Habilitation a Diriger des Recherches

by Igor Gejadze

Advanced uncertainty analysis and optimal design in variational estimation (data assimilation)

Grenoble, 27 February, 2018

Outline

1. Education and background
2. General introduction
3. Uncertainty quantification
4. Design problems
5. Parts not included
6. Current and future work

Background:

1979-1985: Moscow State Technical University (Bauman)
MSc in Mech. Eng.:
Flight dynamics, guidance and control

1985-1991: State Research Institute of Aviation Systems (design bureau)
Engineer – Senior Engineer (project leader)
Autonomous guidance system for TU160

1994-1998: Moscow Aviation Institute
PhD in Mech. Eng.:
Algorithms for real time monitoring of thermal loads in flight vehicle structures

Academic Carrier

1999-2000: Polytechnic school, University of Nantes, PDRF
Inverse problems in polymer extrusion

2000-2003: The Weizmann Institute of Science, Israel, PDRF
Multigrid methods for solving control problems (with Achi Brandt)

2003-2005: University of Strathclyde, Civil Engineering. Dept., Glasgow, PDRF
Development the adjoint for ICOM (Imperial College Ocean Model)

2005-2006: University Joseph Fourier, IMAG, Grenoble, PDRF
Coupling and data assimilation in hydraulic modelling

2006-2013: University of Strathclyde, Civil Engineering. Dept., Glasgow, Senior RF, NERC advanced RF, open end contract
Uncertainty quantification and design in variational DA

2013- : IRSTEA, Montpellier, Senior RF
to be presented later

PhD Supervision

PhD student: Hind Oubanas

Thesis title: "Détermination des débits des fleuves et grands canaux d'irrigation à partir de données acquises au sol ou par télédétection par des méthodes d'assimilation de données".

Defended: 01/2018

Current position: engineer at IRSTEA-Montpellier

Very successful

PhD Student: Hossam Mohamed El-Hanafy

Thesis title: "Sensitivity and uncertainty analysis for flood wave propagation in river channels using the adjoint method".

Defended: 10/2007

Current position: lecturer at Egyptian Military College (MTC), Cairo

Quite successful

PhD student: Kirsty L. Brown

Thesis title: "Efficient multigrid computation of the posterior covariance matrix in large-scale variational data assimilation".

Unsuccessful

Outline

1. Education and background
2. **General introduction**
3. Uncertainty quantification
4. Design problems
5. Parts not included
6. Current and future work

Advanced uncertainty analysis and optimal design in variational estimation (data assimilation)

Uncertainty analysis:

- intrinsic feature of most time-sequential estimation methods (filtering, ensemble and particle methods), however a difficult issue in variational estimation!

Optimal design:

- design of optimal observation arrays or trajectories (in geophysical data assimilation)
- design of test signals (in parameter and structural estimation problems)
- optimal shape design (in aero/hydrodynamics).

A pivotal object in these applications – **Hessian**

Major difficulty – **high dimensions**, **CPU time**

Advanced uncertainty analysis and optimal design in variational estimation (data assimilation)

1. **Hessian** – second derivative of the DA cost-function or design function

2. **Truncated Hessian** – linearised Hessian, obtained by ignoring second-order terms

3. The later is also known as **Controlability Gramian**, **Fisher Information Matrix**

4. **Hessians** are key mathematical objects in optimal and feedback controls, estimation, information and optimal experiment design theories

General formalism - 1

$U \in \mathcal{U}$ - **full set** of model inputs, $\mathcal{U}$ - input space

$X \in \mathcal{X}$ - state variables , $\mathcal{X}$ - state space

$\underline{X = \mathcal{M}(U)}$ - model or input-to-state mapping $\mathcal{M} : \mathcal{U} \to \mathcal{X}$

$\bar{U}$ - 'true'/exact input , $\bar{X} = \mathcal{M}(\bar{U})$ - 'true' model prediction

Instead:

$U^* = \bar{U} + \varepsilon$, where U^* - background/prior , ε - background error

$Y \in \mathcal{Y}$ - observables , $\mathcal{Y}$ - observation space, C - observation operator

$$\underline{Y = C(X) = C(\mathcal{M}(U)) := G(U)}$$

- 'input-to-observations' mapping $G : \mathcal{U} \to \mathcal{Y}$

$\bar{Y} = G(\bar{U})$ - 'true'/exact observations

Instead:

$Y^* = \bar{Y} + \xi = G(\bar{U}) + \xi$, ξ - observation error

$\hat{U} = U|Y^*$ - estimate of U, posterior

General formalism - 2

Variational DA cost-function:

$$J(U) = \frac{1}{2}\|R^{-1/2}(G(U) - Y^*)\|^2_{\mathcal{Y}} + \frac{1}{2}\|B^{-1/2}(U - U^*)\|^2_{\mathcal{U}} \to \inf_U$$

$R = E[\xi\xi^T]$ - observation error covariance

$B = E[\varepsilon\varepsilon^T]$ - background error covariance

$J'(U) = G'^*_U(U)R^{-1}(G(U) - Y^*) + B^{-1}(U - U^*)$ - gradient of $J(U)$

$J'_U(\hat{U}) = 0$ - optimality condition

$J''(U) \cdot v := \mathcal{H}(U) \cdot v =$ - Hessian (second derivative of $J(U)$)

$((G'_U(U))^* R^{-1} G'_U(U) + B^{-1}) \cdot v + [(G'(U))^*]'_U \cdot v \; R^{-1}(G(U) - Y^*_0)$

$H(U) \cdot v = ((G'_U(U))^* R^{-1} G'_U(U) + B^{-1}) \cdot v$ - truncated Hessian

$G'(\cdot)$ and $G'^*(\cdot)$ are the TL and adjoint operators

$\delta U = \hat{U} - \bar{U}$ - estimation error

$E[\delta U] = 0$ - estimation error bias

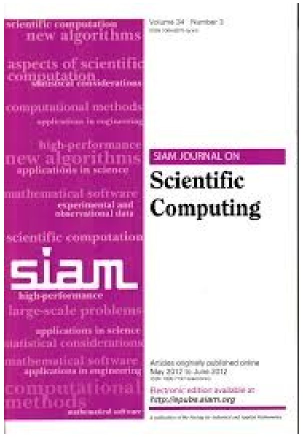
$P_{\delta U} = E[\delta U \delta U^T] \approx H^{-1}(\bar{U}) \approx H^{-1}(\hat{U})$ - estimation error covariance

Outline

1. Education and background
2. General introduction
3. **Uncertainty quantification**
4. Design problems
5. Parts not included
6. Current and future work

Analysis error covariance

SIAM J. SCI. COMPUT.
Vol. 30, No. 4, pp. 1847–1874



ON ANALYSIS ERROR COVARIANCES IN VARIATIONAL DATA ASSIMILATION*

I. YU. GEJADZE[†], F.-X. LE DIMET[‡], AND V. SHUTYAEV[§]

Abstract. The problem of variational data assimilation for a nonlinear evolution model is formulated as an optimal control problem to find the initial condition function (analysis). The equation for the analysis error is derived through the errors of the input data (background and observation errors). This equation is used to show that in a nonlinear case the analysis error covariance operator can be approximated by the inverse Hessian of an auxiliary data assimilation problem which involves the tangent linear model constraints. The inverse Hessian is constructed by the quasi-Newton BFGS algorithm when solving the auxiliary data assimilation problem. A fully nonlinear ensemble procedure is developed to verify the accuracy of the proposed algorithm. Numerical examples are presented.

Key words. data assimilation, optimal control, analysis error, Hessian, covariance operator

AMS subject classifications. 65K10, 35B37

DOI. 10.1137/07068744X

Analysis error covariance

Well known fact:

- analysis error covariance can be approximated by the inverse of the truncated Hessian (Hessian of auxiliary control problem)

Novel results of the paper:

- clean formulations, both in terms of operators and evolutionary equations (continuous time-space)

- computation of the inverse Hessian using the LBFGS

- a unique set of numerical examples, showing dependence of analysis error variance and correlations on transport mechanisms and parameters of the background covariance

Optimality system:

Cost-function and gradient

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

$$\begin{cases} -\dfrac{\partial \varphi^*}{\partial t} - (F'(\varphi))^* \varphi^* = -C^* R^{-1}(C\varphi - \varphi_{obs}), \quad t \in (0,T) \\ \varphi^*\big|_{t=T} = 0. \end{cases}$$

Gauss-Newton versus Newton

Truncated Hessian-vector product:

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

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

$$H(u)v = B^{-1}v - \psi^*\big|_{t=0}.$$

Hessian-vector product:

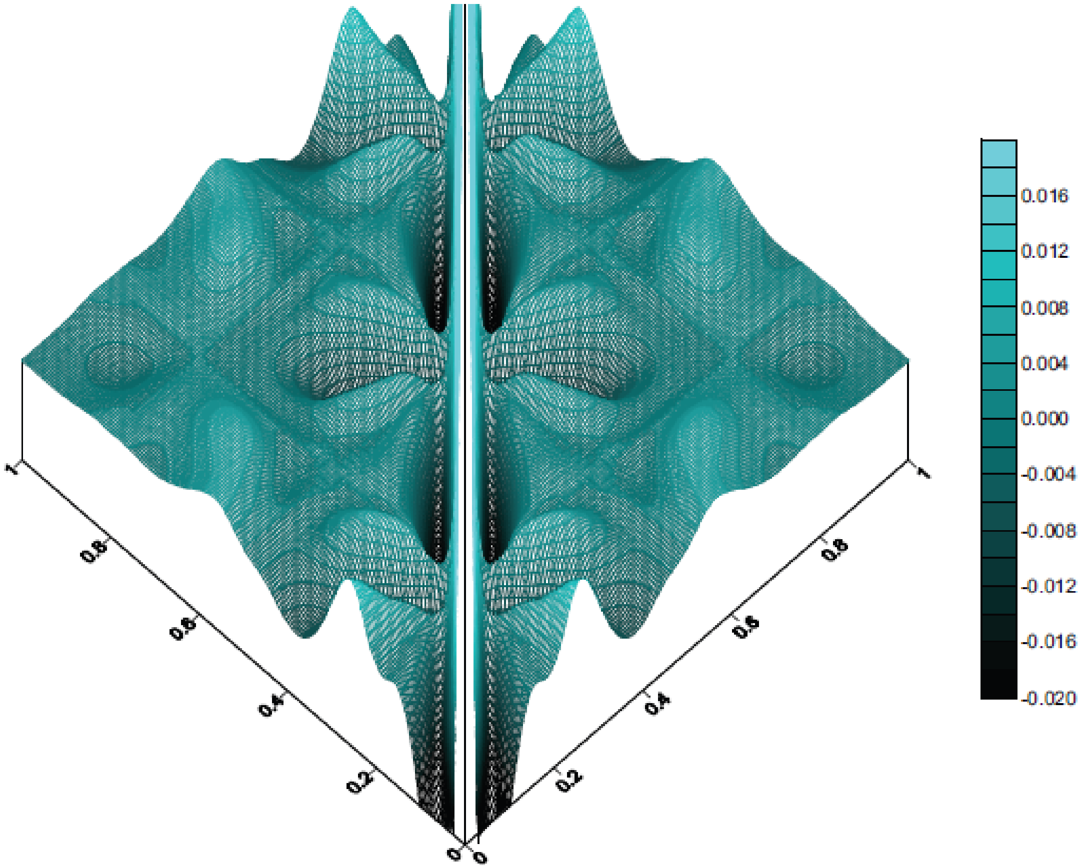
$$\begin{cases} \dfrac{\partial \psi}{\partial t} - F'(\varphi)\psi = 0, \ t \in (0,T), \\ \psi\big|_{t=0} = v, \end{cases}$$

$$\begin{cases} -\dfrac{\partial \psi^*}{\partial t} - (F'(\varphi))^* \psi^* = (F''(\varphi)\psi)^* \varphi^* - C^* R^{-1} C\psi, \ t \in (0,T) \\ \psi^*\big|_{t=T} = 0, \end{cases}$$

$$\mathcal{H}(u)v = B^{-1}v - \psi^*\big|_{t=0}.$$

Example of the inverse Hessian obtained by LBFGS

FIG. 1. *Covariance: linear diffusion problem.*



Analysis error covariance: parameter and boundary value estimation problems

Journal of Computational Physics 229 (2010) 2159–2178

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

On optimal solution error covariances in variational data assimilation problems

I.Yu. Gejadze [a], F.-X. Le Dimet [b], V. Shutyaev [c,*]

[a] Department of Civil Engineering, University of Strathclyde, 107 Rottenrow, Glasgow G4 0NG, UK
[b] MOISE Project (CNRS, INRIA, UJF, INPG); LJK, Université Joseph Fourier, BP 51, 38051 Grenoble Cedex 9, France
[c] Institute of Numerical Mathematics, Russian Academy of Sciences, 119333 Gubkina 8, Moscow, Russia

ARTICLE INFO

Article history:
Received 14 July 2009
Received in revised form 16 November 2009
Accepted 18 November 2009
Available online 24 November 2009

Keywords:
Variational data assimilation
Parameter estimation
Optimal solution error covariances
Hessian preconditioning

ABSTRACT

The problem of variational data assimilation for a nonlinear evolution model is formulated as an optimal control problem to find unknown parameters such as distributed model coefficients or boundary conditions. The equation for the optimal solution error is derived through the errors of the input data (background and observation errors), and the optimal solution error covariance operator through the input data error covariance operators, respectively. The quasi-Newton BFGS algorithm is adapted to construct the covariance matrix of the optimal solution error using the inverse Hessian of an auxiliary data assimilation problem based on the tangent linear model constraints. Preconditioning is applied to reduce the number of iterations required by the BFGS algorithm to build a quasi-Newton approximation of the inverse Hessian. Numerical examples are presented for the one-dimensional convection–diffusion model.

Analysis error variance and correlation matrices for parameter and boundary value estimation problems

Parameter estimation problem

$$\begin{cases} \frac{\partial \varphi}{\partial t} = F(\varphi, \lambda) + f, \quad t \in (0,T) \\ \varphi\big|_{t=0} = u, \\ J(\lambda) = \inf_{v \in \mathcal{Z}_p} J(v). \end{cases}$$

$$J(\lambda) = \frac{1}{2}\|R^{-1/2}(C(\varphi) - Y^*)\|^2_{\mathcal{Y}} + \frac{1}{2}\|B^{-1/2}(\lambda - \lambda^*)\|^2_{\mathcal{X}_p}$$

Truncated Hessian-vector product:

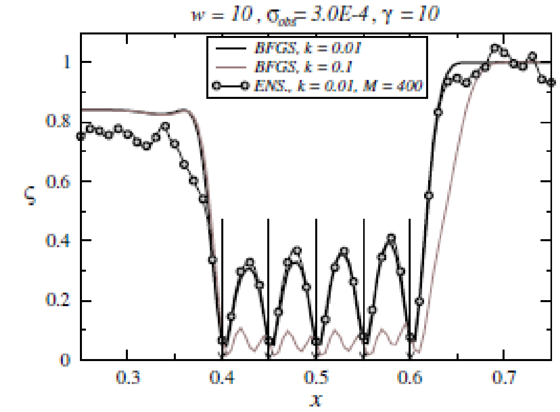
$$\begin{cases} \frac{\partial \psi}{\partial t} - F'_\varphi(\bar{\varphi}, \bar{\lambda})\psi = \boxed{F'_\lambda(\bar{\varphi}, \bar{\lambda})v}, \; t \in (0,T), \\ \psi|_{t=0} = 0, \end{cases}$$

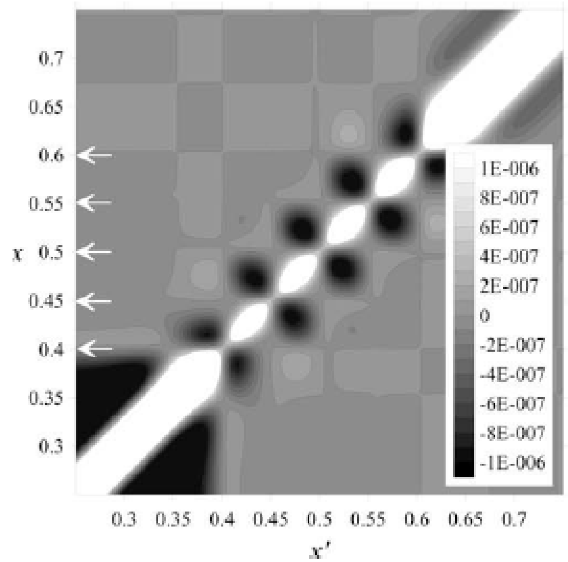
$$\begin{cases} -\frac{\partial \psi^*}{\partial t} - (F'_\varphi(\bar{\varphi}, \bar{\lambda}))^* \psi^* = -(C'(\bar{\varphi}))^* R^{-1} C'(\bar{\varphi})\psi, \quad t \in (0,T) \\ \psi^*\big|_{t=T} = 0, \end{cases}$$

$$H(\hat{\lambda})v = B^{-1}v - \boxed{(F'_\lambda(\bar{\varphi}, \bar{\lambda}))^* \psi^*}.$$

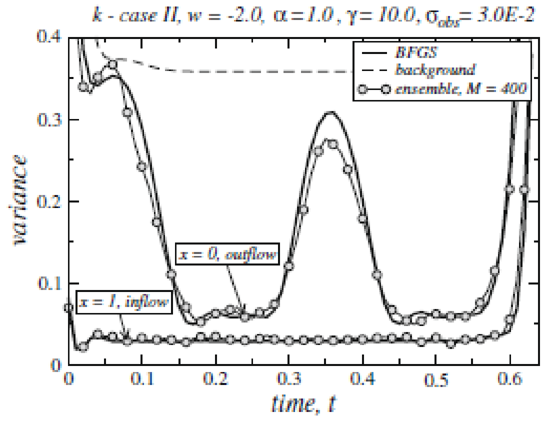
Illustration: variances and correlation matrices

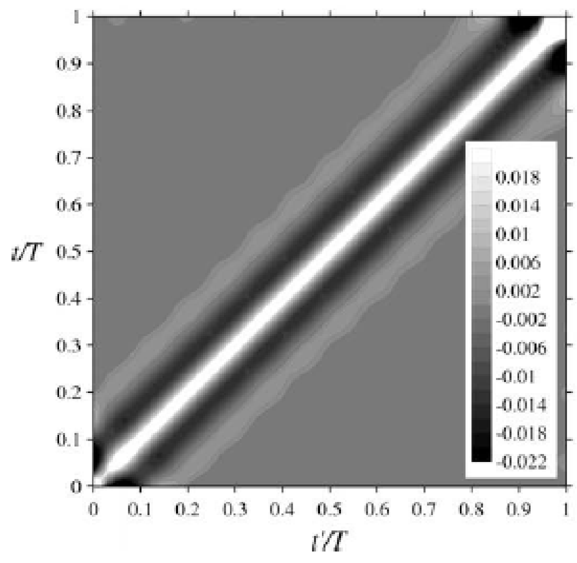
Diffusion coefficient estimation problem

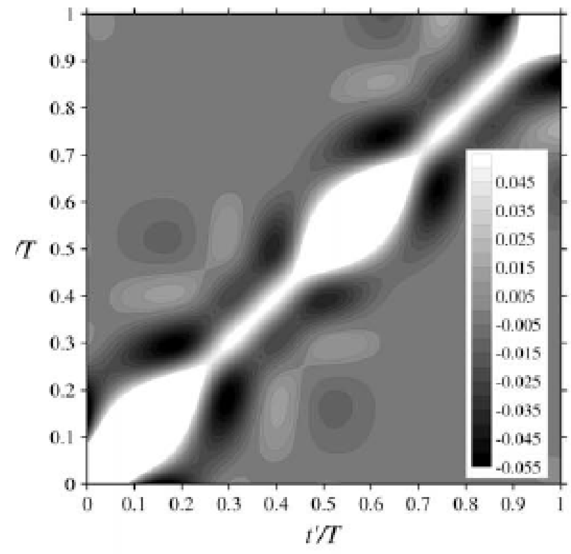




Boundary flux estimation problem







Main results:

1. In the linear case, the estimation/analysis error covariance is equal to the inverse Hessian

2. In the non-linear case, the estimation error covariance can be approximated by the inverse Hessian computed at the “truth”

3. The approximation error consists of:
   a) linearisation error
   b) origin error (due to difference between the optimal solution and the truth). Optimal solution may not always be achieved during minimization procedure

How to reduce this error?

substitute 'local' estimation of the covariance by 'global' estimation

Effective inverse Hessian

Journal of Computational Physics 230 (2011) 7923–7943

Contents lists available at SciVerse ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Computation of the analysis error covariance in variational data assimilation problems with nonlinear dynamics

I.Yu. Gejadze [a,*], G.J.M. Copeland [a], F.-X. Le Dimet [b], V. Shutyaev [c]

[a] *Department of Civil Engineering, University of Strathclyde, 107 Rottenrow, Glasgow, G4 ONG, UK*
[b] *MOISE project (CNRS, INRIA, UJF, INPG); LJK, Université Joseph Fourier, BP 51, 38051 Grenoble Cedex 9, France*
[c] *Institute of Numerical Mathematics, Russian Academy of Sciences, Gubkina 8, Moscow 119333, Russia*

A R T I C L E I N F O

Article history:
Received 26 April 2010
Received in revised form 18 March 2011
Accepted 19 March 2011
Available online 30 March 2011

Keywords:
Large-scale flow models
Nonlinear dynamics
Data assimilation
Optimal control
Analysis error covariance
Inverse Hessian
Ensemble methods
Monte Carlo

A B S T R A C T

The problem of variational data assimilation for a nonlinear evolution model is formulated as an optimal control problem to find the initial condition function. The data contain errors (observation and background errors), hence there will be errors in the optimal solution. For mildly nonlinear dynamics, the covariance matrix of the optimal solution error can often be approximated by the inverse Hessian of the cost functional. Here we focus on highly nonlinear dynamics, in which case this approximation may not be valid. The equation relating the optimal solution error and the errors of the input data is used to construct an approximation of the optimal solution error covariance. Two new methods for computing this covariance are presented: the fully nonlinear ensemble method with sampling error compensation and the 'effective inverse Hessian' method. The second method relies on the efficient computation of the inverse Hessian by the quasi-Newton BFGS method with preconditioning. Numerical examples are presented for the model governed by Burgers equation with a nonlinear viscous term.

