

# ON ANALYSIS ERROR COVARIANCES IN VARIATIONAL DATA ASSIMILATION

I. GEJADZE<sup>†</sup>, F.-X. LE DIMET<sup>‡</sup>, AND V. SHUTYAEV<sup>§</sup>

**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.** ?

**1. Introduction.** The methods of variational data assimilation (DA) were designed to combine models and observational data as sources of information. From the mathematical point of view, DA problems can be formulated as optimal control problems (e.g. [16, 4]). A major advantage of this approach is the derivation of an optimality system which contains all the available information. In practice the optimality system includes background and observation errors, hence the error in optimal solution (analysis error). It is an important practical issue to estimate properties of this error. In this paper we consider a nonlinear evolution model with an unknown initial condition, which must be retrieved using incomplete observations of state evolution. The analysis can be used as a basis to define the background for the next DA cycle or it might be of interest by its own (hind-cast).

The analysis error may be derived through the errors of the input data using the Hessian of the associated DA problem. For a deterministic case this was achieved in [17]. If the errors of the input data are random and normally distributed, then for a linear evolution model, the analysis error covariance matrix is given by the inverse Hessian of the cost functional (see e.g. [29]). For a nonlinear problem no exact relationship exists. In practice it is usually assumed that the evolution of errors can be satisfactorily described by the tangent linear model (TLM). This assumption is called the tangent linear hypothesis (TLH). In this case the analysis error covariance matrix is considered to be approximately equal to the inverse Hessian of the cost functional (see e.g. [29, 14, 23, 30, 28, 11, 32]). These results are given for a discretized problem, while in this paper we prove similar results for the continuous case. Also, our consideration of the nonlinear case is different since we do not rely on the TLH from the very beginning.

Here we develop the ideas of [17] for the case of stochastic errors. We derive the

---

<sup>†</sup>Department of Civil Engineering, University of Strathclyde, John Anderson Building, 107 Rottenrow, Glasgow G4 ONG, UK ([igor.gejadze@strath.ac.uk](mailto:igor.gejadze@strath.ac.uk)). The first author was funded through the Glasgow Research Partnership by the Scottish Funding Council.

<sup>‡</sup>LJK, Université Joseph Fourier, BP 51, 38051 Grenoble Cedex 9, France ([ledimet@inria.fr](mailto:ledimet@inria.fr)). This co-author was funded through the MOISE project (CNRS, INRIA, UJF, INPG).

<sup>§</sup>Institute of Numerical Mathematics, Russian Academy of Sciences, 119991 Gubkina 8, Moscow, Russia ([shutyaev@inm.ras.ru](mailto:shutyaev@inm.ras.ru)). This co-author was funded through the project 06-01-00344 of the Russian Foundation for Basic Research and the program ECO-NET (EGIDE).

exact equation for the analysis error through the errors of the input data without the use of the TLH. This equation involves some operators which are then approximated to derive an estimate of the analysis error covariance operator. It is shown that in a nonlinear case the analysis error covariance operator can be approximated by the inverse Hessian of the auxiliary (linearized) data assimilation problem, which involves the tangent linear model constraints, and it does not coincide in general with the Hessian of the original nonlinear DA problem. This result has not been formulated in an operator form before.

We develop a numerical algorithm to compute the analysis error covariance matrix with the use of the quasi-Newton BFGS method [26, 7]. The following points are essential for this algorithm: choice of input functions, choice of initial guess in the auxiliary DA problem and the exact analytical step search for the minimization along the direction of descent. We also develop a novel, fully nonlinear ensemble method, which is the only way to estimate the actual value of the analysis error covariance matrix in the nonlinear case. We compare the estimates of this matrix obtained by the BFGS and by the ensemble method. This comparison is illustrated by numerical examples given for the 1D nonlinear convection-diffusion model. The tests are specially designed to produce strong nonlinear effects. Other examples show how the analysis error variance depends on the different transport phenomena supported by the model and on the correlation radius of the background error.

