mu_k^2}, \quad w_k = \frac{1}{\sqrt{\mu_k - \alpha}} C^h \varphi_k, \tag{2.21}$$

where μ_k are the eigenvalues of the Hessian H, and φ_k are the fundamental control functions, and

$$r_2 = \max_k \frac{\sqrt{\mu_k - \alpha}}{\mu_k}. \tag{2.22}$$

Example (Shutyaev, Parmuzin): Indian Ocean model



Covariance operators

Consider the error equation (2.14). Since H is invertible, we get

$$\delta\lambda = T_1\xi_1 + T_2\xi_2, \tag{3.1}$$

where $T_i{=}H^{-1}R_i,\ T_1 : Y_p{\rightarrow}Y_p,\ T_2 : Y_{obs}{\rightarrow}Y_p$. We suppose that the errors ξ_1, ξ_2 are normally distributed, unbiased, and mutually uncorrelated. By V_{ξ_i} we denote the covariance operator of the corresponding error $\xi_i,\ i = 1, 2$, i.e. $V_{\xi_1}\cdot = E[(\cdot, \xi_1)_{Y_p}\xi_1]$, $V_{\xi_2}\cdot = E[(\cdot, \xi_2)_{Y_{obs}}\xi_2]$, where E is the expectation. By $V_{\delta\lambda}$ we denote the covariance operator of the optimal solution error: $V_{\delta\lambda}\cdot = E[(\cdot, \delta\lambda)_{Y_p}\delta\lambda]$. From (3.1) we get

$$V_{\delta\lambda} = T_1 V_{\xi_1} T_1^* + T_2 V_{\xi_2} T_2^*. \tag{3.2}$$

To find the covariance operator $V_{\delta u}$, we need to construct the operators $T_i V_{\xi_i} T_i^*,\ i = 1, 2$. Consider the operator $T_1 V_{\xi_1} T_1^*$. Since $T_1 = H^{-1}R_1 = H^{-1}V_1 = T_1^*$, we have $T_1 V_{\xi_1} T_1^* = H^{-1}V_1 V_{\xi_1} V_1 H^{-1}$. Moreover, if $V_1 = V_{\xi_1}^{-1}$, then

$$T_1 V_{\xi_1} T_1^* = H^{-1}V_1 H^{-1} = H^{-1}V_{\xi_1}^{-1}H^{-1}. \tag{3.3}$$

Operator R_2^*

Consider the operator $T_2 V_{\xi_2} T_2^*$. Since $T_2 = H^{-1} R_2$, then

$$T_2 V_{\xi_2} T_2^* = H^{-1} R_2 V_{\xi_2} R_2^* H^{-1}.$$

To determine R_2^*, consider the inner product $(R_2 g, p)_{Y_p}$, $g \in Y_{obs}, p \in Y_p$. From (2.12)–(2.13),

$$(R_2 g, p)_{Y_p} = ((F'_\lambda(\bar{\varphi}, \bar{\lambda}))^* \theta^*, p)_{Y_p} = (C^* V_2 g, \phi)_Y = (g, R_2^* p)_{Y_{obs}},$$

where $R_2^* p = V_2 C \phi$, and ϕ is the solution to the problem

$$\left\{ \begin{array}{rcl} \dfrac{\partial \phi}{\partial t} - F'_\varphi(\bar{\varphi}, \bar{\lambda})\phi & = & F'_\lambda(\bar{\varphi}, \bar{\lambda})p, \; t \in (0, T), \\ \phi|_{t=0} & = & 0. \end{array} \right. \quad (3.4)$$

Operator $T_2 V_{\xi_2} T_2^*$

The operator $T_2 V_{\xi_2} T_2^*$ is defined by successive solutions of the following problems (for a given $v \in Y_p$):

$$Hp = v, \tag{3.5}$$

$$\begin{cases} \dfrac{\partial \phi}{\partial t} - F'_\varphi(\bar\varphi, \bar\lambda)\phi = F'_\lambda(\bar\varphi, \bar\lambda)p, \ t \in (0,T), \\ \phi\big|_{t=0} = 0, \end{cases} \tag{3.6}$$

$$\begin{cases} -\dfrac{\partial \theta^*}{\partial t} - (F'_\varphi(\bar\varphi, \bar\lambda))^*\theta^* = C^* V_2 V_{\xi_2} V_2 C\phi, \ \ t \in (0,T) \\ \theta^*\big|_{t=T} = 0, \end{cases} \tag{3.7}$$

$$Hw = (F'_\lambda(\bar\varphi, \bar\lambda))^*\theta^*, \tag{3.8}$$

then

$$T_2 V_{\xi_2} T_2^* v = w. \tag{3.9}$$

If $V_2 = V_{\xi_2}^{-1}$, then $C^* V_2 V_{\xi_2} V_2 C = C^* V_2 C$ and $(F'_\lambda(\bar\varphi, \bar\lambda))^*\theta^* = Hp - R_1 p.$

Optimal solution error covariance

We get

$$R_2 V_{\xi_2} R_2^* = H - V_1$$

and

$$T_2 V_{\xi_2} T_2^* = H^{-1} R_2 V_{\xi_2} R_2^* H^{-1} = H^{-1}(H - V_1)H^{-1}. \qquad (3.10)$$

From (3.3), (3.10) it follows the result for $V_{\delta\lambda}$:

$$V_{\delta\lambda} = T_1 V_{\xi_1} T_1^* + T_2 V_{\xi_2} T_2^* = H^{-1} V_1 H^{-1} + H^{-1}(H - V_1)H^{-1}.$$

Therefore

$$V_{\delta\lambda} = H^{-1} H H^{-1} = H^{-1}. \qquad (3.11)$$

The last formula gives the optimal-solution-error covariance operator through the Hessian H.

Gejadze, I., Le Dimet, F.-X., Shutyaev, V.P. On analysis error covariances in variational data assimilation. SIAM J. Sci. Comput. (2008), v.30, no.4, 1847-1874.

Shutyaev, V., Le Dimet, F.-X., Gejadze, I. On optimal solution error covariances in variational data assimilation. Russ. J. Numer. Anal. Math. Modelling (2008), v.23, no.2, 197-206.

Numerical algorithm to compute covariances

Consider the covariance operator $V = V_{\delta\lambda}$ defined by (3.11):

$$V = H^{-1}. \tag{4.1}$$

To find the inverse Hessian H^{-1}, the quasi-Newton BFGS method may be used, because it generates an approximation of H^{-1} directly in the course of a minimization process.
Since the Hessian H of the functional S_1 does not depend on functions ξ_1, ξ_2 entering (2.8), we suggest using as follows:

$$\xi_1 = \tilde{\lambda}, \quad \xi_2 = C\delta\tilde{\varphi}, \tag{4.2}$$

where $\delta\tilde{\varphi}$ satisfies the problem (2.17).
In this case, the solution of (2.7) is $\delta\lambda = \tilde{\lambda}$, and $S_1(\tilde{\lambda}) = 0$. The initial guess to start the iterations is $\delta\lambda^0 = 0$.