Effective inverse Hessian - 1

1. New estimate is suggested

$$P = E[H^{-1}(\bar{U} + \delta U)], \; \delta U \sim \rho_a(U) \qquad (\text{not } P = (E[H(\bar{U} + \delta U)])^{-1}$$

where

$$\rho_a(U) = const \cdot exp\left(-\frac{1}{2}\|R^{-1/2}(G(U) - G(\bar{U}))\|^2_{\mathcal{Y}} - \frac{1}{2}\|B^{-1/2}(U - \bar{U})\|^2_{\mathcal{U}}\right)$$

2. A motivation is provided (no rigorous proof is available)

3.1 Implementation: 'pseudo-random' (requires ensemble of opt. solutions)

$$P = \frac{1}{L}\sum_{l=1}^{L} H^{-1}(\hat{U}_l)$$

where $\hat{U}_l$ are elements from the ensemble of estimates $\{\hat{U}_l\}, \; l = 1, \ldots, L$

3.2 Implementation: 'quasi-random' (does not require ensemble of optimal solutions)

$$\left\{ \begin{array}{rcl} P^{k+1} & = & \frac{1}{L}\sum_{l=1}^{L} H^{-1}(\bar{U} + \delta U_l^k), \\ P^0 & = & H^{-1}(\bar{U}), \quad k = 0, 1, \ldots \end{array} \right.$$

$$\delta U_l^k = (P^k)^{1/2}\eta_l, \quad \eta_l \sim N(0, I)$$

Effective inverse Hessian - 2

Efficient implementation using double preconditioning:

$$P = \frac{1}{L}\sum_{l=1}^{L} H^{-1}(\hat{U}_l)$$

$$P = B^{1/2}\tilde{H}^{-1/2}(\bar{U})\left(\frac{1}{L}\sum_{l=1}^{L}\tilde{\tilde{H}}^{-1}(\bar{U}+\delta U_l)\right)\tilde{H}^{-1/2}(\bar{U})(B^{1/2})^*$$

$$\tilde{H}(\cdot) = (B^{1/2})^* H(\cdot) B^{1/2} = (B^{1/2})^*(G'_U(\cdot))^* R^{-1} G'_U(\cdot) B^{1/2} + I$$

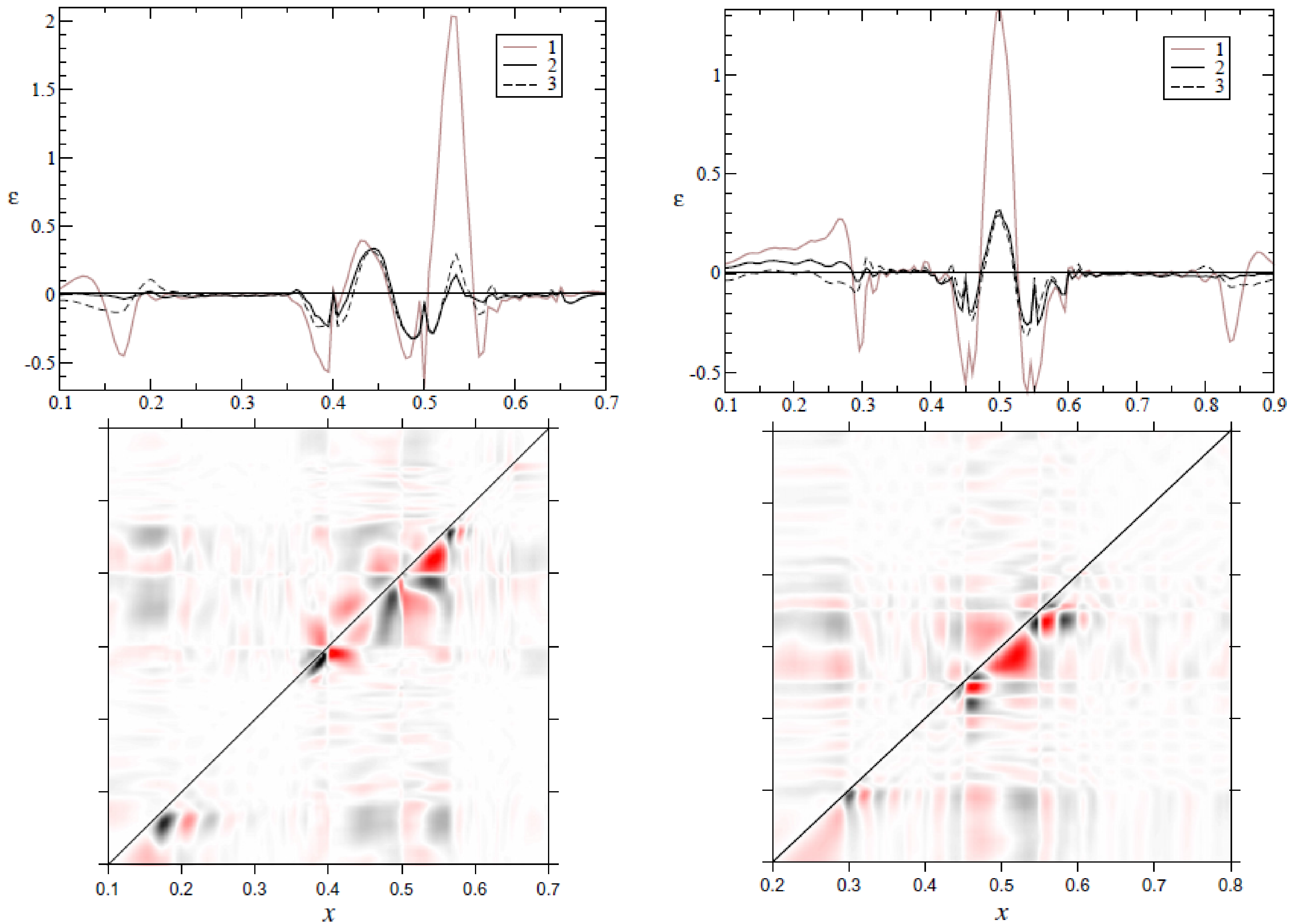
$$\tilde{\tilde{H}}(\bar{U}+\delta U_l) = \tilde{H}^{-1/2}(\bar{U})\tilde{H}(\bar{U}+\delta U_l)\tilde{H}^{-1/2}(\bar{U})$$

All projected Hessians are presented in the limited-memory form via eigenpairs!

$$A^{\beta}\cdot v = I\cdot v + \sum_{k=1}^{K}(\lambda_k^{\beta}-1)W_k(W_k)^*\cdot v.$$

Effective inverse Hessian - 3

Relative error in variance and error in correlation matrix, by the inverse Hessian and the effective inverse Hessian



Bayesian posterior covariance

Quarterly Journal of the Royal Meteorological Society

Q. J. R. Meteorol. Soc. 139: 1826–1841, October 2013 A

Royal Meteorological Society

Analysis error covariance *versus* posterior covariance in variational data assimilation

I. Yu. Gejadze,[a*] V. Shutyaev[b] and F.-X. Le Dimet[c]

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

[b]*Institute of Numerical Mathematics, Russian Academy of Sciences, MIPT, Moscow, Russia*

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

*Correspondence to: I. Yu. Gejadze, Department of Civil Engineering, University of Strathclyde, 107 Rottenrow, Glasgow G4 0NG, UK. E-mail: igor.gejadze@strath.ac.uk

Bayesian posterior covariance - 1

Bayesian posterior covariance and analysis error covariance are different objects (due to different centring of data)

Analysis error covariance: $P = E_a[(U-\bar{U})(U-\bar{U})^T] = E_a[\delta U \delta U^T]$

Bayesian posterior covariance: $\mathcal{P} = E_p\left[(U - E_p[U])(U - E_p[U])^T\right]$

Analysis error pdf:

$$\rho_a(U) = const \cdot exp\left(-\frac{1}{2}\|R^{-1/2}(G(U) - G(\bar{U}))\|^2_{\mathcal{Y}} - \frac{1}{2}\|B^{-1/2}(U-\bar{U})\|^2_{\mathcal{U}}\right)$$

Bayesian estimate pdf:

$$\rho_p(U) = const \cdot exp\left(-\frac{1}{2}\|R^{-1/2}(G(U) - Y_0^*)\|^2_{\mathcal{Y}} - \frac{1}{2}\|B^{-1/2}(U-U_0^*)\|^2_{\mathcal{U}}\right)$$

Bayesian posterior covariance - 2

Bayesian posterior covariance via Hessians: double product formula

a) local (can be found in the literature on statistics)

$$\mathcal{P} = \mathcal{H}^{-1}(\hat{U}_0) H(\hat{U}_0) \mathcal{H}^{-1}(\hat{U}_0)$$

b) global (new)

$$\mathcal{P} = E[\mathcal{H}^{-1}(\hat{U}_l) H(\hat{U}_l) \mathcal{H}^{-1}(\hat{U}_l)], \ ; \delta U \sim \rho_a(U)$$

Efficient implementation using double preconditioning:

$$\mathcal{P} = B^{1/2} \tilde{H}^{-1/2}(\hat{U}_0) \ A(\hat{U}_l) \ \tilde{H}^{-1/2}(\hat{U}_0)(B^{1/2})^*,$$
$$A(\hat{U}_l) = \frac{1}{L} \sum_{l=1}^{L} \tilde{\tilde{H}}^{-1/2}(\hat{U}_l) \tilde{\tilde{\mathcal{H}}}^{-\alpha}(\hat{U}_l) \tilde{\tilde{H}}^{-1/2}(\hat{U}_l)$$

The first-level preconditioning

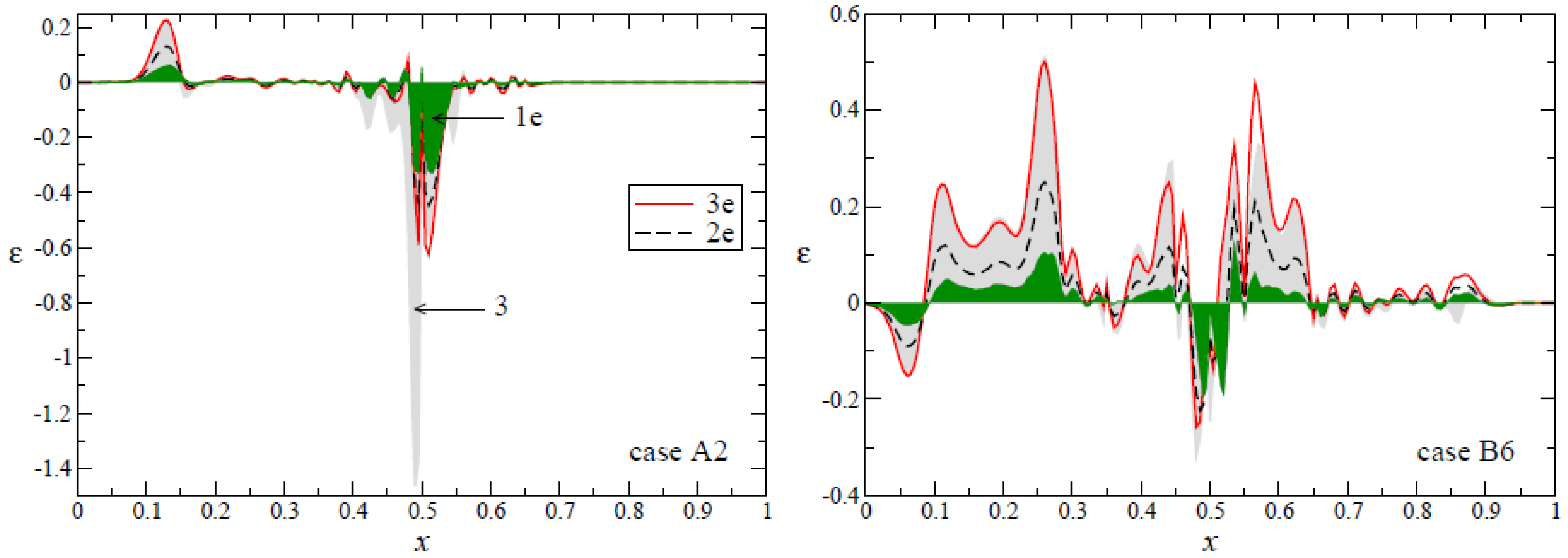
$$\tilde{H}(\cdot) = (B^{1/2})^* H(\cdot) B^{1/2}, \quad \tilde{\mathcal{H}}(\cdot) = (B^{1/2})^* \mathcal{H}(\cdot) B^{1/2}$$

the second-level preconditioning

$$\tilde{\tilde{H}}(\cdot) = \tilde{H}^{-1/2}(\hat{U}_0) \ \tilde{H}(\cdot) \ \tilde{H}^{-1/2}(\hat{U}_0), \quad \tilde{\tilde{\mathcal{H}}}(\cdot) = \tilde{H}^{-1/2}(\cdot) \ \tilde{\mathcal{H}}(\cdot) \ \tilde{H}^{-1/2}(\cdot)$$

Bayesian posterior covariance - 3

Numerical illustration in terms of the relative error:



Red – approximation by the 'effective' inverse Hessian:

$$P = \frac{1}{L}\sum_{l=1}^{L} H^{-1}(\hat{U}_l)$$

Green – approximation by the 'effective' double-product formula

$$\mathcal{P} = \frac{1}{L}\sum_{l=1}^{L} \mathcal{H}^{-1}(\hat{U}_l) H(\hat{U}_l) \mathcal{H}^{-1}(\hat{U}_l)$$

Bayesian posterior covariance - 4

Main results:

1. Bayesian posterior covariance and analysis error covariance are different objects (due to different centring of data)

2. They coincide in the linear case

3. Analysis error covariance relies on the truncated Hessian and, therefore, no 'second-order' information is involved

4. Posterior covariance involves both truncated and full Hessians (second derivative)

5. In practice, the double-product formula is unstable in local version and is only useful in 'effective' implementation

6. 'Effective' estimates can be built using a relatively small sample ($L<100$)

Since we obtain a sample of Hessians (either truncated or full), one may access higher moments of the posterior distribution !

Gauss-verifiability

Journal of Computational Physics 280 (2015) 439–456

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

ELSEVIER

On gauss-verifiability of optimal solutions in variational data assimilation problems with nonlinear dynamics

CrossMark

I.Yu. Gejadze [a,*], V. Shutyaev [b]

[a] UMR G-EAU, IRSTEA-Montpellier, 361 Rue J.F. Breton, BP 5095, 34196, Montpellier, France
[b] Institute of Numerical Mathematics, Russian Academy of Sciences, Moscow Institute for Physics and Technology, 119333 Gubkina 8, Moscow, Russia

A R T I C L E I N F O

Article history:
Received 12 June 2014
Received in revised form 24 September 2014
Accepted 26 September 2014
Available online 2 October 2014

Keywords:
Large-scale geophysical flow model
Nonlinear dynamics
Data assimilation
Optimal control
Identifiability
Confidence region
Analysis error covariance
Non-gaussianity

A B S T R A C T

The problem of variational data assimilation for a nonlinear evolution model is formulated as an optimal control problem to find the initial condition. The optimal solution (analysis) error arises due to the errors in the input data (background and observation errors). Under the gaussian assumption the confidence region for the optimal solution error can be constructed using the analysis error covariance. Due to nonlinearity of the model equations the analysis pdf deviates from the gaussian. To a certain extent the gaussian confidence region built on a basis of a non-gaussian analysis pdf remains useful. In this case we say that the optimal solution is "gauss-verifiable". When the deviation from the gaussian further extends, the optimal solutions may still be partially (locally) gauss-verifiable. The aim of this paper is to develop a diagnostics to check gauss-verifiability of the optimal solution. We introduce a relevant measure and propose a method for computing decomposition of this measure into the sum of components associated to the corresponding elements of the control vector. This approach has the potential for implementation in realistic high-dimensional cases. Numerical experiments for the 1D Burgers equation illustrate and justify the presented theory

Gauss-verifiability - 1

To check gaussianity of the estimator one can use classical tests for multivariate normality (Jarque-Berra, etc.):

1. Local (computed at a reference solution)
2. One integral value characterising normality of the optimal solution vector
3. Requires inversion of the sample covariance (deficient rank matrix)

New multivariate normality test is suggested:

1. Global (computed in the vicinity of a reference solution)
2. Distributed value (defined for each element of the control vector)
3. Requires inversion of the Hessian (full rank matrix)

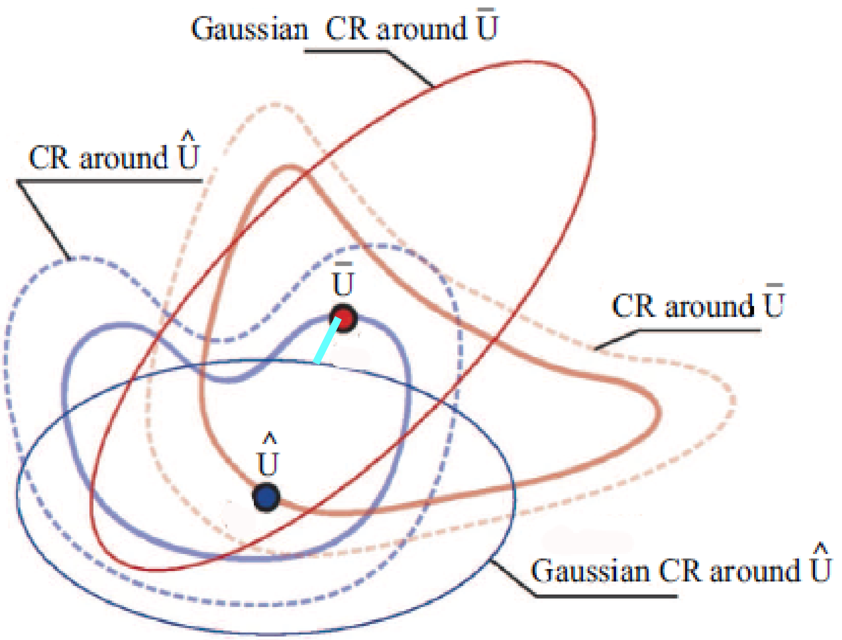
Gauss-verifiability - 2

Coexistence principle:

- estimate belongs to a certain confidence level of the 'true' distribution;
- 'truth' belongs to the corresponding confidence level built for the correct posterior distribution

Coexistence breach:

- 'truth' falls outside the confidence region built for the Gaussian posterior distribution, which approximates the correct one



Gauss-verifiability - 3

Coexistence measure (to quantify coexistence breach):

$$E[\theta(v,\hat{U})] = \int \theta(v,\hat{U})\rho_a(\hat{U}+v,\hat{U})dv$$

$$\theta(v,\hat{U}) = \frac{1}{2}\|P^{-1/2}(\hat{U}+v)v\|^2_{\mathcal{U}} - J(\hat{U}+v,\hat{U},G(\hat{U}))$$

where

$$J(U,U^*,Y^*) = \frac{1}{2}\|R^{-1/2}(G(U)-Y^*)\|^2_{\mathcal{Y}} + \frac{1}{2}\|B^{-1/2}(U-U^*)\|^2_{\mathcal{U}}$$

$$\rho_a(U) = const \cdot exp\left(-\frac{1}{2}\|R^{-1/2}(G(U)-G(\bar{U}))\|^2_{\mathcal{Y}} - \frac{1}{2}\|B^{-1/2}(U-\bar{U})\|^2_{\mathcal{U}}\right)$$

P - covariance of $\rho_a(U)$

Approximate coexistence measure:

$$\mathcal{D} = \frac{1}{2}tr\{E[P^{-1}(\hat{U}+v)P(\hat{U})]\} - C_1, \quad \text{where} \quad C_1 = E[J(\hat{U}+v,\hat{U},G(\hat{U}))]$$

Deconvolved coexistence measure:

$$\mathcal{D} = \sum_{i=1}^{n} d_i, \quad \text{where} \quad d_i = \frac{1}{2}(Ae_i,e_i)_{\mathcal{U}} - \frac{C_1}{n}, \quad A = E[Q^T P^{-1}(\hat{U}+v)Q]$$

Gauss-verifiability - 4

Deconvolved coexistence measure:

$$\mathcal{D} = \sum_{i=1}^{n} d_i, \qquad d_i = \frac{1}{2}(Ae_i, e_i)_{\mathcal{U}} - \frac{C_1}{n}, \quad A = E[Q^T P^{-1}(\hat{U} + v)Q]$$

d_i - contribution to coexistence measure by i-th element

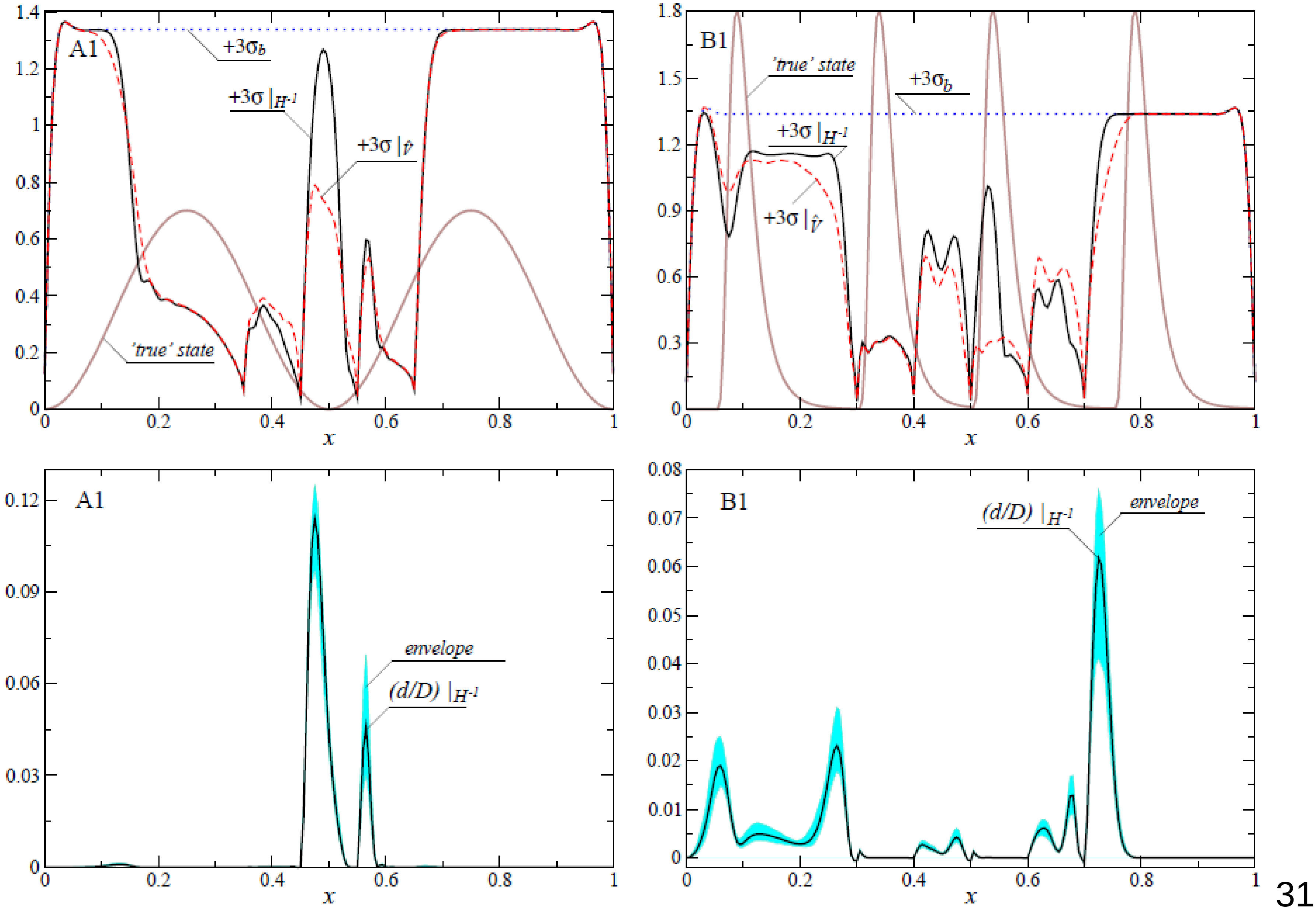
square-root decomposition $P(\hat{U}) = QQ^T, \quad Q : \mathcal{U} \to \mathcal{U}$

Assume: $P^{-1}(\hat{U} + v) \approx H(\hat{U} + v)$

$Q = H^{-1/2}(\hat{U})$

Finally: $\boxed{A = E[\tilde{\tilde{H}}(\hat{U} + v)]}$

Gauss-verifiability - 5



Gauss-verifiability – 5: Main results

1. New multivariate normality test is suggested: global, distributed value (i.e. for each element of the optimal solution vector)

2. This test allows to reveal the areas where the deviation from gaussianity is most noticeable. Knowing these areas is important (different options, e.g. local application of particle method of MCMC)

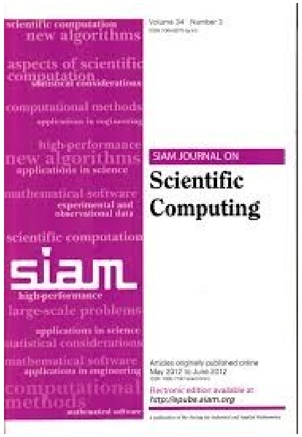
3. If the analysis error covariance is estimated via the 'effective' inverse Hessian, the distributed coexistence measure comes as by-product for free

Outline

1. Education and background
2. General introduction
3. Uncertainty quantification
4. **Design problems**
5. Parts not included
6. Current and future work

Sensor-location design problem

SIAM J. SCI. COMPUT.
Vol. 34, No. 2, pp. B127–B147



ON COMPUTATION OF THE DESIGN FUNCTION GRADIENT FOR THE SENSOR-LOCATION PROBLEM IN VARIATIONAL DATA ASSIMILATION*

I. YU. GEJADZE† AND V. SHUTYAEV‡

Abstract. The optimal sensor-location problem is considered in the framework of variational data assimilation for a large-scale dynamical model governed by partial differential equations. This problem is formulated as an optimization problem for the design function defined on the limited-memory approximation of the inverse Hessian of the data assimilation cost function. The expression for the gradient of the design function with respect to the sensor-location coordinates is derived via the adjoint to the Hessian derivative. An efficient algorithm for the gradient evaluation suitable for large-scale applications is suggested. This algorithm exploits the special structure of the limited-memory inverse Hessian defined by a small number of Ritz pairs obtained by the Lanczos method. If additional memory is allocated and certain data are stored during the computation of the Ritz pairs, no additional runs of the tangent linear model are required to evaluate the gradient. The accuracy of the gradients is checked in the numerical experiments. These gradients can be used for the gradient-based optimization of the design function within the chosen global optimization procedure.

Key words. optimal experiment design, sensor-location problem, design function gradient, large-scale flow models, variational data assimilation, limited-memory inverse Hessian, Lanczos method

AMS subject classifications. 65K10, 35B37

DOI. 10.1137/110825121

Problem statement

Assume having L sensors located at co-ordinates $\bar{x} = (x_1, \ldots, x_L)^T,\ x_i(t) \in \Omega$
Then, the observation operator C is defined as follows:

$$Y = CX = ((\phi(x, x_1), X)_{\mathcal{X}}, \ldots, (\phi(x, x_L), X)_{\mathcal{X}})^T$$

Data assimilation cost-function:

$$J(U) = \frac{1}{2}\|R^{-1/2}(C\mathcal{M}(U) - Y^*)\|^2_{\mathcal{Y}} + \frac{1}{2}\|B^{-1/2}(U - U^*)\|^2_{\mathcal{U}} \to \inf_U$$

Truncated Hessian associated with cost-function:

$$H(\cdot) = (\mathcal{M}'_U(\cdot))^* \ C^* R^{-1} C \ \mathcal{M}'_U(\cdot) + B^{-1}$$

Truncated Hessian equivalent to Fisher Information Matrix (FIM)!
Most (alphabetical) design criteria are based on FIM.

$$\Psi(\cdot, \bar{x}) = \sum_{i=1}^{m} p_i \left(H^{-1}(\cdot, \bar{x})e_i, e_i\right)_{R^m}, \ p_i = const \geq 0$$

If $p_i = 1, \ i = 1, \ldots, m$, then Ψ is the trace of H^{-1}

$\inf_{\bar{x}} \Psi(\cdot, \bar{x})$ is the classical A-optimality design criterion

Design function gradient

$$\Psi'_{\bar{x}}(\cdot,\bar{x}) \text{ - gradient } \Psi_{\bar{x}}(\cdot,\bar{x}) \text{ with respect to } \bar{x}$$

Due to the absence of inexpensive methods for computing the gradient of the design function, its minimization is generally performed using gradient-free methods !

Theorem. The gradient of the design function $\Psi(\cdot,\bar{x})$ with respect to $\bar{x}$ can be expressed via the gradients in the eigenvalues and eigenvectors of the projected Hessian by the formula

$$\Psi'_{\bar{x}}(\cdot,\bar{x}) = \sum_{k=1}^{K} \Psi'_{\bar{x},k}(\cdot,\bar{x}),$$

where $\Psi'_{\bar{x},k}(\cdot,\bar{x}) \approx s'_{\bar{x},k}(B^{1/2}U_k)^T P(B^{1/2}U_k) + 2(s_k - 1)(B^{1/2}U_k)^T P B^{1/2}\bar{U}V'_{\bar{x},k}$

with $s'_{\bar{x},k}$ and $V'_{\bar{x},k}$ satisfying some expressions

All elements required for computing these expressions are accumulated during construction of the Hessian itself (via eigenpairs by Lanczos) !

A new mathematical object: adjoint to the Hessian derivative - $(H'v)^*w$

Main results

The gradient of the design function can be computed at a negligible cost (in terms of CPU time) in the process of computing the design function itself

This is an important step toward the gradient-based optimization in optimal sensor location problem

Approach that can be used in other design problems, such as the optimal test-signal design for parameter identification, or optimal shape design

Control set design problem

Journal of Computational Physics 325 (2016) 358–379

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Design of the control set in the framework of variational data assimilation

I.Yu. Gejadze *, P.-O. Malaterre

UMR G-EAU, IRSTEA-Montpellier, 361 Rue J.F. Breton, BP 5095, 34196, Montpellier, France

A R T I C L E   I N F O

Article history:
Received 25 May 2016
Received in revised form 17 August 2016
Accepted 23 August 2016
Available online 29 August 2016

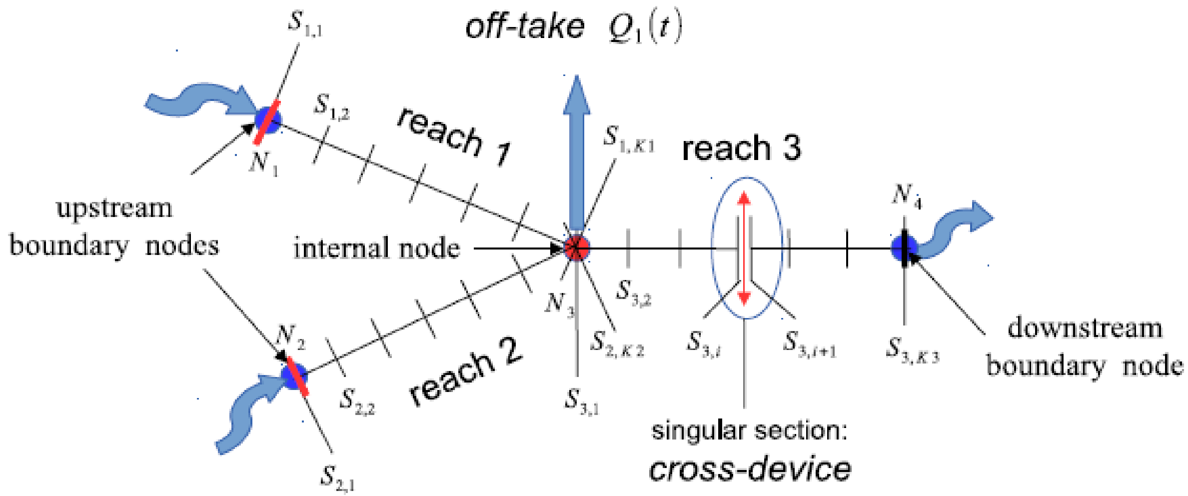
Keywords:
Control set design
Uncertainty quantification
Variational data assimilation
1D hydraulic network model
Automatic differentiation

A B S T R A C T

Solving data assimilation problems under uncertainty in basic model parameters and in source terms may require a careful design of the control set. The task is to avoid such combinations of the control variables which may either lead to ill-posedness of the control problem formulation or compromise the robustness of the solution procedure. We suggest a method for quantifying the performance of a control set which is formed as a subset of the full set of uncertainty-bearing model inputs. Based on this quantity one can decide if the chosen 'safe' control set is sufficient in terms of the prediction accuracy. Technically, the method presents a certain generalization of the 'variational' uncertainty quantification method for observed systems. It is implemented as a matrix-free method, thus allowing high-dimensional applications. Moreover, if the Automatic Differentiation is utilized for computing the tangent linear and adjoint mappings, then it could be applied to any multi-input 'black-box' system. As application example we consider the full Saint-Venant hydraulic network model SIC^2, which describes the flow dynamics in river and canal networks. The developed methodology seem useful in the context of the future SWOT satellite mission, which will provide observations of river systems the properties of which are known with quite a limited precision.

Motivation

River flow model and discharge estimation problem - 1



St-Venant equations:

a) Continuity

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = q$$

b) Momentum

$$\frac{\partial Q}{\partial t} + \frac{\partial Q^2/A}{\partial x} + gA\frac{\partial Z}{\partial x} = -gAS_f + kqV$$

Q - discharge

$A(Z, p_{geo})$ - wet cross-section area, p_{geo} - geometry parameters

Z - water surface elevation

$V = Q/A$ - mean velocity

$S_f(C_s)$ - friction function dependent on Strickler coefficient

q - lateral discharge

Motivation:

River flow model and discharge estimation problem - 2

Internal node equation:

off-take or tributary

$$Q|_{S_{1,k1}} + Q|_{S_{2,k2}} - Q|_{S_{3,1}} = \boxed{Q_1}$$

$$Z|_{S_{1,k1}} = Z|_{S_{3,1}}, \quad Z|_{S_{2,k2}} = Z|_{S_{3,1}}$$

or

$$H|_{S_{1,k1}} = H_{S_{3,1}}, \; H|_{S_{2,k2}} = H_{S_{3,1}}; \; H = V^2/g + Z$$

Cross-device (singular section) equation:

$$Q|_{S_{3,i}} - Q|_{S_{3,i+1}} = 0$$

or

$$Q|_{S_{3,i}} = \mathcal{F}(Z|_{S_{3,i}}, Z|_{S_{3,i+1}}, \; C_d \;)$$

Boundary conditions:

a) Upstream boundary nodes: inflow discharge $\mathcal{Q}(t)$, elevation $Z(t)$

b) Downstream boundary node: rating curve $Q(Z, p_{rc})$, elevation $Z(t)$

where p_{rc} - rating curve parameters

Initial conditions: Q_0, Z_0 - steady-state flow solution

Full input vector:

$$U = \{ \; \boxed{\mathcal{Q}, \; Q_i, \; q, \; \{Q_0, Z_0\},} \; \boxed{C_s, \; k, \; C_d, \; p_{rc}, \; p_{geo},} \; \boxed{p_{num}} \; \} \in \mathcal{U}$$

Flow controls, i.e. controls associated to state variables: $\{Q, Z\}$

Physical parameters

Parameters of numerical scheme

Motivation:

discharge estimation problem – what to control and how?

$$U = \{ \underbrace{Q,\ Q_i,\ q,\ \{Q_0, Z_0\},}_{\text{Flow controls, i.e. controls associated to state variables: } \{Q,Z\}}\ \underbrace{C_s,\ k,\ C_d,\ p_{rc},\ p_{geo},}_{\text{Physical parameters}}\ \underbrace{p_{num}}_{\text{Parameters of numerical scheme}} \} \in \mathcal{U}$$

SIC model does not support supercritical flow (as majority of models using the Preissmann scheme), and drying/wetting processes. As long as these flow states are encountered, the code execution stops with an appropriate error message.

These states does not normally occur when the model runs to describe physically meaningful conditions. However, during solving estimation problems, the current estimates may not be supported by the model ! (comment)
As a result, the estimation process breaks down.

Some parameters enter the model in a highly non-linear way, for example geometric parameters. Introducing these parameters into the control vector leads to ill-posedness (non-uniqueness (equifinality) / multiple local minima), which requires the global minimum search !

Careful design of the control set is necessary !

Question: – what to control and how?

It is impossible to control the full input U !

1. Technical/implementation issues:
   a) dimension of the control vector
   b) convergence rate
2. Fundamental:
   a) identifiability / equifinality;
   b) convexity;
   c) connectivity of the solution domain

Aim: design of a sufficient control set

tradeoff: accuracy versus robustness and solvability

Control set design: design function

Notations:

$\Psi \in \mathcal{D}$ - design function or Quantity of Interest (QoI) , $\mathcal{D}$ is a 'design' space

$\Psi = D(X)$ - 'state-to-design' mapping $D : \mathcal{X} \to \mathcal{D}$

$\delta\Psi = D(\mathcal{M}(\hat{U})) - D(\mathcal{M}(\bar{U}))$ - posterior design function error

$\delta\Psi = D'_X(\bar{X})\mathcal{M}'_U(\bar{U})\delta U$ - design function error linearized

$$P_{\delta\Psi} = E[\delta\Psi\delta\Psi^T] = D'_X(\bar{X})\mathcal{M}'_U(\bar{U})P_{\delta U}(\mathcal{M}'_U(\bar{U}))^*(D'_X(\bar{X}))^*$$

- design function error covariance for small δU

$P_{\delta U} = B$ - without DA

$P_{\delta U} \simeq H^{-1}(\bar{U}) \simeq H^{-1}(\hat{U})$ - after DA

Design function error covariance can be represented by a limited number of its eigenvalues / eigenvectors, given the product .

The above presented material is more or less well established results known as "**variational uncertainty quantification**".