In this paper we consider the background error and the observation error, but there are other potential errors, such as 'representation' or 'model' error and an error due to discretization of the model equations. Errors of these types may also be taken into account by the technique based on the inverse Hessian. For a deterministic case the 'model' error was considered in [17].

The paper is organized as follows. In Sect.2, we provide the statement of the variational DA problem for a nonlinear evolution model to identify the initial condition. In Sect.3, the equation of the analysis error is derived through the errors of the input data using the Hessian of the auxiliary DA problem. In Sect.4, we derive the formulas for the covariance operator through the covariance operators of the input errors involving this Hessian. The numerical algorithm to compute the analysis error covariance matrix as the inverse Hessian and a fully nonlinear ensemble method are described in Sect.5. Section 6 presents the model employed (the 1D nonlinear convection-diffusion equation) and other details of numerical implementation. Numerical examples are presented in Sect.7. A particular Sub-section in Sect.7 investigates the validity of the TLH condition. The Appendix contains the definition of the Hessian of the original nonlinear data assimilation problem using the second-order adjoint technique.

From this point on we shall refer to 'analysis error covariance/variance' simply as 'covariance/variance'.

Part of the theory presented in this paper has been published in [18].

**2. Statement of the problem.** Consider the mathematical model of a physical process that is described by the evolution problem

$$(2.1) \quad \begin{cases} \frac{\partial \varphi}{\partial t} = F(\varphi) + f, & t \in (0, T) \\ \varphi|_{t=0} = u, \end{cases}$$

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  $X$  into  $X$ . Let  $Y = L_2(0, T; X)$ ,  $\|\cdot\|_Y = (\cdot, \cdot)_Y^{1/2}$ ,  $f \in Y$ . Suppose that for a given  $u \in X$ ,  $f \in Y$  there exists a unique solution  $\varphi \in Y$  to (2.1).

Let us formulate the following DA problem (optimal control problem) with the aim to identify the initial condition: find  $u \in X$  and  $\varphi \in Y$  such that they satisfy (2.1), and on the set of solutions to (2.1), a cost functional  $S(u)$  takes the minimum value, i.e.

$$(2.2) \quad \begin{cases} \frac{\partial \varphi}{\partial t} = F(\varphi) + f, & t \in (0, T) \\ \varphi|_{t=0} = u, \\ S(u) = \inf_{v \in X} S(v), \end{cases}$$

where

$$(2.3) \quad S(u) = \frac{1}{2}(V_1(u - u_0), u - u_0)_X + \frac{1}{2}(V_2(C\varphi - \varphi_{obs}), C\varphi - \varphi_{obs})_{Y_{obs}}.$$

In this formulation  $u_0 \in X$  is a prior initial-value function (background state),  $\varphi_{obs} \in Y_{obs}$  is a prescribed function (observational data),  $Y_{obs}$  is a Hilbert space (observation space),  $C : Y \rightarrow Y_{obs}$  is a linear bounded operator,  $V_1 : X \rightarrow X$  and  $V_2 : Y_{obs} \rightarrow Y_{obs}$  are symmetric positive definite operators.

The necessary optimality condition reduces the problem (2.2)-(2.3) to the following system [19], [1]:

$$(2.4) \quad \begin{cases} \frac{\partial \varphi}{\partial t} = F(\varphi) + f, & t \in (0, T) \\ \varphi|_{t=0} = u, \end{cases}$$

$$(2.5) \quad \begin{cases} -\frac{\partial \varphi^*}{\partial t} - (F'(\varphi))^* \varphi^* = -C^* V_2 (C\varphi - \varphi_{obs}), & t \in (0, T) \\ \varphi^*|_{t=T} = 0, \end{cases}$$

$$(2.6) \quad V_1(u - u_0) - \varphi^*|_{t=0} = 0$$

with the unknowns  $\varphi, \varphi^*, u$ , where  $(F'(\varphi))^*$  is the adjoint to the Frechet derivative of  $F$ , and  $C^*$  is the adjoint to  $C$  defined by  $(C\varphi, \psi)_{Y_{obs}} = (\varphi, C^*\psi)_Y$ ,  $\varphi \in Y, \psi \in Y_{obs}$ . We assume that the system (2.4)–(2.6) has a unique solution. We shall denote by  $\mathcal{H}$  the Hessian of the nonlinear DA problem (2.2)–(2.3). The definition of  $\mathcal{H}$  is provided in the Appendix.