BFGS method

Applied for solving the auxiliary DA problem (2.7)-(2.8), the BFGS method has the form:

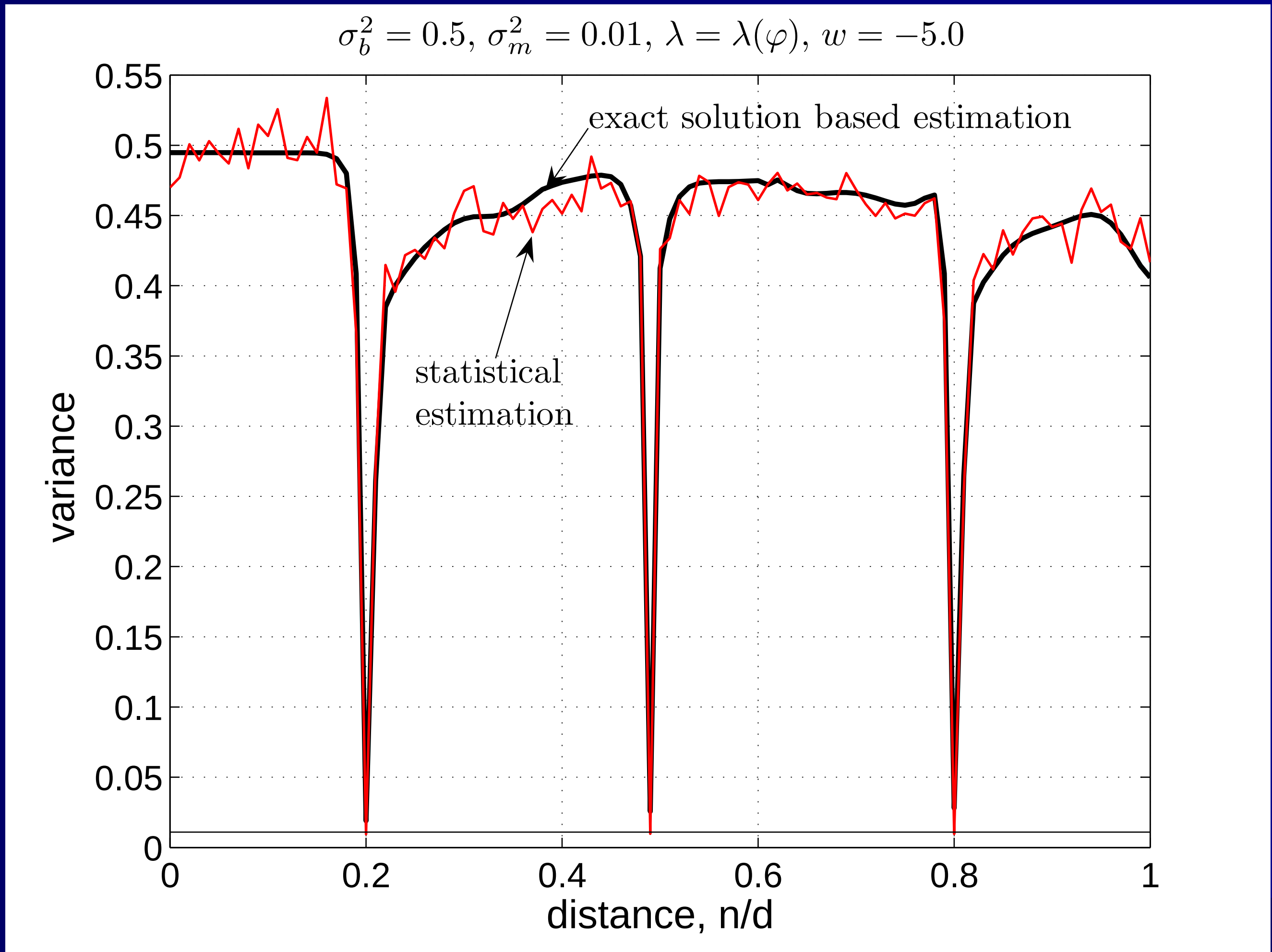
$$d^k = H_k^{-1} S_1'(\delta\lambda^k), \tag{4.3}$$

$$\delta\lambda^{k+1} = \delta\lambda^k - \alpha^k d^k, \tag{4.4}$$

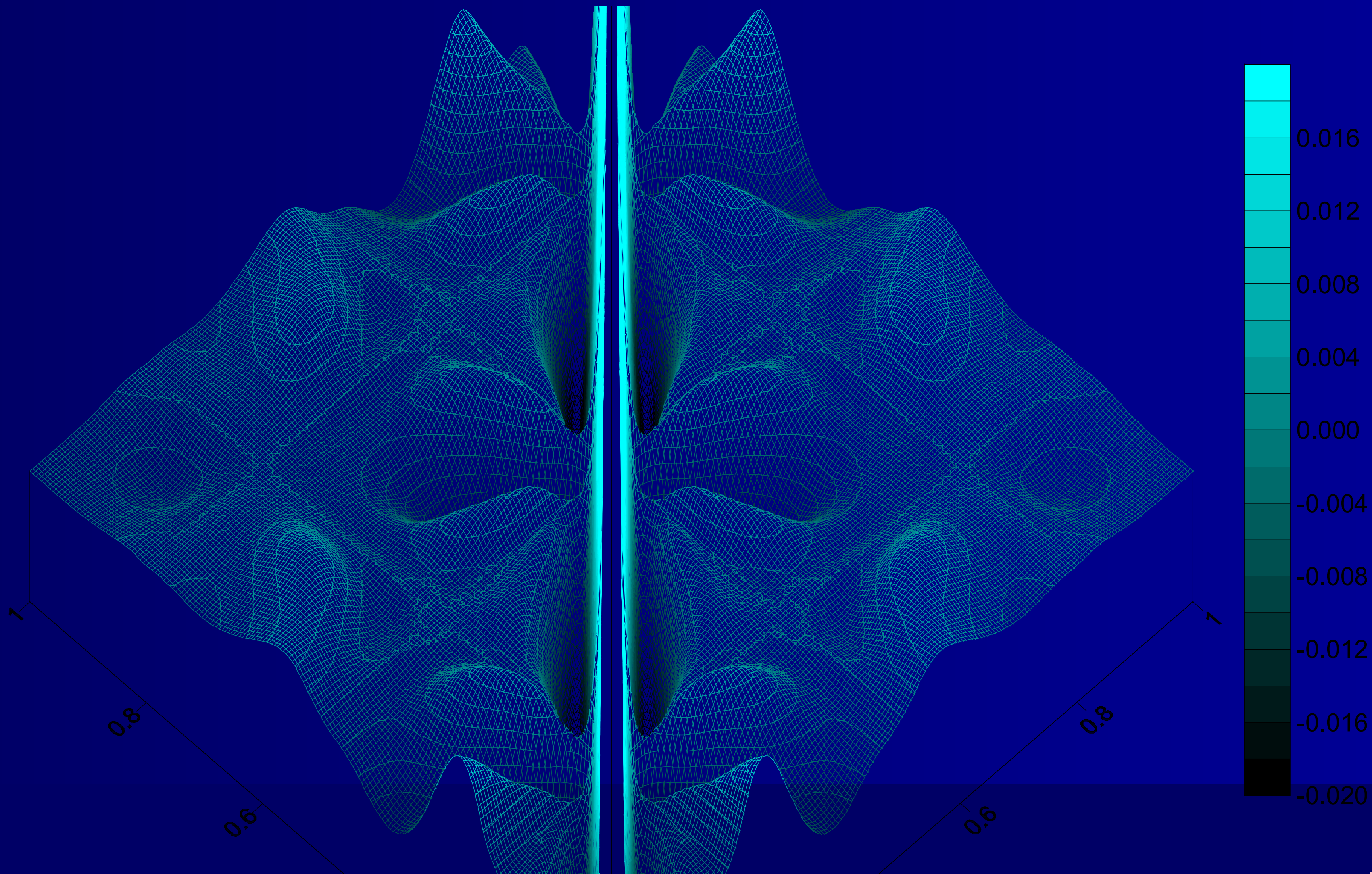
$$H_{k+1}^{-1} = \left(I - \frac{sy^T}{y^T s}\right) H_k^{-1} \left(I - \frac{ys^T}{y^T s}\right) + \frac{ss^T}{y^T s}, \tag{4.5}$$