Control set design: partial control -1

$U_a \in A$ - active subset of the full control vector U

$U_p = U \setminus U_a$ - passive subset of U

U_a^*, U_p^* - background/prior

$\bar{U}_a, \bar{U}_p$ - 'truth'/exact components of U

$\varepsilon_a = U_a^* - \bar{U}_a$, $\varepsilon_p = U_p^* - \bar{U}_p$ - background error in components of U

If $E[\varepsilon_a \varepsilon_p^T] = 0$, then B is block-diagonal with blocks B_a and B_p, where $B_a = E[\varepsilon_a \varepsilon_a^T]$ and $B_p = E[\varepsilon_p \varepsilon_p^T]$

$$J(U_a) = \frac{1}{2}\|R^{-1/2}(G(U_a, U_p^*) - Y^*)\|_{\mathcal{Y}}^2 + \frac{1}{2}\|B_a^{-1/2}(U_a - U_a^*)\|_A^2 \to \inf_{U_a}$$

- cost-function for U_a

$$\delta U_a \simeq H_a{}^{-1}(\bar{U})(G_{U_a}'^*(\bar{U})R^{-1}\xi + B_a^{-1}\varepsilon_a - G_{U_a}'^*(\bar{U})R^{-1}G_{U_p}'(\bar{U})\varepsilon_p)$$

- estimation error via ε_a, ε_p and ξ

$$H_a(\bar{U}) = G_{U_a}'^*(\bar{U})R^{-1}G_{U_a}'(\bar{U}) + B_a^{-1}$$

- Hessian of an auxiliary control problem for U_a

Control set design: partial control -2

$$P_{\delta\Psi} = E[\delta\Psi\delta\Psi^T] = D'_X(\bar{X})\mathcal{M}'_U(\bar{U})P_{\delta U}(\mathcal{M}'_U(\bar{U}))^*(D'_X(\bar{X}))^*$$

$\delta U = (\delta U_a, \varepsilon_p)^T$ - posterior error in the input vector

$P_{\delta U} = E[\delta U \cdot \delta U^T]$ - error covariance:

$$P_{\delta U} = \begin{pmatrix} E[\delta U_a \cdot \delta U_a^T] & E[\delta U_a \cdot \varepsilon_p^T] \\ E[\varepsilon_p \cdot \delta U_a^T] & E[\varepsilon_p \cdot \varepsilon_p^T] \end{pmatrix} = \begin{pmatrix} P_{\delta U_a} & P_{\delta U_{ap}} \\ P_{\delta U_{pa}} & B_p \end{pmatrix}$$

Elements of the covariance are expressed via blocks of the Hessian !

$$P_{\delta U_a} = E[\delta U_a \delta U_a^T] = H_a{}^{-1} + H_a{}^{-1} H_{ap} B_p H_{pa} H_a{}^{-1}$$

$$P_{\delta U_{ap}} = E[\delta U_a \varepsilon_p^T] = -H_a{}^{-1} H_{ap} B_p$$

$$P_{\delta U_{pa}} = E[\varepsilon_p \delta U_a^T] = -B_p H_{pa} H_a{}^{-1}$$

Control set design: NA results, case A

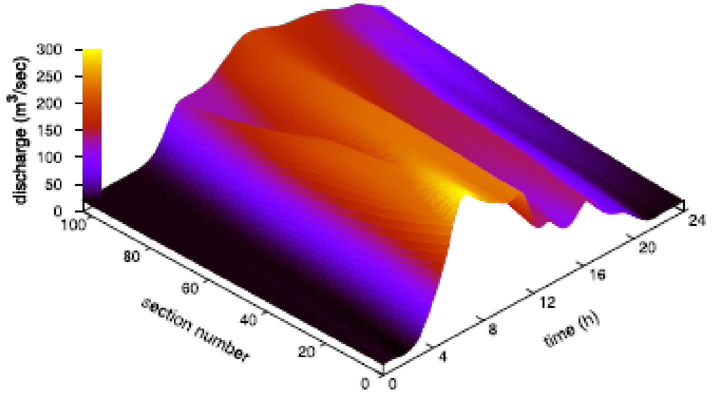
$$U = (Q_1(t), z_b(k), b(k), C_s(k), U_*)^T, \; k = 1, \ldots, K_s$$

cc1 - full control case: i.e. $U_a = U \setminus U_*$

$cc2$ – partial control case: $U_a = Q_1(t)$;

$cc3$ – partial control case: $U_a = (Q_1(t), z_b(k))^T$;

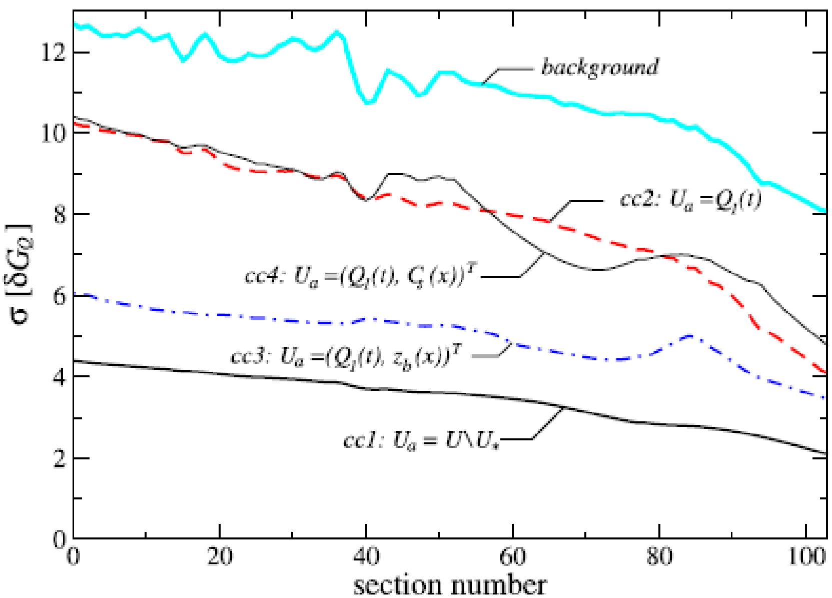
$cc4$ – partial control case: $U_a = (Q_1(t), C_s(k))^T$.

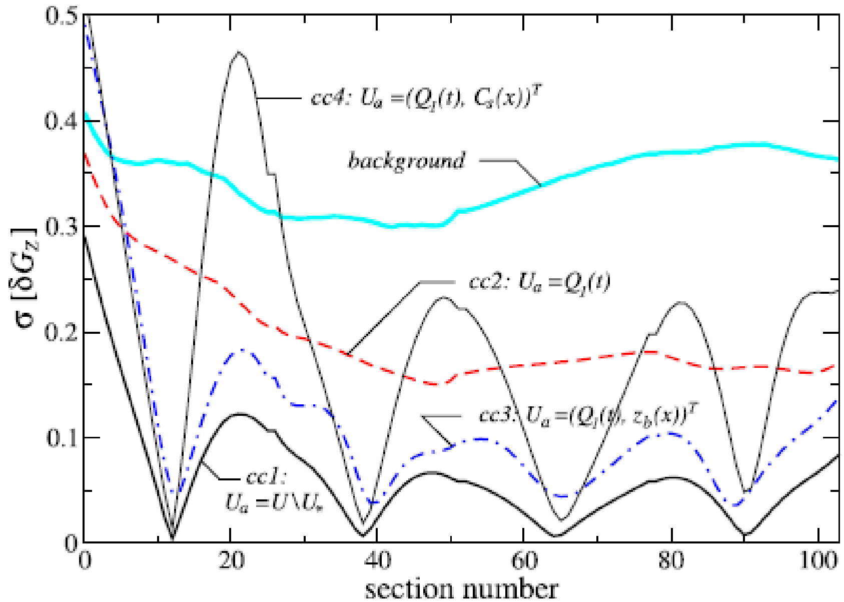


Standard deviation for uncertainties:

$$\sigma[\delta Q_1] = 25 m^3/s$$

$$\sigma[\delta z_b] = 0.33m, \; \sigma[\delta C_s] = 3.3 \;, \; \sigma[\delta b] = 0.07$$





Outline

1. Education and background
2. General introduction
3. Analysis error/posterior covariance
4. Design problems
5. **Parts not included**
6. Current and future work

Implicit (idle) control

Quarterly Journal of the Royal Meteorological Society *Q. J. R. Meteorol. Soc.* 143: 000–000, April 2017 B DOI:10.1002/qj.3102

RMetS
Royal Meteorological Society

Implicit treatment of model error using inflated observation-error covariance

I. Gejadze,[a]* H. Oubanas[a] and V. Shutyaev[b]

[a] *IRSTEA, Montpellier, France*
[b] *Institute of Numerical Mathematics, Russian Academy of Sciences, MIPT, Moscow, Russia*

*Correspondence to: I. Gejadze, 361 Rue J.F. Breton, BP 5095, 34196 Montpellier, France.
E-mail: igor.gejadze@irstea.fr

Inverse Hessian by multigrid approach

SIAM J. SCI. COMPUT.
Vol. 0, No. 0, pp. 000–000

A MULTILEVEL APPROACH FOR COMPUTING THE LIMITED-MEMORY HESSIAN AND ITS INVERSE IN VARIATIONAL DATA ASSIMILATION*

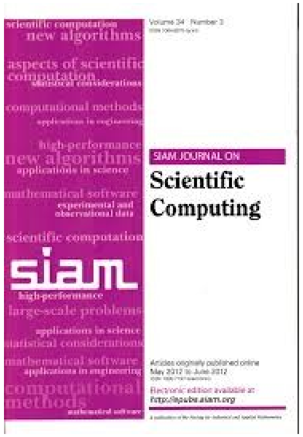
KIRSTY L. BROWN[†], IGOR GEJADZE[‡], AND ALISON RAMAGE[†]

Abstract. Use of data assimilation techniques is becoming increasingly common across many application areas. The inverse Hessian (and its square root) plays an important role in several different aspects of these processes. In geophysical and engineering applications, the Hessian-vector product is typically defined by sequential solution of a tangent linear and adjoint problem; for the inverse Hessian, however, no such definition is possible. Frequently, the requirement to work in a matrix-free environment means that compact representation schemes are employed. In this paper, we propose an enhanced approach based on a new algorithm for constructing a multilevel eigenvalue decomposition of a given operator, which results in a much more efficient compact representation of the inverse Hessian (and its square root). After introducing these multilevel approximations, we investigate their accuracy and demonstrate their efficiency (in terms of reducing memory requirements and/or computational time) using the example of preconditioning a Gauss–Newton minimization procedure.

Key words. data assimilation, inverse Hessian, limited memory, preconditioning, multigrid

AMS subject classifications. 65K05, 65K10, 15A09, 15A29

DOI. 10.1137/15M1041407



New mathematical concepts introduced ...

1. Multilevel eigenvalue decomposition

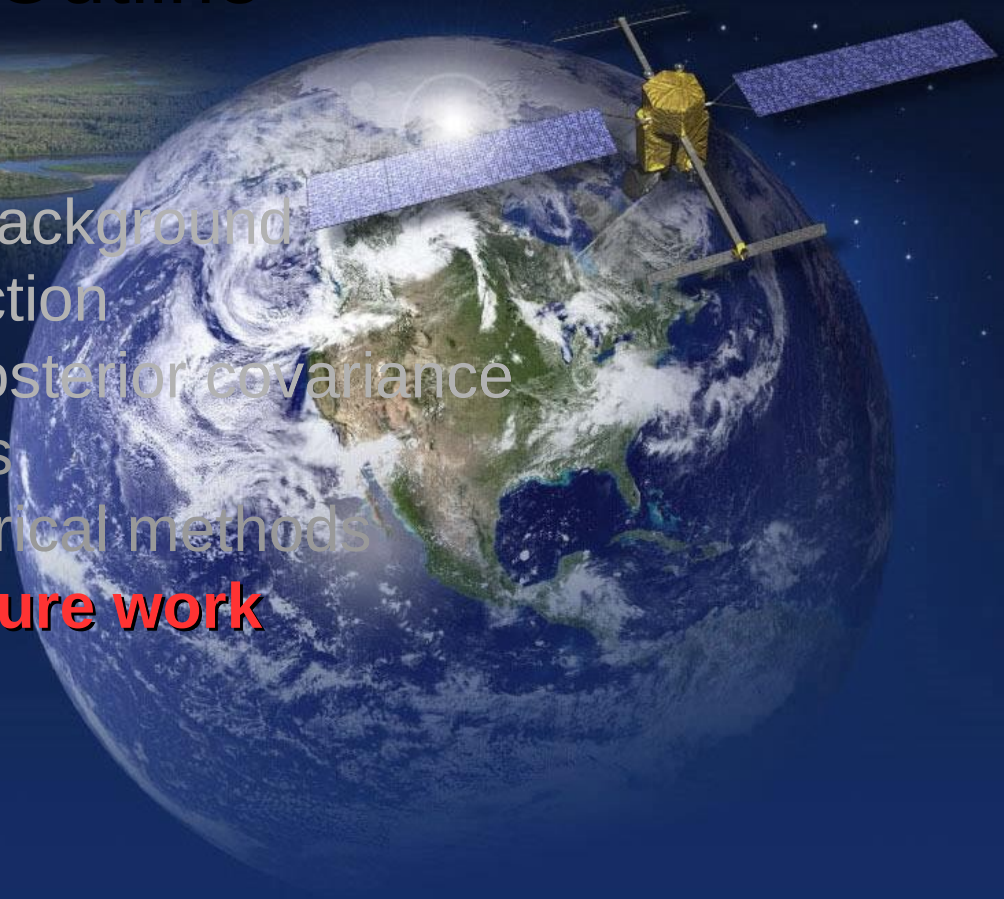
2. 'Effective' approximations for analysis error/ posterior covariances

3. New multivariate normality test: coexistence measure

4. Adjoint to Hessian derivative

5. Control set design

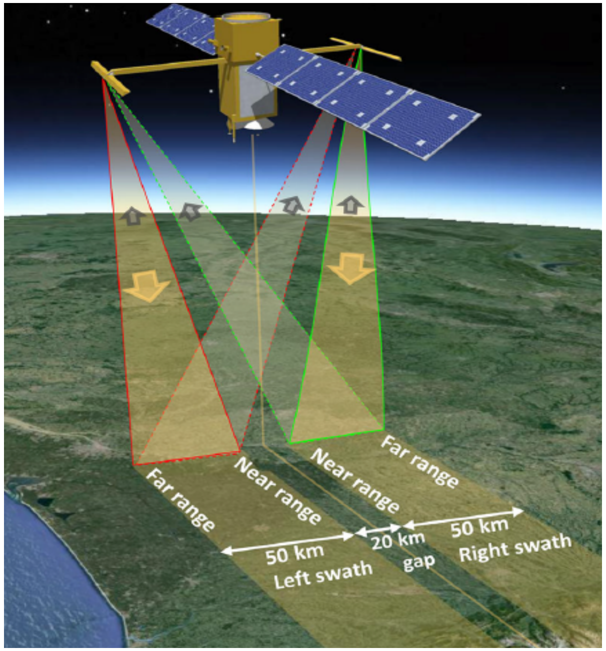
Outline

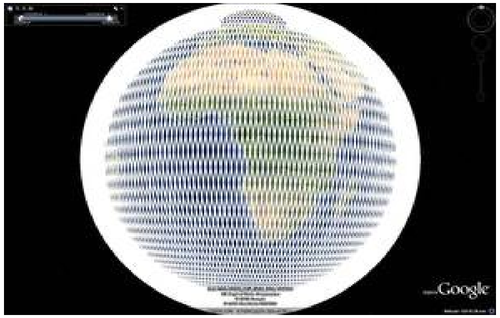


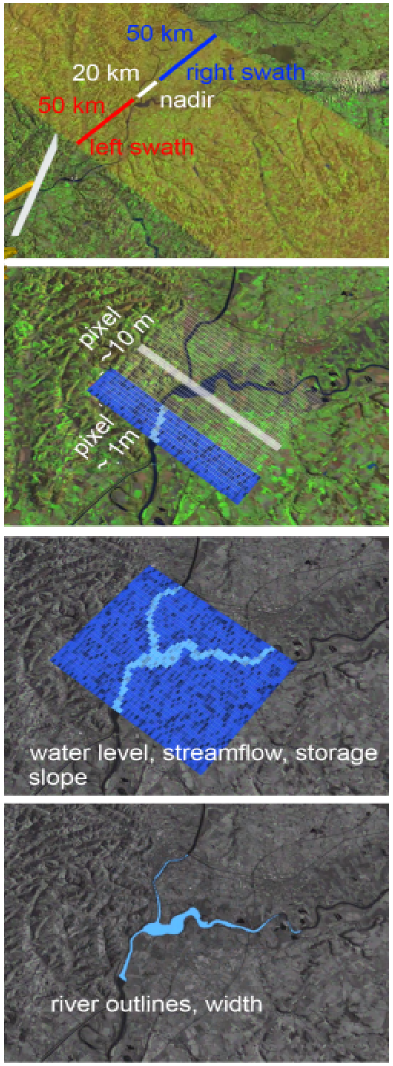
1. Education and background
2. General introduction
3. Analysis error/posterior covariance
4. Design problems
5. Advanced numerical methods
6. **Current and future work**

Variational DA in river hydraulics

Application: Surface Water and Ocean Topography (SWOT) mission







One of tasks of the mission is to monitor river discharge in ungauged basins.

Main difficulties:

- low accuracy of observations, low temporal frequency
- insufficient data on the river bed geometry and roughness, properties of the catchment area of interest and lateral inflows / off-takes

Variational DA in river hydraulics

Model: SIC^2 (Simulation and Integration of Control for Canals) - hydraulic network model, developed by IRSTEA (former CEMAGREF), since 1990.
Based on Saint-Venant equation (for a single reach), includes:

- coupling between reaches
- complex description of river bed
- storage areas, most of known hydraulic devices

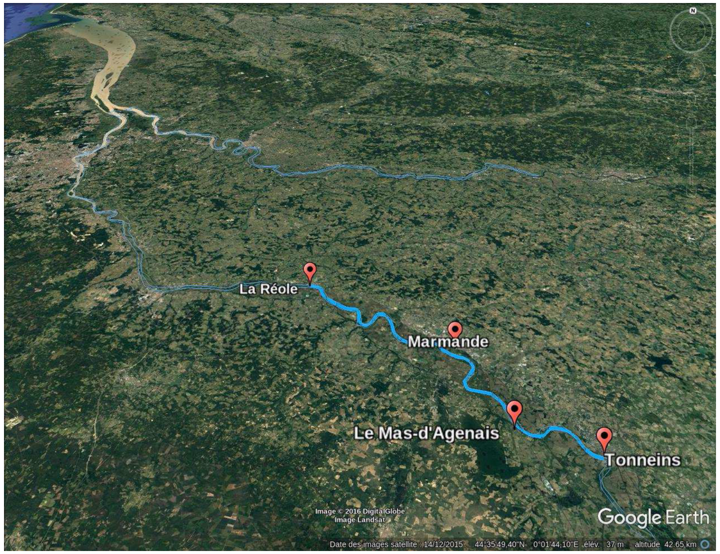
Main contributions:

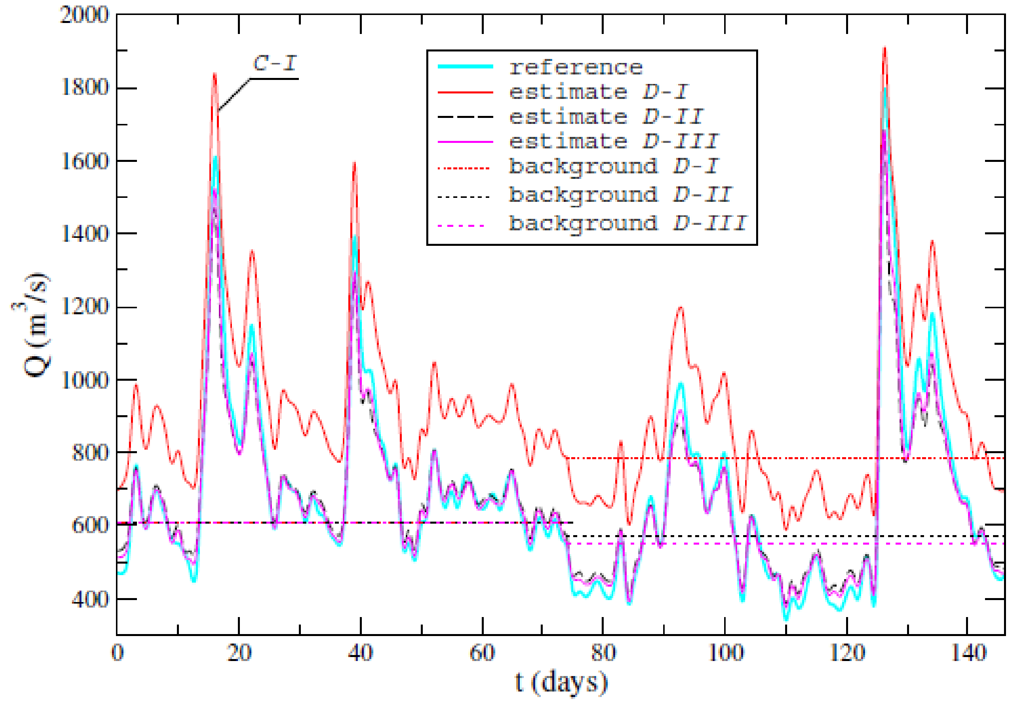
1. development of the adjoint and TL counterparts for SIC^2 (using Automatic Differentiation). Presently, it is the only known stable adjoint for models of this type (Mascaret, Mike11, ISIS, etc.)!