where $s = \delta\lambda^{k+1} - \delta\lambda^k,\ y = S_1'(\delta\lambda^{k+1}) - S_1'(\delta\lambda^k)$, H_k^{-1} is the approximation to H^{-1} on the k-th iteration, $S_1'(\delta\lambda^k)$ is the value of the gradient of S_1 in $\delta\lambda$ at the point $\delta\lambda^k$, α^k are iterative parameters, I is the identity operator.

Statistical variance σ_a^2 and variance via H^{-1}



Covariance


0.016
0.012
0.008
0.004
0.000
-0.004
-0.008
-0.012
-0.016
-0.020
1
0.8
0.6
1
0.8
0.6

Conclusions

The error of the optimal initial-value function in variational data assimilation for a nonlinear evolution model may be expressed by an equation through the errors of the input data without the tangent linear hypothesis. The approximation of the error equation allows to derive the analysis error covariance operator which turnes to be the inverse Hessian of the auxiliary (linearized) error assimilation problem. This Hessian does not coincide in general with the Hessian of the original cost functional. With the use of the quasi-Newton BFGS method, a numerical algorithm is developed to compute the analysis error covariance operator as the inverse Hessian. The algorithm is based on a special choice of input functions in the auxiliary data assimilation problem and the analytical step search for the minimization along the direction of descent. This leads to obtain the covariance operator which perfectly matches the one obtained by the statistical (ensemble) method.

Shutyaev, V., Le Dimet, F.-X., Gejadze, I. On optimal solution error covariances in variational data assimilation. Russ. J. Numer. Anal. Math. Modelling (2008), v.23, no.2, 197-206.

References

[1] Greenstadt J. Variations on variable-metric methods. Math. Comp. 24 (1970) 1-22 (Appendix by J. Bard).

[2] Le Dimet F.-X., Navon I.M., Daescu D.N. Second-order information in data assimilation. *Monthly Weather Review*, 2002, v.130, no.3, pp.629–648.

[3] Le Dimet F.-X., Shutyaev V. On deterministic error analysis in variational data assimilation. *Nonlinear Processes in Geophysics*, 2005, 14, p. 1-10.

[4] Le Dimet F.X., Talagrand O. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus*, 1986, v.38A, pp.97-110.

[5] Liu D.C., Nocedal J. On the limited memory BFGS method for large scale minimization, Math. Prog. 45 (1989) 503-528.

[6] Rabier F., Courtier P. Four-dimensional assimilation in the presence of baroclinic instability. *Quart. J. Roy. Meteorol. Soc.*, 1992, v.118, pp.649-672.

[7] Thacker W.C. The role of the Hessian matrix in fitting models to measurements. *J. Geophys. Res.*, 1989, v.94, no.C5, pp.6177-6196.

Posterior covariance in variational DA for nonlinear evolution models

F.-X. Le Dimet[1] , I. Gejadze[2], V.P. Shutyaev[3]

[1] *MOISE project, LJK, University of Grenoble, France*

[2] *Department of Civil Engineering, University of Strathclyde, Glasgow, UK*

[3] *Institute of Numerical Mathematics, Russian Academy of Science, Moscow, Russia*

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

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

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

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

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

Important question: *how far it works ?*

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 $\xi_{b,l}, \quad \xi_{o,l}$

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

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 page 10.
Otherwise, this method is very expensive and can not be used in its original form for large-scale applications.

The BFGS method is used to build the inverse Hessian.

Example 1: Initialization problem

Model (non-linear convection-diffusion):

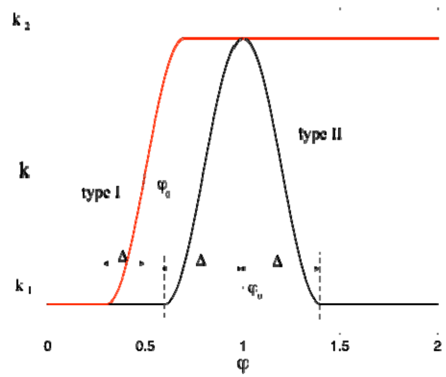
$$\frac{\partial \varphi}{\partial t} + \frac{\partial (w\varphi)}{\partial x} - \frac{\partial}{\partial x}\left(k(\varphi)\frac{\partial \varphi}{\partial x}\right) = Q(\varphi)$$

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

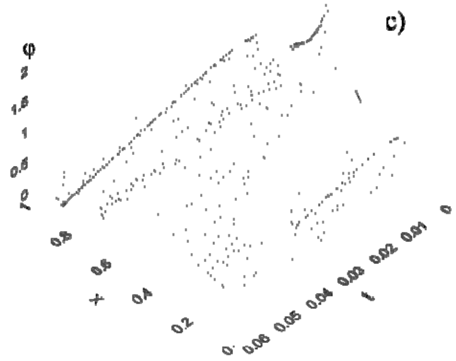
$$\varphi(x,0) = u$$

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

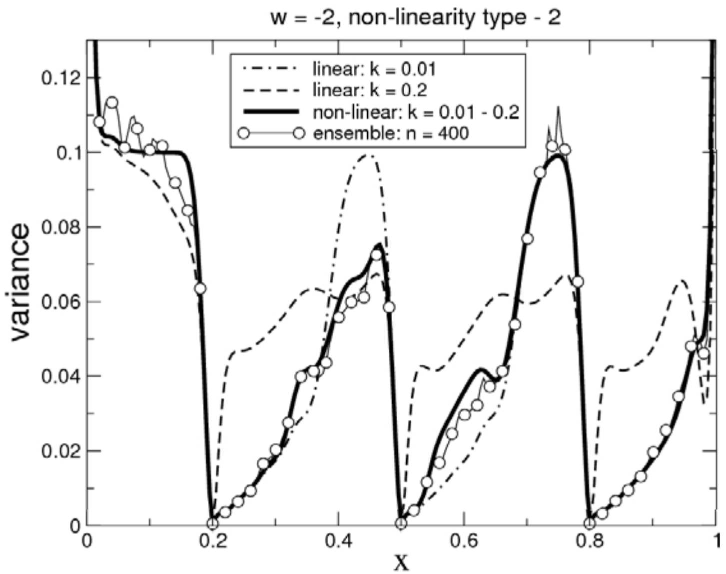
Nonlinear diffusion coefficient



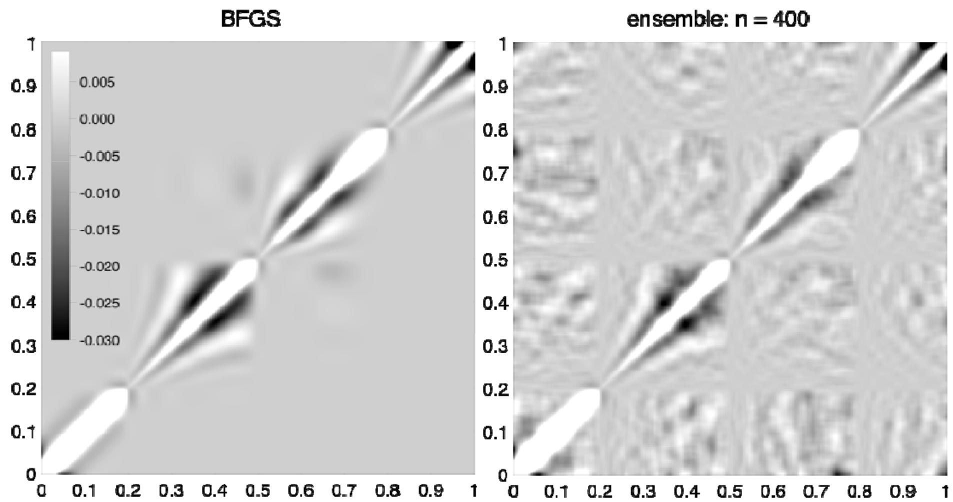
Field evolution



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



H^{-1} and ensemble covariance



About the Tangent Linear Hypothesis

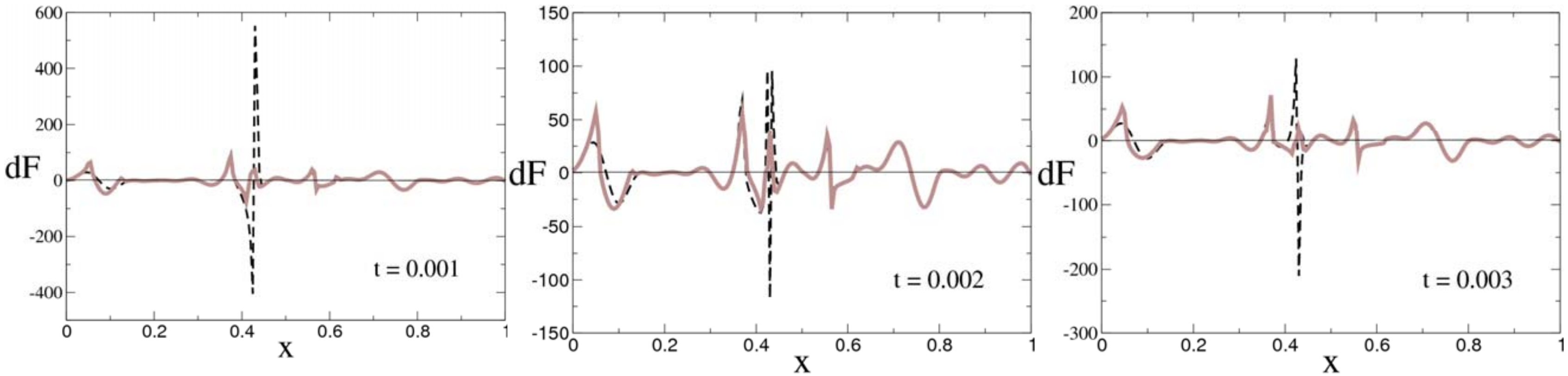
It is said that the inverse Hessian of the objective function is a good approximation of the posterior covariance if the 'tangent linear hypothesis' (TLH) is valid, i.e. the error dynamics is adequately represented by the tangent linear model.

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):

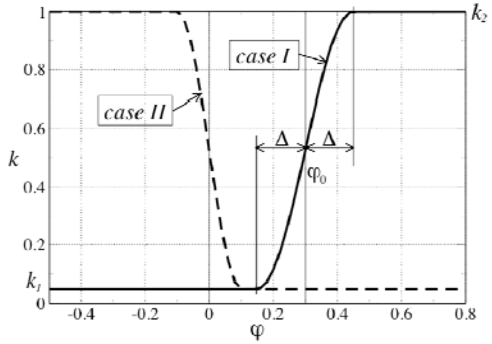
$$\frac{\partial \varphi}{\partial t} + \frac{\partial (w\varphi)}{\partial x} - \frac{\partial}{\partial x}\left(k(\varphi)\frac{\partial \varphi}{\partial x}\right) = Q(\varphi)$$