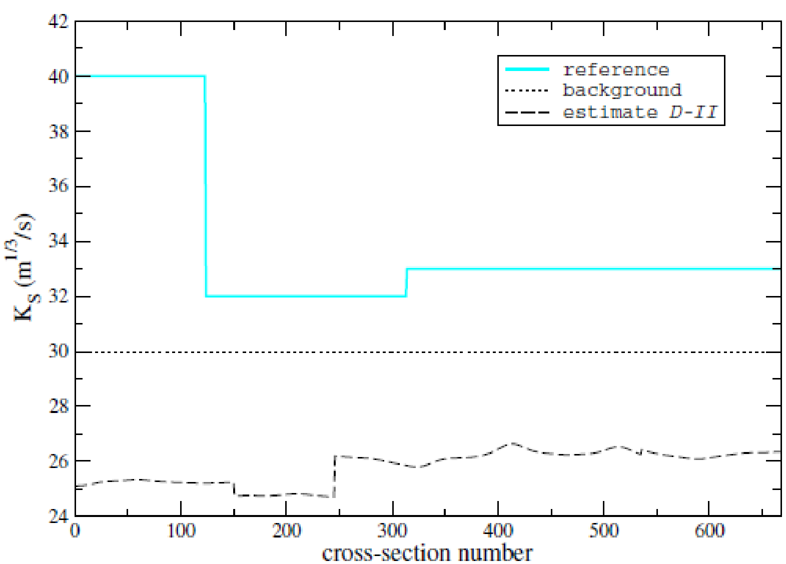
2. approach for solving discharge estimation problem under strong uncertainty in major model parameters: simultaneous estimation of discharge, bed elevation and bed roughness coefficient, iterative regularization

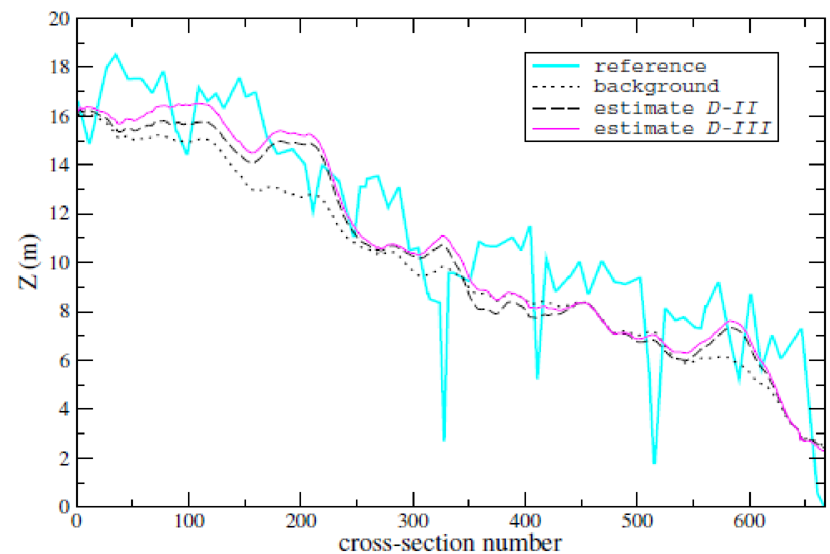
3. PhD by Hind Oubanas

Variational DA in river hydraulics



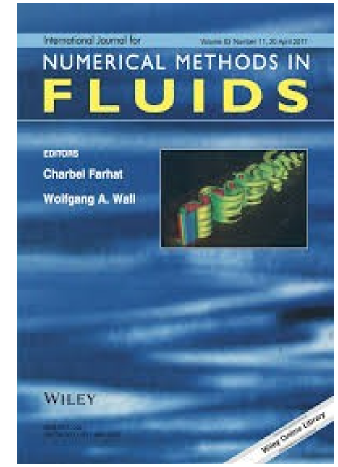






Variational DA in river hydraulics

INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids 2017; **83**:405–430
Published online 28 July 2016 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/fld.4273



Discharge estimation under uncertainty using variational methods with application to the full Saint-Venant hydraulic network model

Igor Gejadze and Pierre-Olivier Malaterre

UMR G-EAU IRSTEA-Montpellier, 361 Rue J.F. Breton, BP 5095, 34196 Montpellier, France

SUMMARY

Estimating river discharge from *in situ* and/or remote sensing data is a key issue for evaluation of water balance at local and global scales and for water management. Variational data assimilation (DA) is a powerful approach used in operational weather and ocean forecasting, which can also be used in this context. A distinctive feature of the river discharge estimation problem is the likely presence of significant uncertainty in principal parameters of a hydraulic model, such as bathymetry and friction, which have to be included into the control vector alongside the discharge. However, the conventional variational DA method being used for solving such extended problems often fails. This happens because the control vector iterates (i.e., approximations arising in the course of minimization) result into hydraulic states not supported by the model. In this paper, we suggest a novel version of the variational DA method specially designed for solving estimation-under-uncertainty problems, which is based on the ideas of iterative regularization.

The method is implemented with SIC^2, which is a full Saint-Venant based 1D-network model. The SIC^2 software is widely used by research, consultant and industrial communities for modeling river, irrigation canal, and drainage network behavior. The adjoint model required for variational DA is obtained by means of automatic differentiation. This is likely to be the first stable consistent adjoint of the 1D-network model of a commercial status in existence.

The DA problems considered in this paper are offtake/tributary estimation under uncertainty in the cross-device parameters and inflow discharge estimation under uncertainty in the bathymetry defining parameters and the friction coefficient. Numerical tests have been designed to understand identifiability of discharge given uncertainty in bathymetry and friction. The developed methodology, and software seems useful in the context of the future Surface Water and Ocean Topography satellite mission.

Received 29 October 2015; Revised 22 February 2016; Accepted 21 June 2016

Variational DA in river hydraulics

Accepted Manuscript

River discharge estimation from synthetic SWOT-type observations using variational data assimilation and the full Saint-Venant hydraulic model

Hind Oubanas, Igor Gejadze, Pierre-Olivier Malaterre, Franck Mercier

PII: S0022-1694(18)30080-5
DOI: https://doi.org/10.1016/j.jhydrol.2018.02.004
Reference: HYDROL 22553

To appear in: *Journal of Hydrology*

Please cite this article as: Oubanas, H., Gejadze, I., Malaterre, P-O., Mercier, F., River discharge estimation from synthetic SWOT-type observations using variational data assimilation and the full Saint-Venant hydraulic model, *Journal of Hydrology* (2018), doi: https://doi.org/10.1016/j.jhydrol.2018.02.004

Variational DA in hydrology (GR4J)

GR4J is a daily lumped four-parameter rainfall - runoff model, global (catchment)
Inputs: *P* – the rainfall depth, *E* – potential evapotranspiration (PE)
Output: *Q* - runoff

Parameters:

X_1 - maximum capacity of the production store;
X_2 - coefficient in groundwater exchange term;
X_3 - the routing store reference capacity ;
X_4 - time-base of unit hydrographs

State variables:

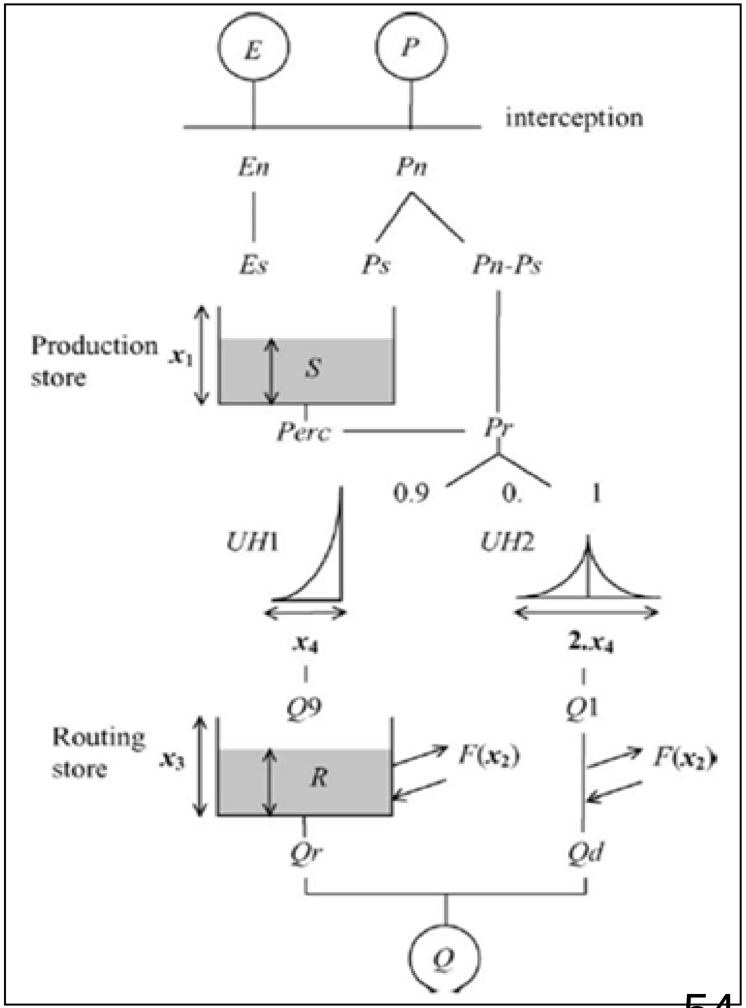
S - water content in the production store;
R - water content in the routing store

Initial state:

$$S(t=0), R(t=0), P_r(t=0, t=-dt, t=-2dt, \dots)$$

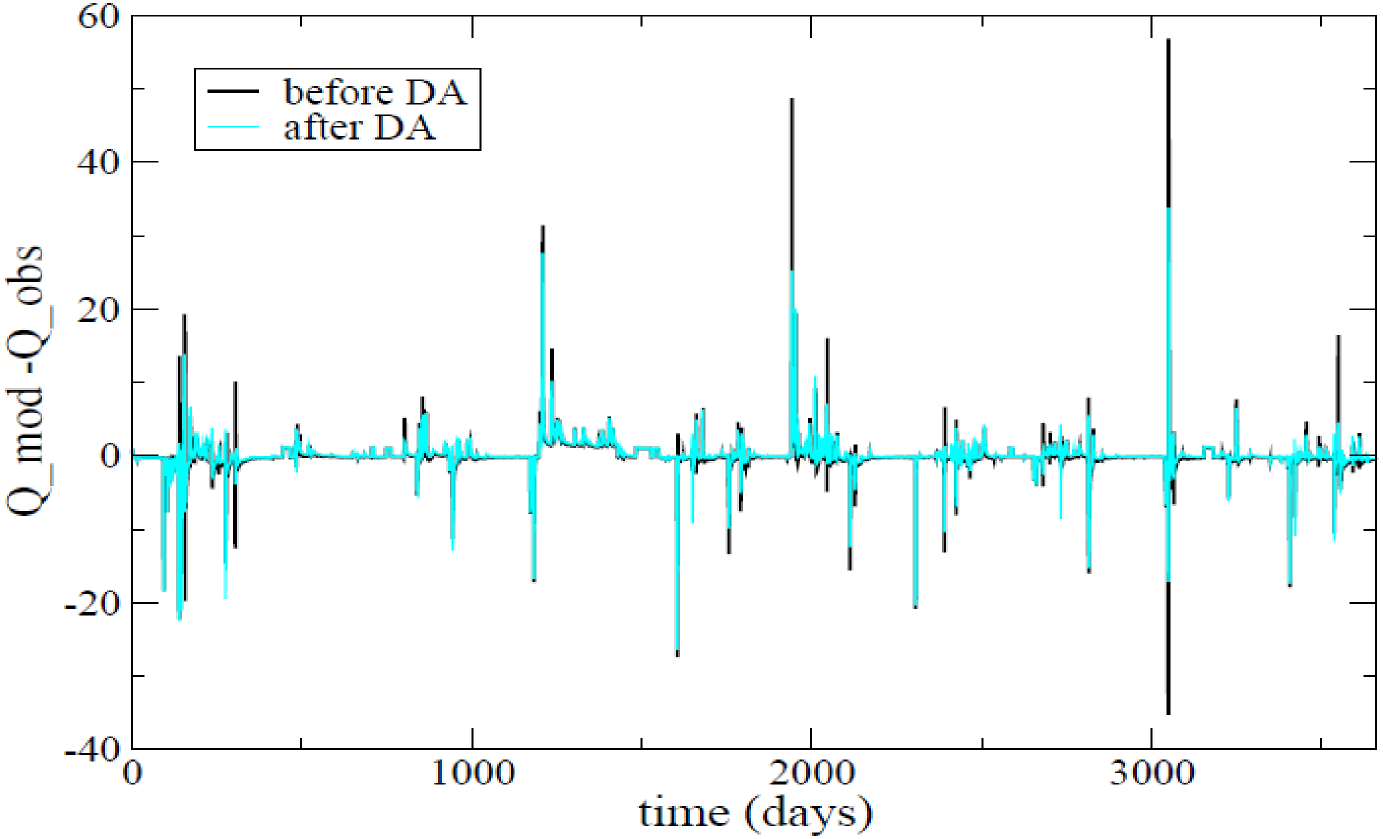
Types of estimation problems:

- parameter calibration;
- initialization
- boundary control: $P(t), E(t)$



Variational DA in hydrology (GR4J)

Parameter calibration problem:



COST_INI = 18103.0
COST_FNL = 11534.0

Variational DA in hydrology (AIGA)

AIGA is an hourly lumped three-parameter rainfall - runoff model, distributed over catchment, represented by pixels

Inputs:

P – the rainfall depth given by discretized precipitation map

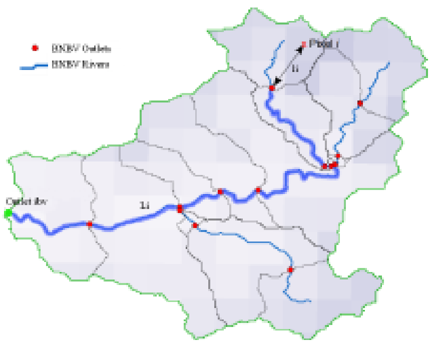
E – potential evapotranspiration (PE)

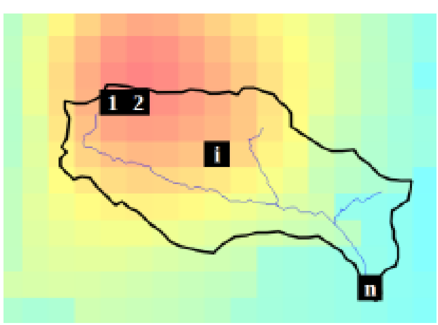
Output: *Q* – runoff

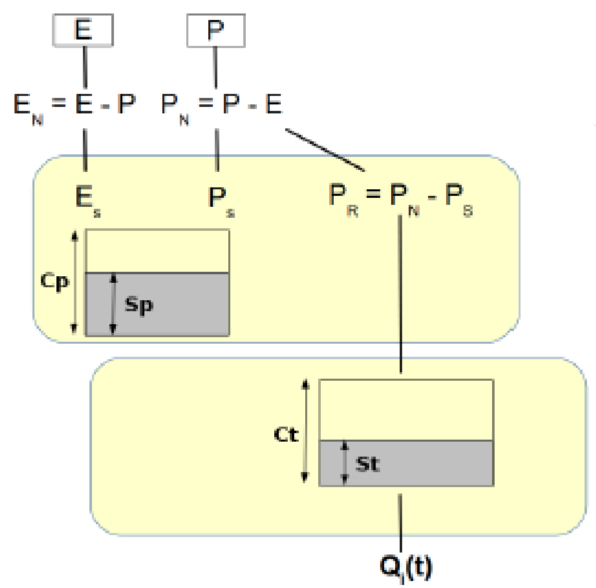
Routing scheme – $$Q_{ibv} = \sum_{i=1}^{n} q_i(t - lag) \;\; avec \;\; lag = L_{ibv} / v_{ibv}$$

Resolution - 1 km^2,

parameters: capacities of transfer and production reservoirs, transport speed

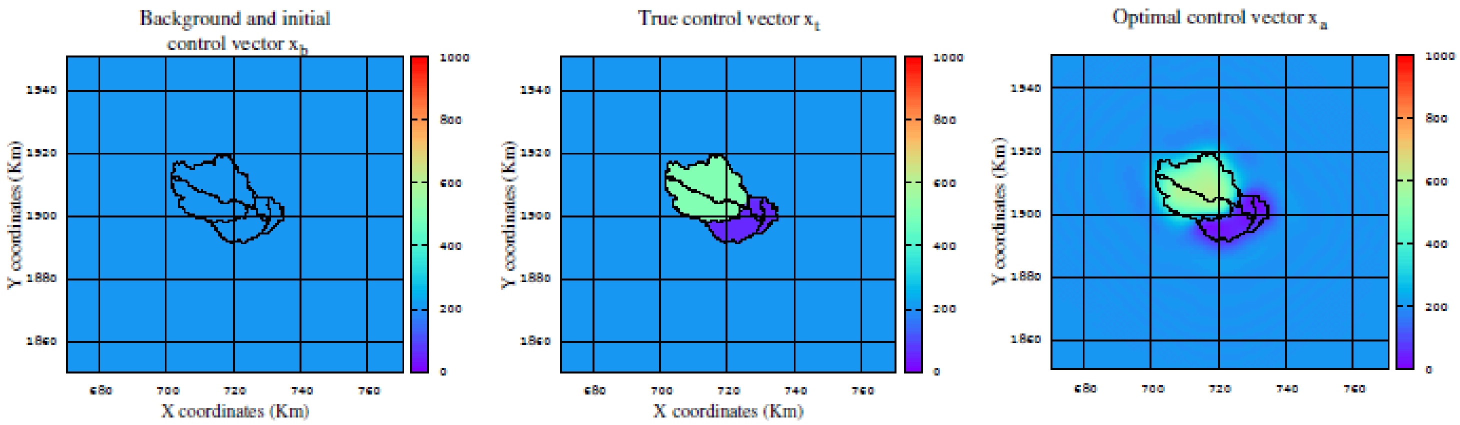






Variational DA in hydrology (AIGA)

Illustration:



Subject of the PHD thesis started at the end of 2017

Thesis title:
Data assimilation applied to a distributed hydrological model: regional calibration and assimilation of flows observed in the AIGA method

PhD student: Maxime Jay-Allemand

The thesis will be co-supervised by Patrick Arnaud (Irstea, UR RECOVER, Aix-en-Provence)

Future plans

Another PhD funding anticipated

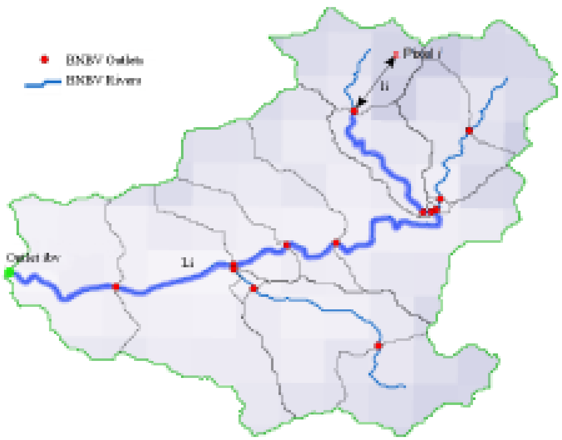
Thesis title:

Assimilation de données pour améliorer les modèles de qualité de l'eau : vers un indicateur de pression azoté

2018-2021 in collaboration with Irstea Antony

Global sensitivity and functional uncertainty analysis using numerical derivatives
(LEFE-MANU project, in collaboration with Victor Shutyaev)
Important results obtained, to be presented in Clermont-Ferrand

General direction of the group G_EAU:
development of integrated hydraulic-hydrology model capable of assimilating different types of data, including satellite-born, radar, in-situ (gauge if any), etc.