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

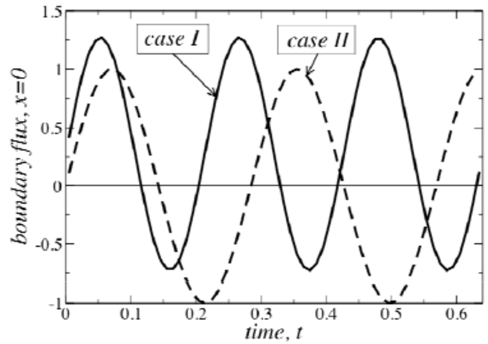
$$\varphi(x,0) = 0$$

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

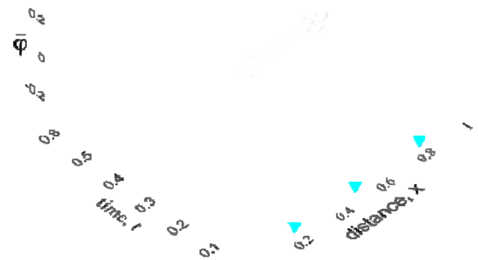
Nonlinear diffusion coefficient



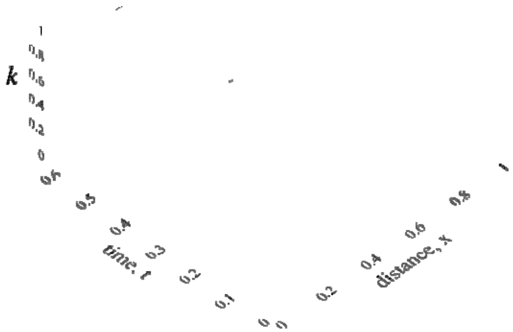
'True' boundary condition



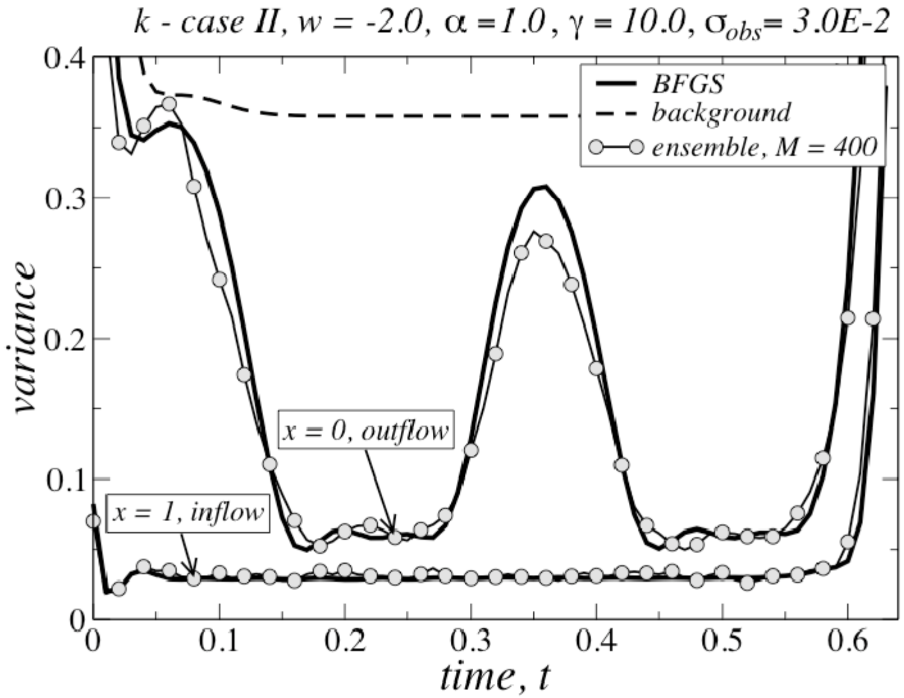
Field evolution



Field of the diffusion coefficient



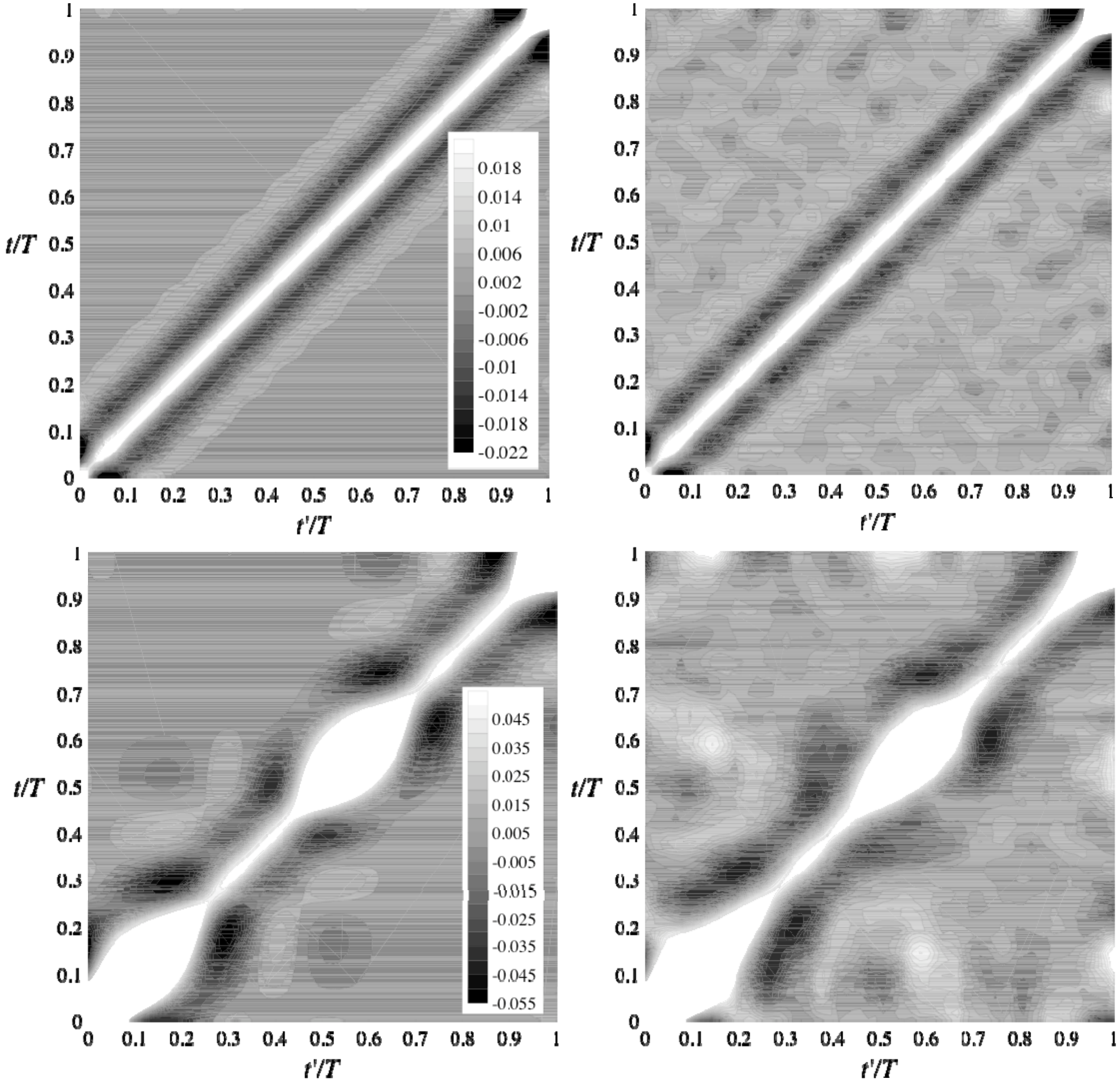
$diag(H^{-1})$ and ensemble variance



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):

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

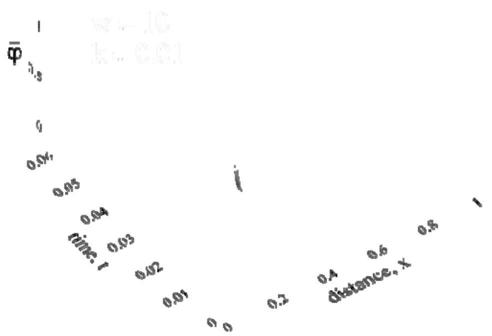
$$x \in (0,1), \quad t \in (0,T]$$

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

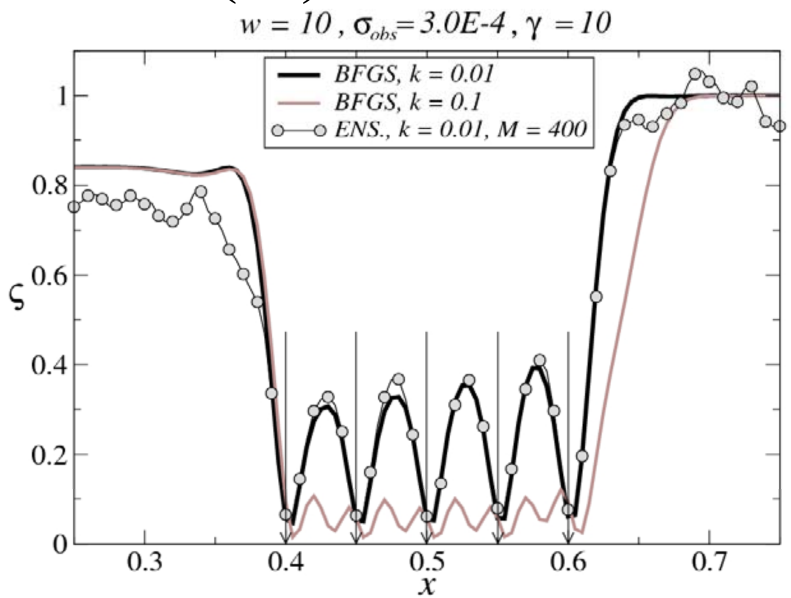
$$\frac{\partial \varphi(0,t)}{\partial x} = 0, \quad \frac{\partial \varphi(1,t)}{\partial x} = 0$$

Parameter estimation problem is nonlinear even for linear dynamics !

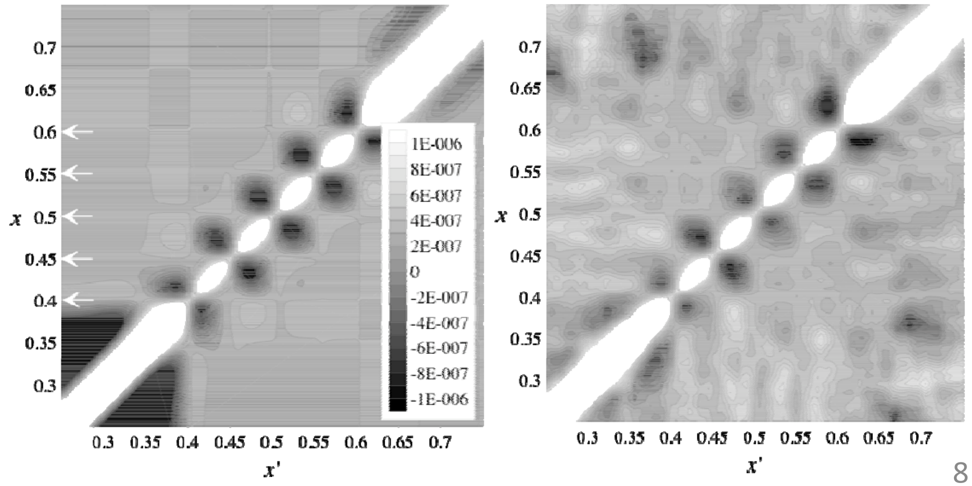
Field evolution



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



H^{-1} and ensemble covariance



When the main result is not valid

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

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

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

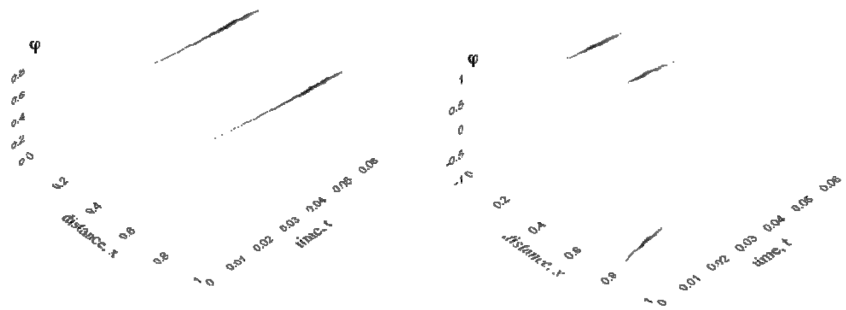
$$x \in (0,1), \quad t \in (0,T]$$

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

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

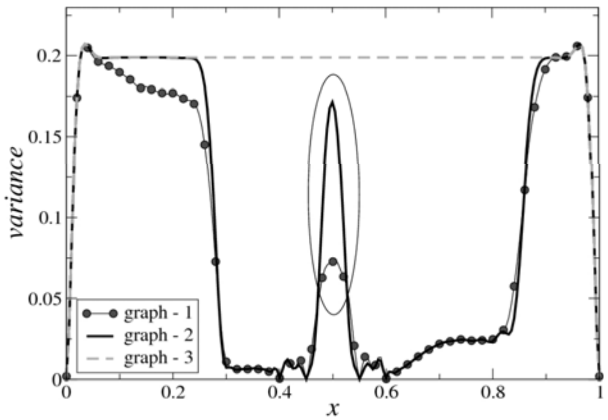
$$k = k_0 + k_1(\partial\varphi/\partial x)^2$$

Field evolution: case A and case B

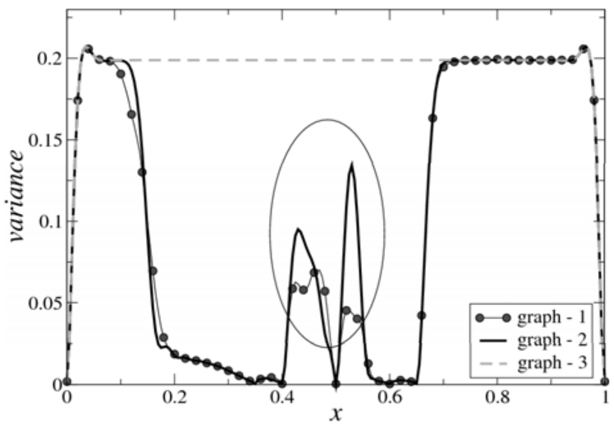


$diag(H^{-1})$ and ensemble variance for initialization problem

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, background variance – dashed line

Compensation of the sampling error

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

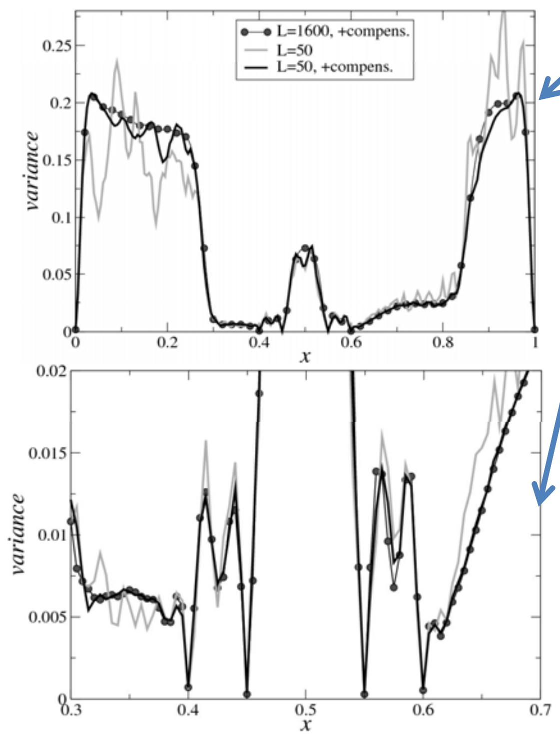
$$\delta V = \frac{1}{L}\sum_{l=1}^{L}\delta u_l^1 \delta u_l^{1,T} - H^{-1}$$

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.