Thank you !
OLYMPIC ATHLETE
FROM RUSSIA

Question: – what to control and how?

Explicit control:
inclusion certain input variables into the active control set.

Implicit (idle) control:
a way of taking into account uncertainty in certain input variables without considering them as active controls: idea related to the 'nuisance parameter' concept, common in classical and Bayesian statistics.

Nuisance parameters – inputs which affect the design function (QoI) indirectly, i.e. via estimates of other inputs, otherwise are out of interest.
Examples:
- discharge estimation under uncertainty in bathymetry and bed roughness;
- heat flux at the ocean surface in the presence of model error.

Counterexample:
- forecasting.

Implicit treatment of uncertainty in nuisance parameters is generally achieved by modifying the likelihood function. In Gaussian case this can be done by inflating the observation covariance matrix.

Inverse Hessian by multigrid approach

SIAM J. SCI. COMPUT.
Vol. 0, No. 0, pp. 000–000

A MULTILEVEL APPROACH FOR COMPUTING THE LIMITED-MEMORY HESSIAN AND ITS INVERSE IN VARIATIONAL DATA ASSIMILATION*

KIRSTY L. BROWN†, IGOR GEJADZE‡, AND ALISON RAMAGE†

Abstract. Use of data assimilation techniques is becoming increasingly common across many application areas. The inverse Hessian (and its square root) plays an important role in several different aspects of these processes. In geophysical and engineering applications, the Hessian-vector product is typically defined by sequential solution of a tangent linear and adjoint problem; for the inverse Hessian, however, no such definition is possible. Frequently, the requirement to work in a matrix-free environment means that compact representation schemes are employed. In this paper, we propose an enhanced approach based on a new algorithm for constructing a multilevel eigenvalue decomposition of a given operator, which results in a much more efficient compact representation of the inverse Hessian (and its square root). After introducing these multilevel approximations, we investigate their accuracy and demonstrate their efficiency (in terms of reducing memory requirements and/or computational time) using the example of preconditioning a Gauss–Newton minimization procedure.

Key words. data assimilation, inverse Hessian, limited memory, preconditioning, multigrid

AMS subject classifications. 65K05, 65K10, 15A09, 15A29

DOI. 10.1137/15M1041407

Inverse Hessian by multigrid approach - 1

Consider symmetric operator in the limited-memory form:

$$A^{\beta} \cdot v = I \cdot v + \sum_{k=1}^{K} (\lambda_k^{\beta} - 1) W_k (W_k)^* \cdot v$$

Basic idea:

1. represent operator on the coarsest grid level

2. use available local preconditioner to improve its eigen-spectrum

3. build a limited-memory approximation to its inverse, which forms the basis for the local preconditioner at the next finer level

4. move up one grid level and repeat

Inverse Hessian by multigrid approach - 2

Multigrid structure:

$$N_e = (n_0, n_1, \ldots, n_{k_c})$$

$$[\Lambda, \mathcal{W}] = MLEVD(A_0, N_e)$$

Output:

$$\Lambda = \left[\lambda_{k_c}^1, \ldots, \lambda_{k_c}^{n_{k_c}}, \lambda_{k_c-1}^1, \ldots, \lambda_{k_c-1}^{n_{k_c-1}}, \ldots, \lambda_0^1, \ldots, \lambda_0^{n_0}\right],$$

$$\mathcal{W} = \left[W_{k_c}^1, \ldots, W_{k_c}^{n_{k_c}}, W_{k_c-1}^1, \ldots, W_{k_c-1}^{n_{k_c-1}}, \ldots, W_0^1, \ldots, W_0^{n_0}\right]$$

Size of eigenvectors:

$$m_k = m/2^k + 1$$

Inverse Hessian by multigrid approach - 2

Consider symmetric operator in the limited-memory form:

$$A^\beta \cdot v = I \cdot v + \sum_{k=1}^{K} (\lambda_k^\beta - 1) W_k (W_k)^* \cdot v$$

Prolongation operator:

$$v_{k-i} = S_{k,k-i}\, v_k, \qquad S_{k,k} = I_k$$

Restriction operator:

$$\tilde{v}_k = S^*_{k,k-i}\, v_{k-i}, \qquad S^*_{k,k} = I_k$$

Projection of A_k at a finer grid level $k-i$, $0 < i \le k$

$$\begin{aligned} P_{k-i}(A_k) &= S_{k,k-i}(A_k - I_k) S^*_{k,k-i} + I_{k-i}, \qquad 0 < i \le k \\ P_k(A_k) &= A_k, \end{aligned}$$

Projection of A_{k-i} at a coarser grid level k, $0 < i \le k_c$

$$\begin{aligned} Q_k(A_{k-i}) &= S^*_{k,k-i}(A_{k-i} - I_{k-i}) S_{k,k-i} + I_k, \qquad 0 < i \le k_c \\ Q_k(A_k) &= A_k. \end{aligned}$$

Inverse Hessian by multigrid approach - 3

Multigrid structure:

$$N_e = (n_0, n_1, \ldots, n_{k_c})$$

Multilevel Eigenvalue Decomposition (MLEVD) Algorithm

$$[\Lambda, \mathcal{W}] = MLEVD(A_0, N_e)$$

for $k = k_c, k_c - 1, \ldots, 0$

compute by the Lanczos method and store in memory:

$\{\lambda_k^i, U_k^i\}$, $i = 1, \ldots, n_k$ of $\tilde{Q}_k(A_0)$ in (4--7)

using $T_{k,k+1}$ and $T^*_{k,k+1}$ from (4--11)

end

$$\tilde{Q}_k(A_0) = T^*_{k,k+1} Q_k(A_0) T_{k,k+1}. \qquad (4\text{–}7)$$

$$T_{k,k+1} = \begin{cases} P_k(T_{k+1,k+2}\hat{Q}_{k+1}^{-1/2}(A_0)) & k = 0, 1, \ldots, k_c - 1; \\ I_{k_c}, & k = k_c; \end{cases} \qquad (4\text{–}11a)$$

$$T^*_{k,k+1} = \begin{cases} P_k(\hat{Q}_{k+1}^{-1/2}(A_0) T^*_{k+1,k+2}), & k = 0, 1, \ldots, k_c - 1; \\ I_{k_c}, & k = k_c; \end{cases} \qquad (4\text{–}11b)$$

$$\hat{Q}_k^{-1/2}(A_0) = I_k + \sum_{i=1}^{n_k} ((\lambda_k^i)^{-1/2} - 1) W_k^i (W_k^i)^*$$

Inverse Hessian by multigrid approach - 4

Multigrid structure:

$$N_e = (n_0, n_1, \ldots, n_{k_c})$$

$$[\Lambda, \mathcal{W}] = MLEVD(A_0, N_e)$$

Output:

$$\Lambda = \left[\lambda_{k_c}^1, \ldots, \lambda_{k_c}^{n_{k_c}}, \lambda_{k_c-1}^1, \ldots, \lambda_{k_c-1}^{n_{k_c-1}}, \ldots, \lambda_0^1, \ldots, \lambda_0^{n_0}\right],$$

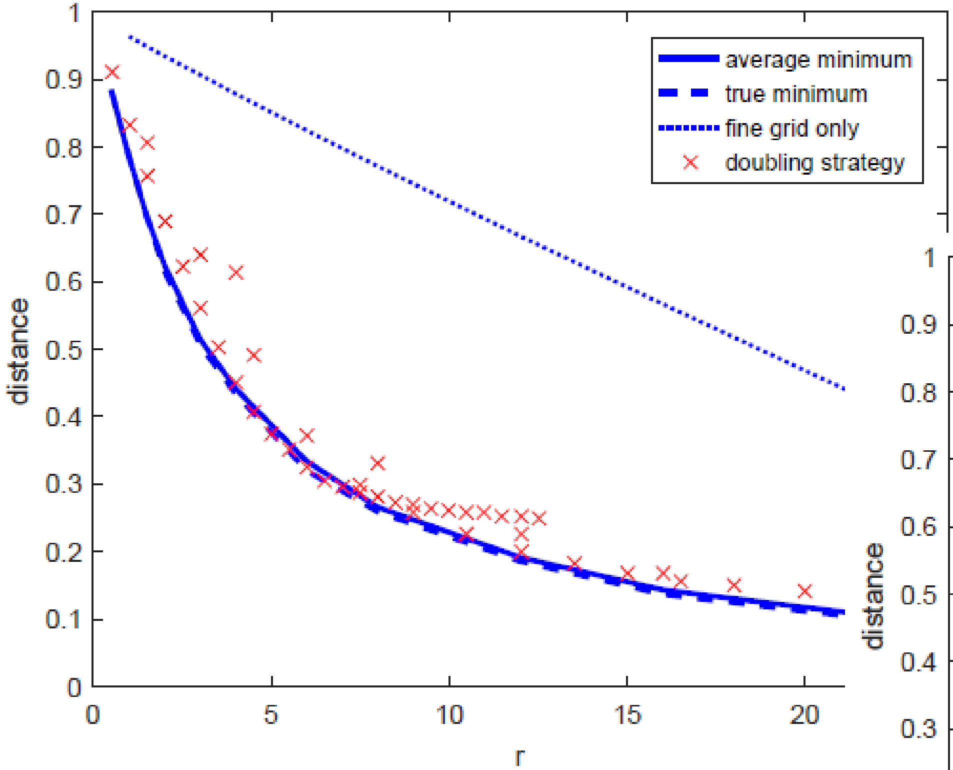
$$\mathcal{W} = \left[W_{k_c}^1, \ldots, W_{k_c}^{n_{k_c}}, W_{k_c-1}^1, \ldots, W_{k_c-1}^{n_{k_c-1}}, \ldots, W_0^1, \ldots, W_0^{n_0}\right]$$

Size of eigenvectors:

$$m_k = m/2^k + 1$$

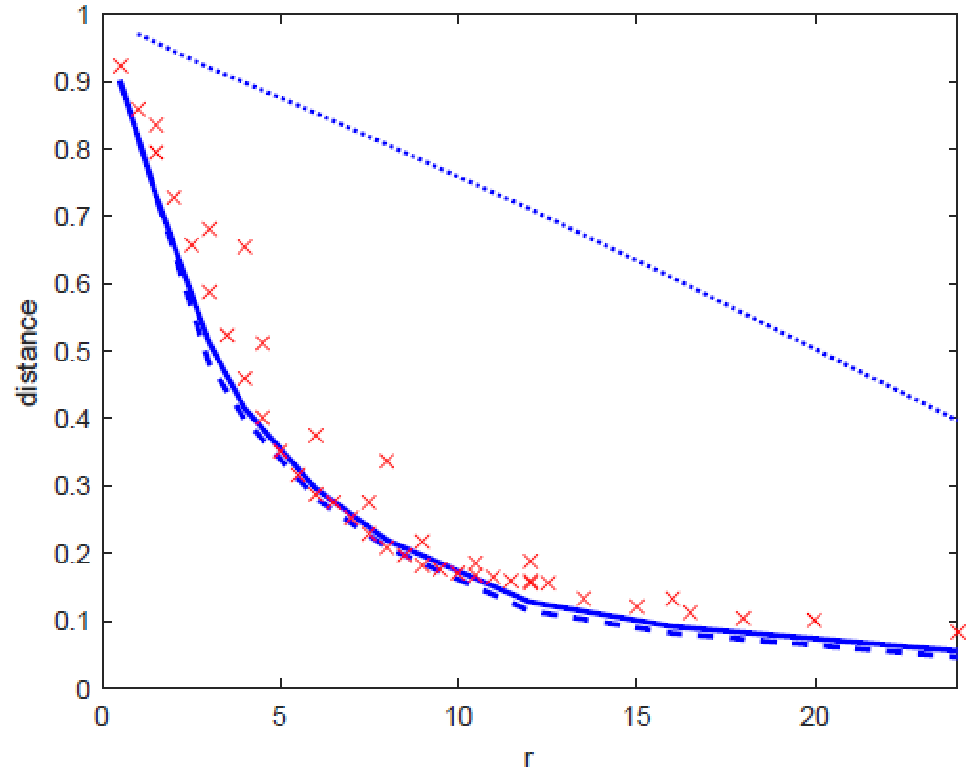
Inverse Hessian by multigrid approach - 5

Illustration: super-compact storage



Memory ratio:

$$r = \sum_{k=0}^{k_c} \frac{n_k}{2^k}$$

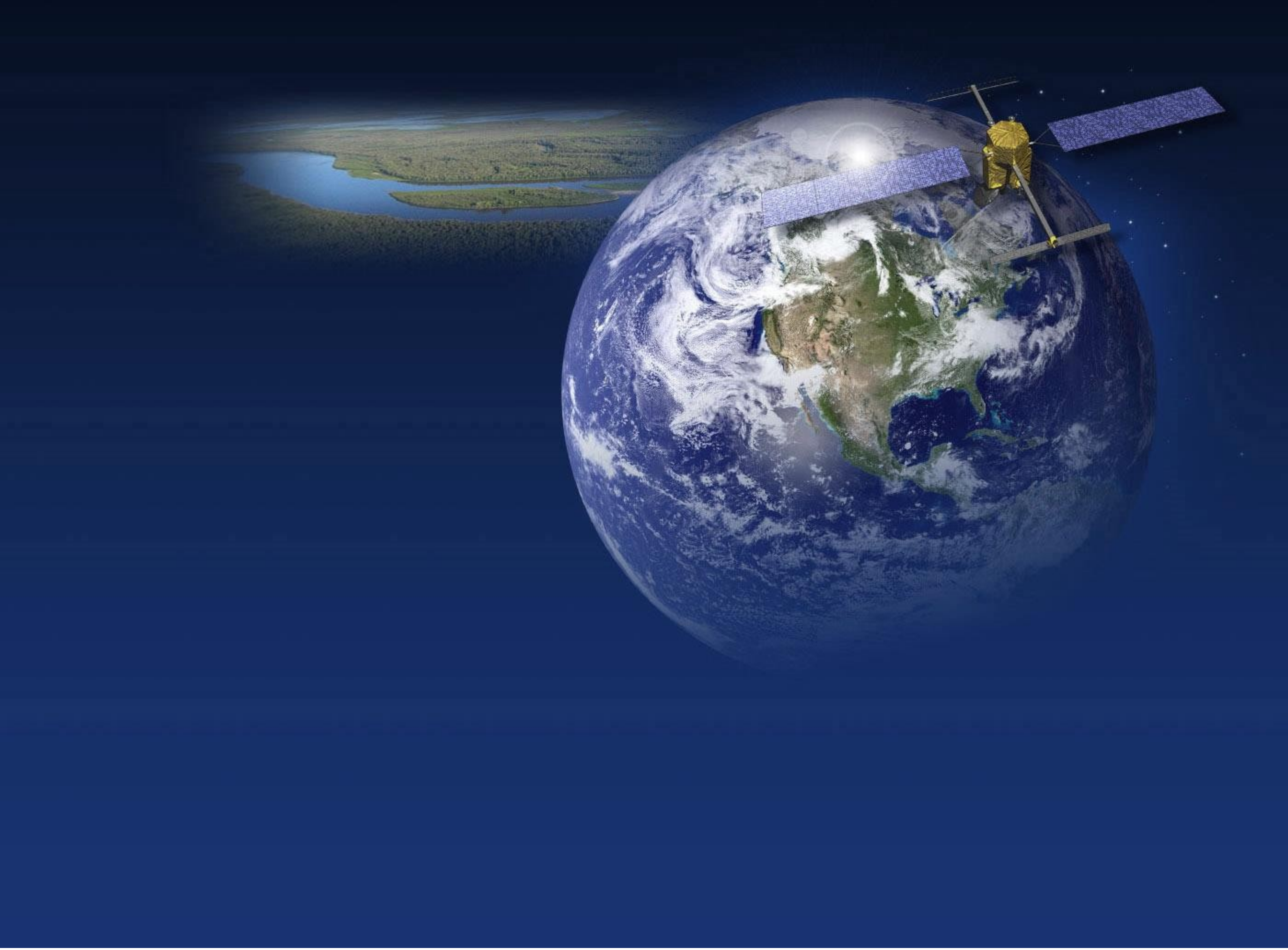


Inverse Hessian by multigrid approach - remarks

1. 'Multilevel eigenvalue decomposition' is a new concept in linear algebra
2. This decomposition can be used for super-compact representation of symmetric operators (inverses, square-roots) arising in discretization of partial differential equations: Hessian, Schrodinger operator, Laplace
3. Multilevel eigenvalue decomposition can be used for observation space decomposition in Gauss-Newton and Newton solvers, thus enables direct parallelization of control problems

Outline

1. Education and background
2. General introduction
3. Analysis error/posterior covariance
4. Design problems
5. Advanced numerical methods
6. **Current and future work**



Implicit control - motivation

Explicit control:
inclusion certain input variables into the active control set.

Implicit (idle) control:
a way of taking into account uncertainty in certain input variables without considering them as active controls: idea related to the 'nuisance parameter' concept, common in classical and Bayesian statistics.

Nuisance parameters – inputs which affect the design function (QoI) indirectly, i.e. via estimates of other inputs, otherwise are out of interest.
Examples:
- discharge estimation problem under uncertainty in bathymetry and bed roughness;
- heat flux at the ocean surface in the presence of model error.
Counterexample:
- forecasting.

Implicit treatment of uncertainty in nuisance parameters is generally achieved by modifying the likelihood function. In Gaussian case this can be done by inflating the observation covariance matrix.

Implicit control - notations

$$J(U) = \frac{1}{2}\|R^{-1/2}(G(U) - Y^*)\|^2_{\mathcal{Y}} + \frac{1}{2}\|B^{-1/2}(U - U^*)\|^2_{\mathcal{U}} \to \inf_{U}$$

- cost-function for U

$U_a \in A$ - active subset of the full control vector U

$U_p = U \setminus U_a$ - passive subset of U

$$J(U_a) = \frac{1}{2}\|R^{-1/2}(G(U_a, U_p^*) - Y^*)\|^2_{\mathcal{Y}} + \frac{1}{2}\|B_a^{-1/2}(U_a - U_a^*)\|^2_{A} \to \inf_{U_a}$$

- cost-function for U_a

Introduce $U_q = U \setminus U_a$ - 'idle' subset of U

$\bar{U}_q$, U_q^* - 'true'/exact and background for U_q

$\varepsilon_a = U_a^* - \bar{U}_a$, $\varepsilon_q = U_q^* - \bar{U}_q$ - background errors

Assume as before $E[\varepsilon_a \varepsilon_q^T] = 0$, i.e. B is block-diagonal with B_a and B_q

Important: $E[\varepsilon_q] = 0$!

Implicit control – formulation and main result - 1

Rewrite $J(U)$ as follows

$$J(U_a, U_q) = \frac{1}{2}\|R^{-1/2}(G(U_a, U_q) - Y^*)\|^2_{\mathcal{Y}}$$
$$+\frac{1}{2}\|B_a^{-1/2}(U_a - U_a^*)\|^2_{\mathcal{U}} + \frac{1}{2}\|B_q^{-1/2}(U_q - U_q^*)\|^2_{\mathcal{U}} \to \inf_{U_a, U_q}$$

Consider a modified cost-function

$$J(U_a|U_q) = \frac{1}{2}\|{R_g}^{-1/2}(G(U_a, U_q^*) - Y^*)\|^2_{\mathcal{Y}} + \frac{1}{2}\|B_a^{-1/2}(U_a - U_a^*)\|^2_{A} \to \inf_{U_a}$$

$$\Updownarrow$$

$$R_g = R + G'_{U_q}(\bar{U})B_q G'^{*}_{U_q}(\bar{U})$$

Theorem:

The estimators associated to cost-functions $J(U_a^{(1)}, U_q)$ and $J(U_a^{(2)}|U_q)$ are equivalent in the following sense:
for a linear mapping *G* the optimal solution errors $\delta U_a^{(1)}$ and $\delta U_a^{(2)}$ are identical; thus the solutions $\hat{U}_a^{(1)}$ and $\hat{U}_a^{(2)}$ are identical;
For a non-linear *G* these solutions match approximately, if the tangent linear hypothesis is valid.