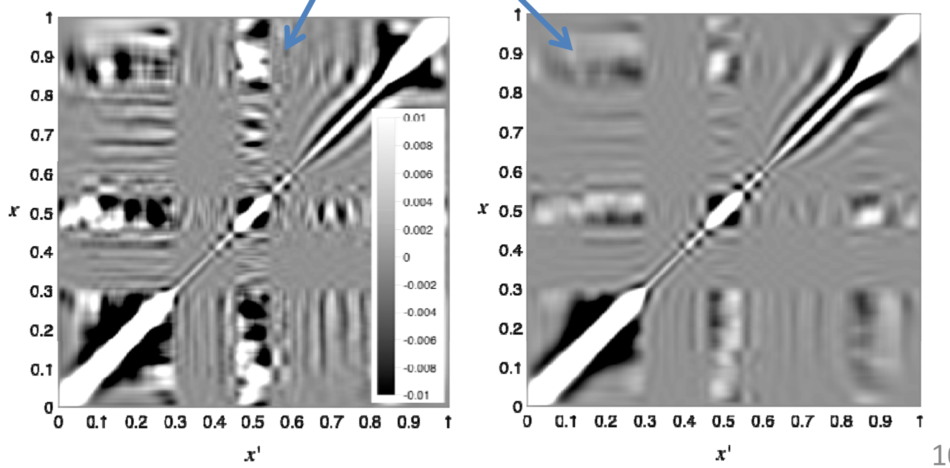
<u>This simple approach allows us to reduce significantly the sample size.</u>



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 :

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

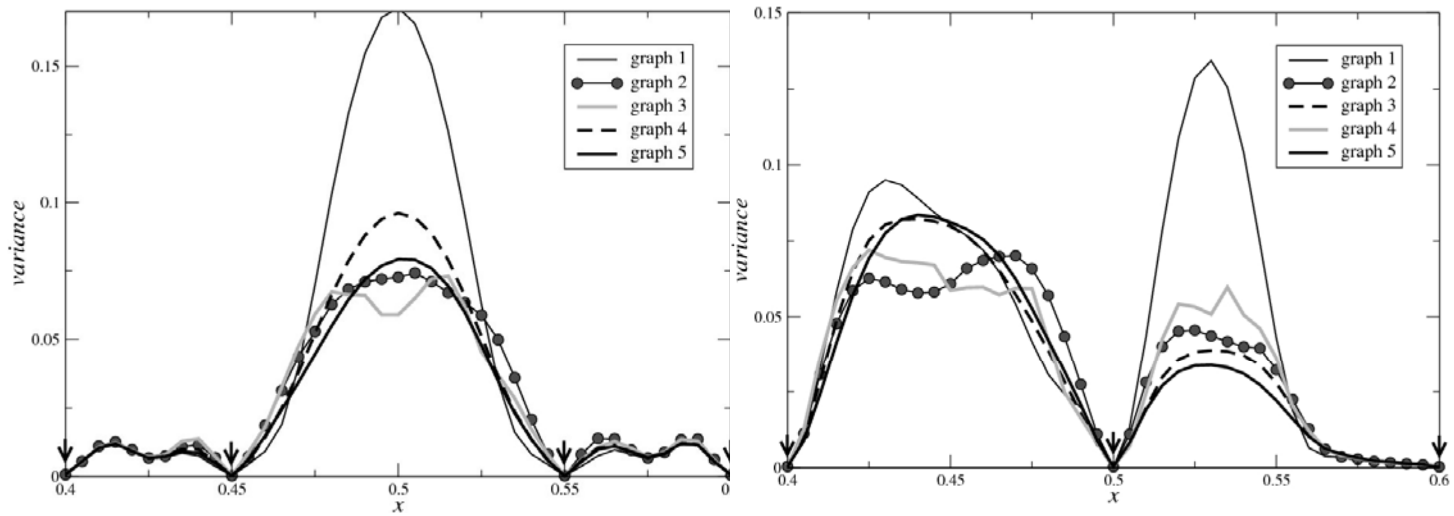
Moreover, instead of optimal solutions it is possible to use a sample of functions $\tilde{u}_l$, such that