Implicit control – formulation and main result - 2

Corollary. Under conditions of the Theorem,

$$E(\delta U_a \delta U_a^T) \approx E(\delta \hat{U}_a \delta \hat{U}_a^T) = {H_g}^{-1}(\bar{U})$$

where

$$H_g = B_a^{-1} + G'^{*}_{U_a}(\bar{U}){R_g}^{-1} G'_{U_a}(\bar{U})$$

In practice, $\bar{U}$ is not known. We use $U^* = (U_a^*, U_q^*)^T$ or $\tilde{U} = (\hat{U}_a, U_q^*)^T$

Computing ${R_g}^{-1}v$:

$$\tilde{R}_g v = R^{-1/2} R_g R^{-1/2} v = v + R^{-1/2} G'_{U_q}(\bar{U}) B_q G'^{*}_{U_q}(\bar{U}) R^{-1/2} v, \; v \in \mathcal{Y}$$

$${R_g}^{-1} v = R^{-1/2} \left(I + \sum_{i=1}^{N_1} (\beta_i^{-1} - 1) z_i z_i^T \right) R^{-1/2} v,$$

where $\{\beta_i, z_i\}$, $i = 1, \ldots, N_1$ are leading eigenpairs of $\tilde{R}_g$

$\delta U = (\delta U_a, \varepsilon_q)^T$ - posterior error in the input vector

$$\delta U_a = {H_g}^{-1}(\bar{U}) \left(G'^{*}_{U_a}(\bar{U}) {R_g}^{-1} \xi + B_a^{-1} \varepsilon_a - G'^{*}_{U_a}(\bar{U}) {R_g}^{-1} G'_{U_q}(\bar{U}) \varepsilon_q \right)$$

Implicit control – numerical model

$$\frac{\partial \varphi}{\partial t} + (a + b\varphi)\frac{\partial \varphi}{\partial x} = \frac{\partial}{\partial x}\left(d(\varphi)\frac{\partial \varphi}{\partial x}\right) + c\varphi + \eta$$

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

- 1D generalised Burgers' equation

$\varphi(x,0) = u(x)$ - initial condition

$(d\varphi/dx)|_{x=0} = (d\varphi/dx)|_{x=1} = 0$ - Neumann boundary conditions

$d(\varphi) = d_0 + d_1\,(d\varphi/dx)^2,\ \ d_0, d_1 = const > 0$ - viscosity coefficient

$a(x,t)$ - advection uncertainty, $a(x,t) \sim \mathcal{N}(0, B_q)$

The initial condition $u(x)$ is the quantity of interest, thus included into the active control set, whereas the advection uncertainty $a(x,t)$ is subjected to implicit treatment as a nuisance parameter !

Two cases:
Bylinear case: $b = 0,\ d_1 = 0$, nonlinear - otherwise

Implicit control – conditions of numerical tests

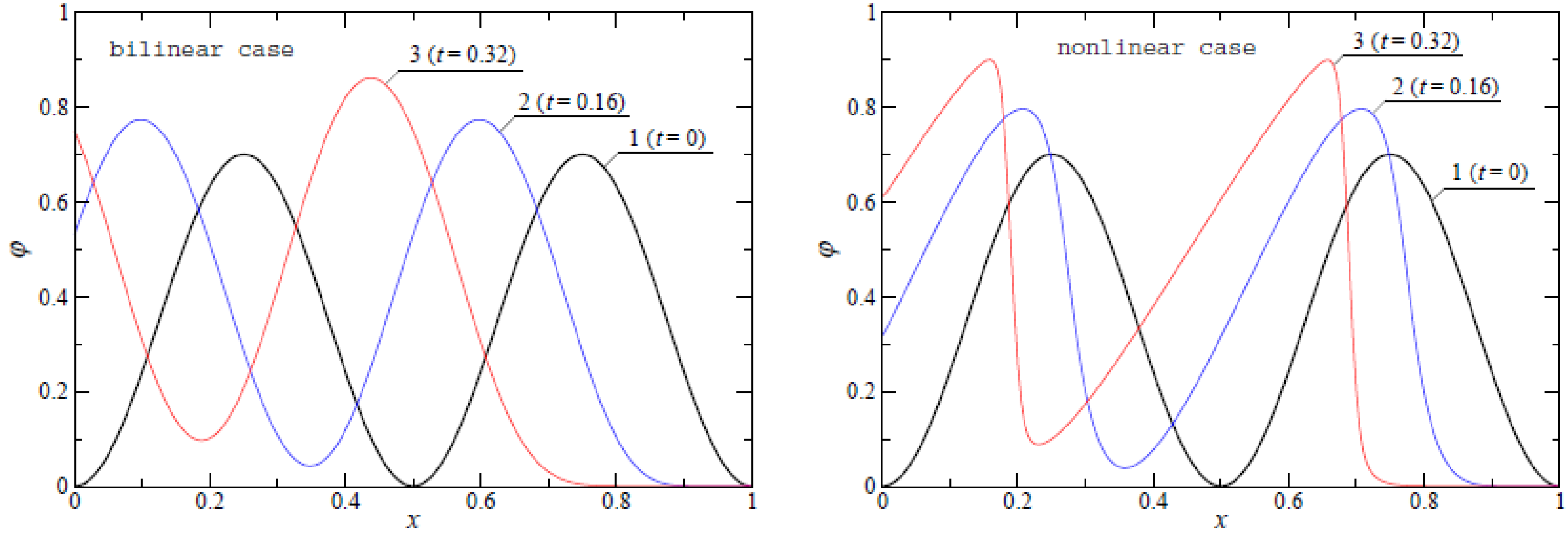


Figure 1. Initial and evolved states: *bilinear* case (left) and *nonlinear* case (right).

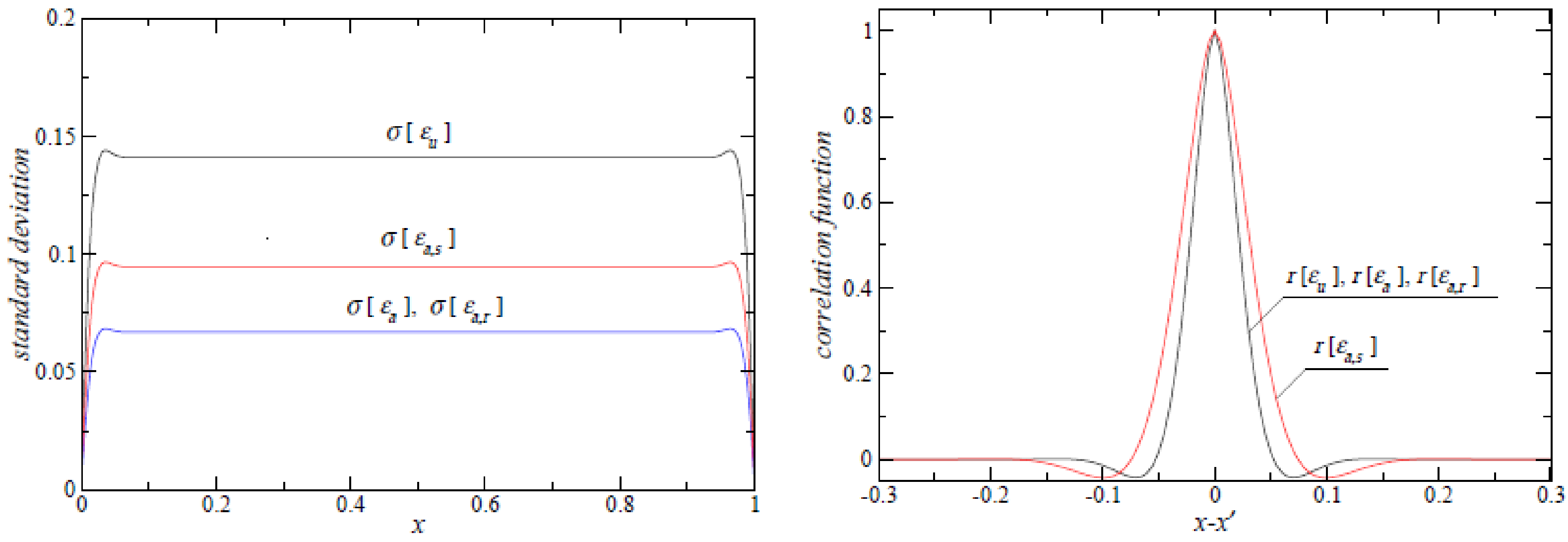


Figure 2. Background error standard deviation (left), and correlation functions (right).

Implicit control – inflated observation covariance

$$\frac{\partial \varphi}{\partial t} + (a + b\varphi)\frac{\partial \varphi}{\partial x} = \frac{\partial}{\partial x}\left(d(\varphi)\frac{\partial \varphi}{\partial x}\right) + c\varphi + \eta$$

$\bar{w}^j = \mathcal{T}(w^j, \bar{w}^{j-1}) = k_t \bar{w}^{j-1} + (1 - k_t) w^j$ - filter to introduce temporal correlations

$k_t = 1$ - full temporal correlation - $a(x,t) \equiv a(x)$

$k_t = 0$ - no temporal correlation

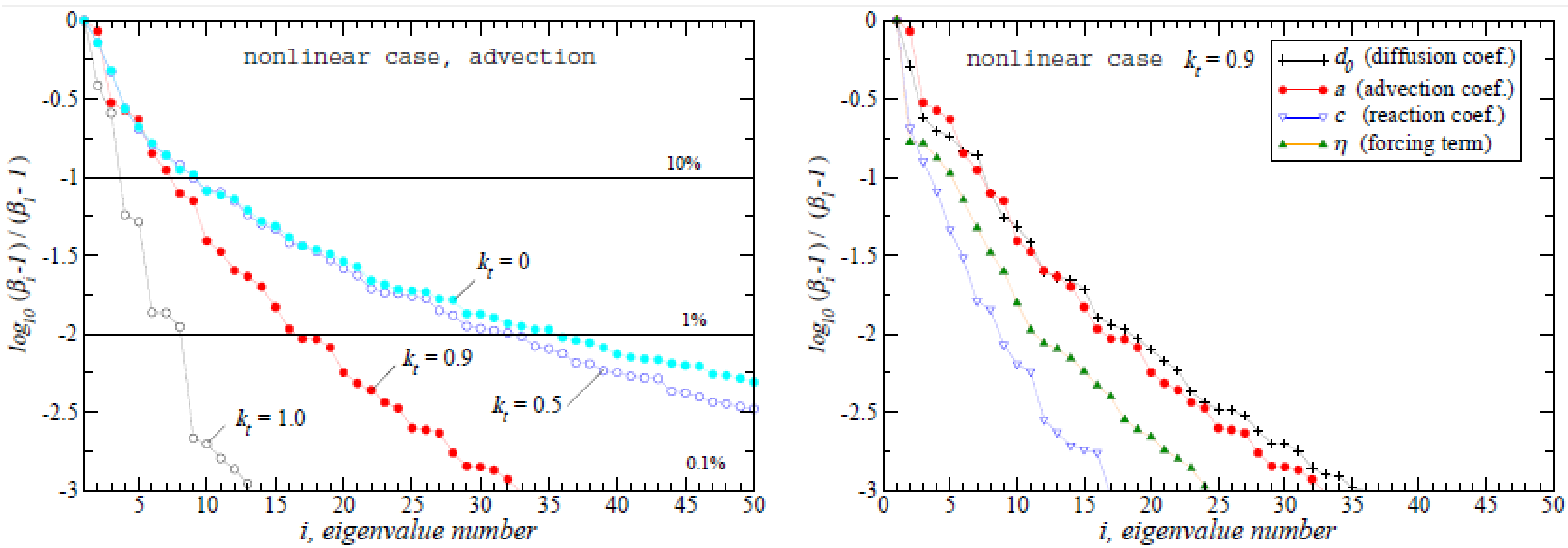


Figure 3. Scaled eigenvalues of $\bar{R}_g$ for different k_t (left), and for different parameters (right). Case A/*nonlinear*.

$$R_g^{-1} v = R^{-1/2}\left(I + \sum_{i=1}^{N_1}(\beta_i^{-1} - 1) z_i z_i^T\right) R^{-1/2} v,$$

Implicit control – numerical results (error)

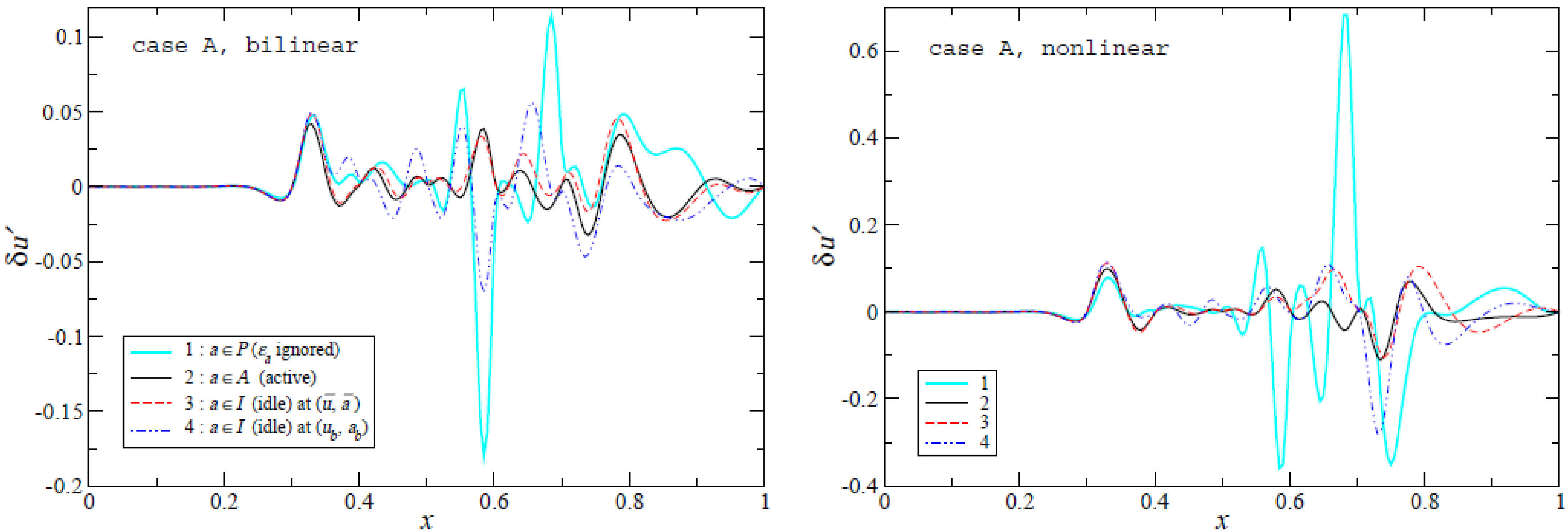


Figure 4. Estimation error $\delta u'$ by different methods for case A/*bilinear* (left), and case A/*nonlinear* (right).

Implicit control – numerical results (st. dev.)

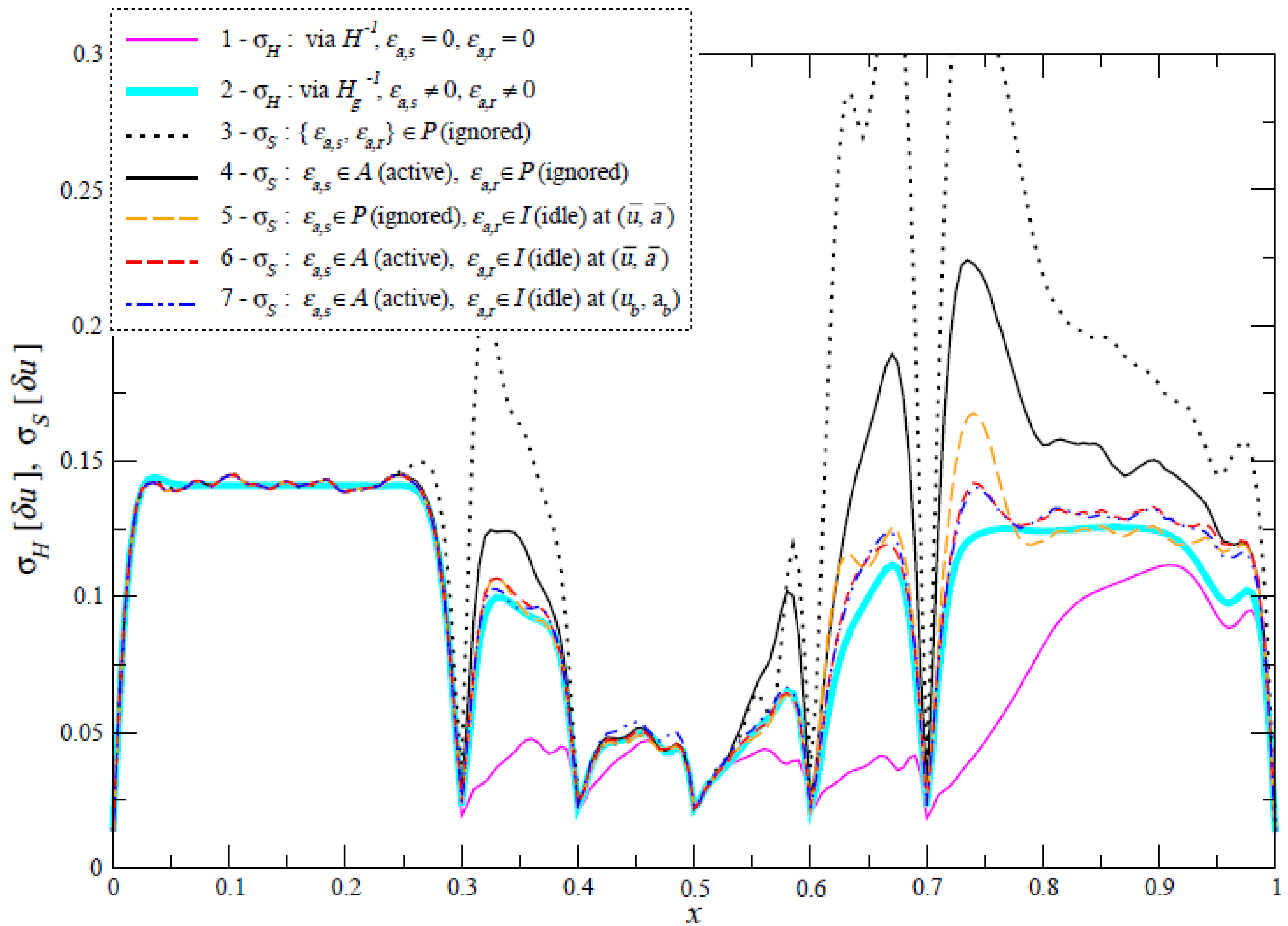


Figure 6. Standard deviation $\sigma_S[\delta u]$ and $\sigma_H[\delta u]$ by different methods for case B/*nonlinear*

Implicit control – numerical results (discharge)

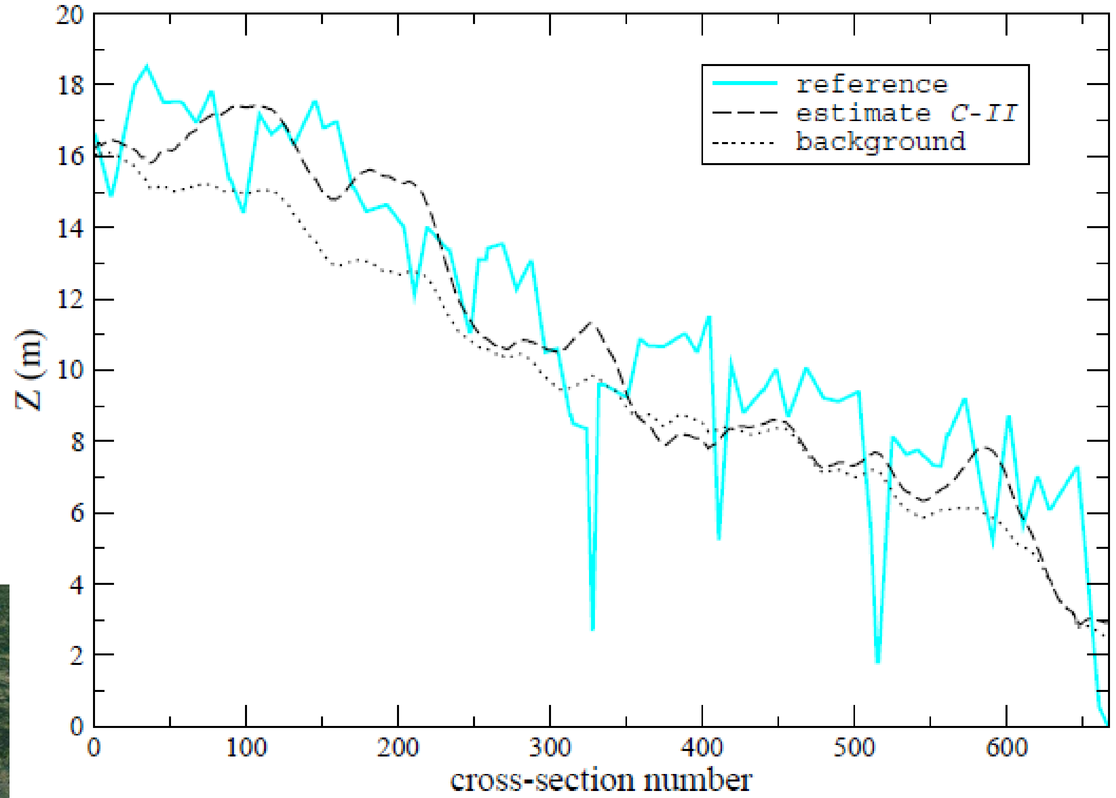


Figure 11. Garonne bed elevation.