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

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

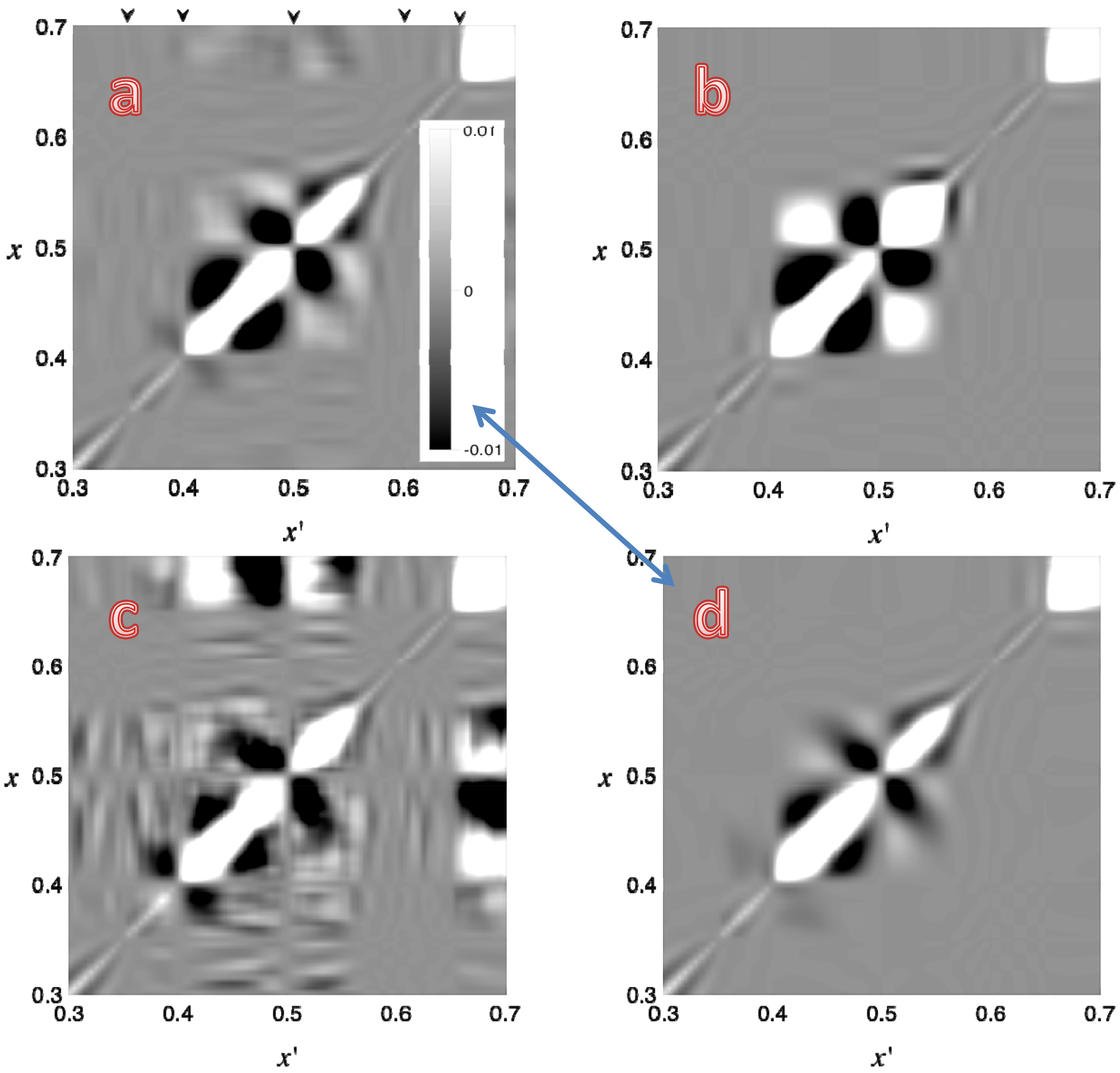


Fig. a - estimate by the fully nonlinear ensemble method ($L=1600$) with compensation;
Fig. b - estimate by the inverse Hessian on the ‘exact’ solution;
Fig. c - estimate by the fully nonlinear ensemble method ($L=50$) with compensation;
Fig. d – ‘effective’ inverse Hessian method using sample of randomly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

Nonlinearity criteria

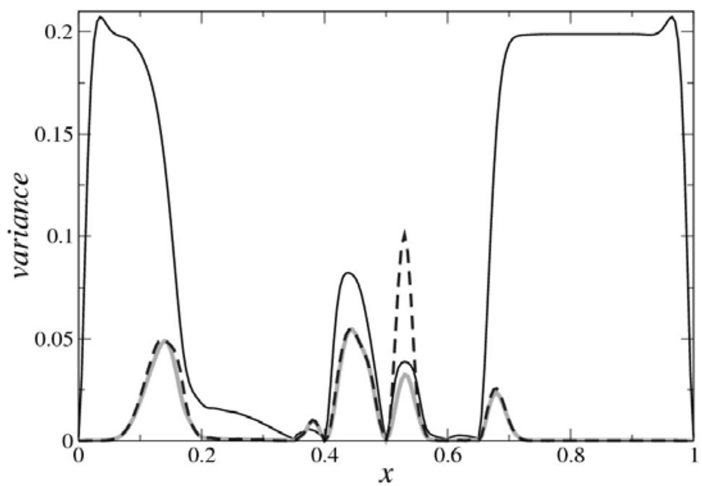
In the process of evaluating the 'effective' inverse Hessian one can also compute the following second order statistics:

$$D_{\delta u}^2 = \frac{1}{L}\sum_{l=1}^{L}\Big(diag\{H^{-1}(u_l)\} - diag\{V_{\delta u}\}\Big)^2$$

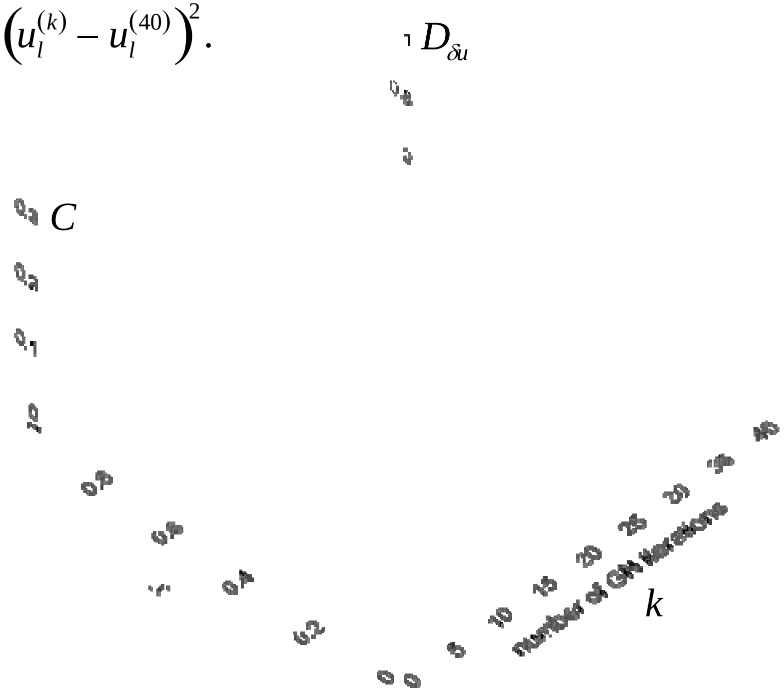
$$\tilde{D}_{\delta u}^2 = \frac{1}{L}\sum_{l=1}^{L}\Big(diag\{H^{-1}(u_l)\} - diag\{H^{-1}(\overline{u})\}\Big)^2$$

These statistics characterize the stability of the variance in respect to the position of the solution on the solution locus surface and are useful as distributed ***nonlinearity criteria***. For example, one can notice a similarity between $D_{\delta u}$ and an averaged convergence rate of the Gauss-Newton method evaluated as follows:

$$C^2 = \frac{1}{L}\sum_{l=1}^{L}\Big(u_l^{(k)} - u_l^{(40)}\Big)^2.$$



In fugure above: solid bold line - $V_{\delta u}$ by the 'effective' inverse Hessian approach, dashed line - $\tilde{D}_{\delta u}$, faint solid line - $D_{\delta u}$.



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.

References

- Gejadze I., Le Dimet F.-X., Shutyaev V.P. On analysis error covariances in variational data assimilation. *SIAM J. Sci. Comp.* , 2008, v.30/4, pp.1847-1874.
- Gejadze I., Le Dimet F.-X., Shutyaev V.P. On optimal solution error covariances in variational data assimilation problems. *J. Comp. Physics* , 2009, 30pp, in review.
- Gejadze I., Le Dimet F.-X., Shutyaev V.P. On computing a posteriori covariances in variational data assimilation for nonlinear evolution models. *Tellus A*, 2009, in preparation.