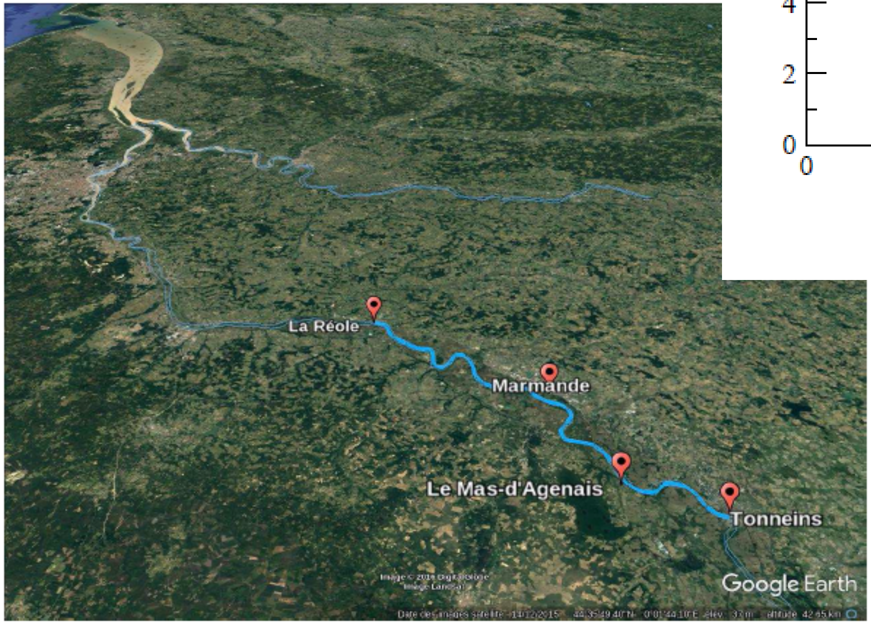


Figure 9. Garonne river, downstreem part.

Discharge estimation:
the Garonne river at Tonneins, 2010

Implicit control – numerical results (discharge)

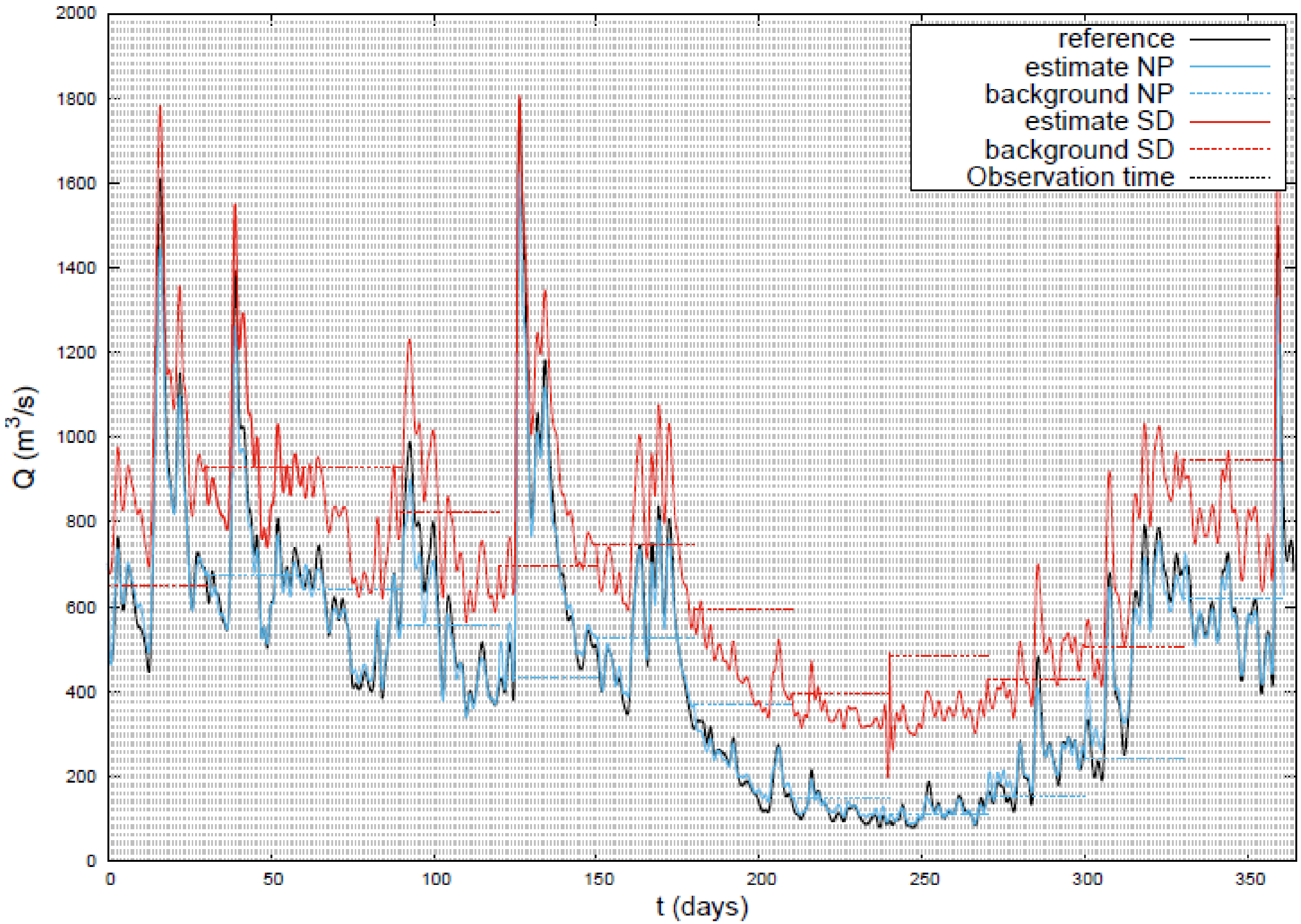


Figure 10. Discharge at Tonneins, 2010.

Implicit control – applicability to forecasting

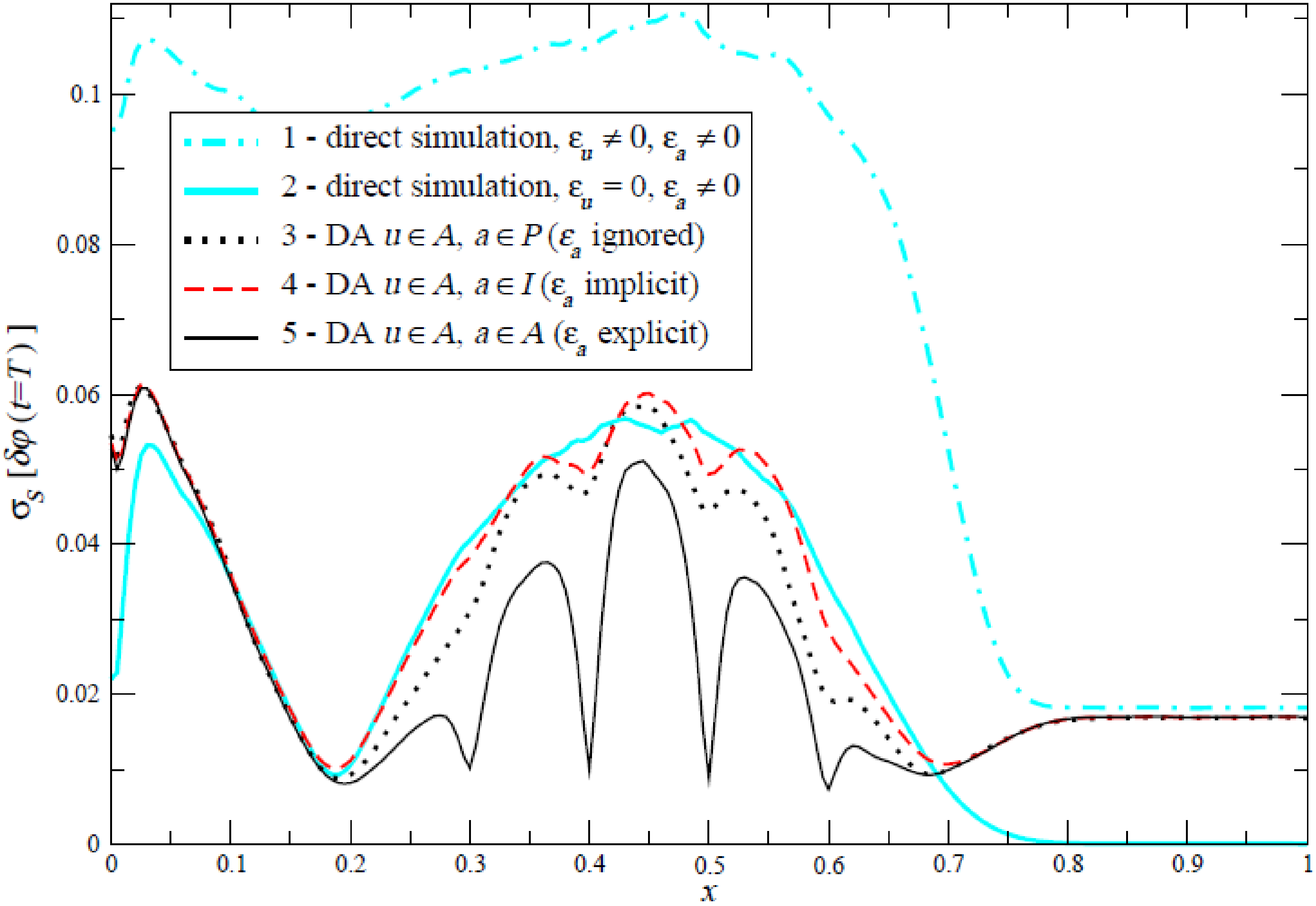


Figure 8. Forecast $(t = T)$ standard deviation for case A/*bilinear*.

Implicit control – biased uncertainty

Case of a biased error $\varepsilon_q \sim \mathcal{N}(\varepsilon_{q,s}, B_q)$

Then, it can be represented as $\varepsilon_q = \varepsilon_{q,s} + \varepsilon_{q,r}$,
where $\varepsilon_{q,s}$ and $\varepsilon_{q,r} \sim \mathcal{N}(0, B_q)$ are respectively the systematic and random components of ε_q.

One way to deal with a biased uncertainty would be to include both components into the control vector and consider minimizing the following cost-function:

$$J(U_a, \varepsilon_{q,s}, \varepsilon_{q,r}) = \frac{1}{2}\|B_a^{-1/2}(U_a - U_a^*)\|_{\mathcal{U}}^2 + \frac{\alpha}{2}\|W\varepsilon_{q,s}\|_{\mathcal{U}}^2$$
$$+\frac{1}{2}\|B_q^{-1/2}\varepsilon_{q,r}\|_{\mathcal{U}}^2 + \frac{1}{2}\|R^{-1/2}(G(U_a, U_q^* - \varepsilon_{q,s} - \varepsilon_{q,r}) - Y^*)\|_{\mathcal{Y}}^2 \to \inf_{U_a,\, \varepsilon_{q,s}, \varepsilon_{q,r}}$$

An alternative way is to include only the systematic part:

$$J(U_a, \varepsilon_{q,s}) = \frac{1}{2}\|B_a^{-1/2}(U_a - U_a^*)\|_{\mathcal{U}}^2 + \frac{\alpha}{2}\|W\varepsilon_{q,s}\|_{\mathcal{U}}^2$$
$$+\frac{1}{2}\|R_g{}^{-1/2}(G(U_a, U_q^* - \varepsilon_{q,s}) - Y^*)\|_{\mathcal{Y}}^2 \to \inf_{U_a,\, \varepsilon_{q,s}}$$

$$R_g = R + G'_{U_{q,r}}(\bar{U})B_q G'^*_{U_{q,r}}(\bar{U})$$

Implicit control – numerical results (st. dev.)

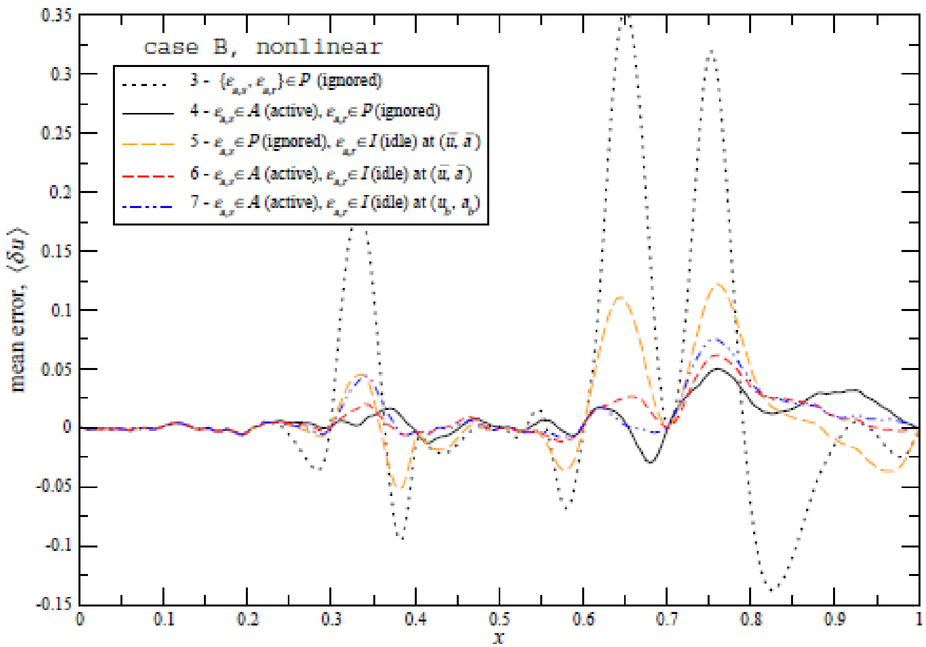
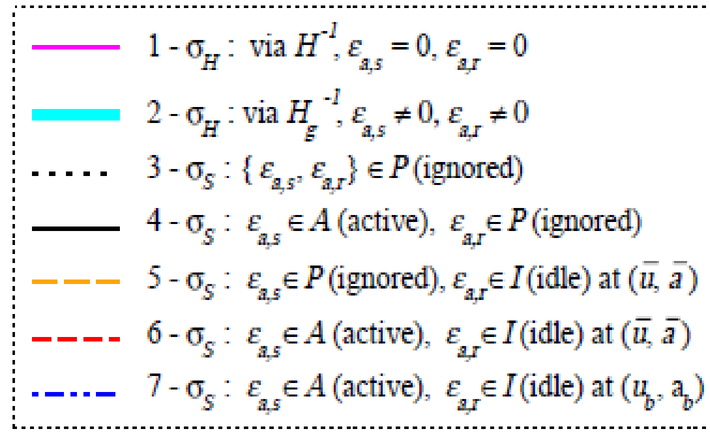


Figure 7. Mean error by different methods for case B/*nonlinear*.



If uncertainty bias is not included into the active control set, the QoI estimate is also biased ! However, this does not significantly affect its variance !

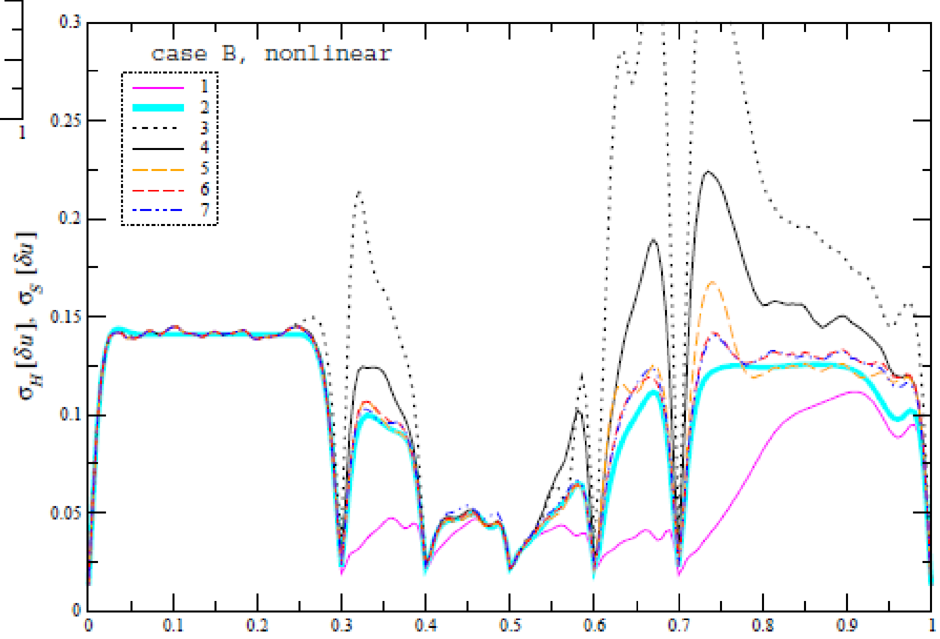


Figure 6. Standard deviation $\sigma_S[\delta u]$ and $\sigma_H[\delta u]$ by different methods case B/*nonlinear*

Conclusions on the implicit control method

- influence of uncertainties in certain variables on the estimates of other variables can be eliminated/reduced by considering the former as 'idle' controls, which implies the implicit treatment via the inflated observation covariance (in essence, the modified likelihood)

- the difficulties with the control vector extension method include oversizing, solvability and robustness. The suggested method allows us to alleviate these difficulties, however the method would only be useful if the active and idle control sets are properly defined. Here we must use the control set design approach suggested at the beginning

- the proposed method is feasible for high-dimensional models since the inflated observation covariance is represented by a relatively small set of its largest eigenpairs obtained by means of the Lanczos algorithm. For mildly nonlinear problem this covariance has to be computed only once

- so far, the method is primarily suited to the case when the quantities of interest coincide with the active controls or largely dominated by them. As it stands, the method is not useful for forecasting

Future work

- The suggested control set design method can be generalized to the case when the full input simultaneously contains **active**, **passive** and **idle** (nuisance) controls

- The method can be generalized to include **integrated controls**, i.e. control inputs which are not originally presented in the model. This allows us to assess the performance of existing DA methods, such as optimal nudging or sub-window technique

- A few developments to analyse **global** rather than local design functions can be suggested

- Important: the presented methodology can be used as a basis for a general **control space decomposition approach** (work in progress), which might eventually lead to a better forecasting algorithm