Suppose that  $u_0 = \bar{u} + \xi_1$ ,  $\varphi_{obs} = C\bar{\varphi} + \xi_2$ , where  $\xi_1 \in X$ ,  $\xi_2 \in Y_{obs}$ , and  $\bar{\varphi}$  is the ("true") solution to the problem (2.1) with  $u = \bar{u}$ :

$$(2.7) \quad \begin{cases} \frac{\partial \bar{\varphi}}{\partial t} = F(\bar{\varphi}) + f, & t \in (0, T) \\ \bar{\varphi}|_{t=0} = \bar{u}. \end{cases}$$

The functions  $\xi_1, \xi_2$  may be treated as the errors of the input data  $u_0, \varphi_{obs}$  (background and observation errors, respectively). For  $V_1$  and  $V_2$  in (2.3), we consider  $V_1 = V_{\xi_1}^{-1}$ ,  $V_2 = V_{\xi_2}^{-1}$ , where  $V_{\xi_i}$  is the covariance operator of the corresponding error  $\xi_i, i = 1, 2$ .

Having assumed that the solution of the problem (2.4)–(2.6) exists, we will study the impact of the errors  $\xi_1, \xi_2$  on the optimal solution  $u$  and develop the theory presented in [17] for the case of stochastic errors.

**3. Equation for the analysis error via Hessian.** The system (2.4)–(2.6) with the three unknowns  $\varphi, \varphi^*, u$  may be treated as an operator equation of the form

$$(3.1) \quad \mathcal{F}(U, U_d) = 0,$$

where  $U = (\varphi, \varphi^*, u)$ ,  $U_d = (u_0, \varphi_{obs}, f)$ .

The following equality holds for the exact solution ('true state'):

$$(3.2) \quad \mathcal{F}(\bar{U}, \bar{U}_d) = 0,$$

with  $\bar{U} = (\bar{\varphi}, \bar{\varphi}^*, u)$ ,  $\bar{U}_d = (\bar{u}, C\bar{\varphi}, f)$ ,  $\bar{\varphi}^* = 0$ . The system (3.2) is the necessary optimality condition of the following optimal control problem: find  $u$  and  $\varphi$  such that

$$\begin{cases} \frac{\partial \varphi}{\partial t} = F(\varphi) + f, & t \in (0, T) \\ \varphi|_{t=0} = u, \\ \bar{S}(u) = \inf_v \bar{S}(v), \end{cases}$$

where

$$\bar{S}(u) = \frac{1}{2}(V_1(u - \bar{u}), u - \bar{u})_X + \frac{1}{2}(V_2(C\varphi - C\bar{\varphi}), C\varphi - C\bar{\varphi})_{Y_{obs}}.$$

From (3.1)–(3.2), we get

$$(3.3) \quad \mathcal{F}(U, U_d) - \mathcal{F}(\bar{U}, \bar{U}_d) = 0.$$

Let  $\delta U = U - \bar{U}$ ,  $\delta U_d = U_d - \bar{U}_d$ . Then (3.3) gives

$$(3.4) \quad \mathcal{F}(\bar{U} + \delta U, \bar{U}_d + \delta U_d) - \mathcal{F}(\bar{U}, \bar{U}_d) = 0.$$

Let  $\delta\varphi = \varphi - \bar{\varphi}$ ,  $\delta u = u - \bar{u}$ ; then  $\delta U = (\delta\varphi, \varphi^*, \delta u)$ ,  $\delta U_d = (\xi_1, \xi_2, 0)$ . Let us suppose that  $F$  is continuously Frechet differentiable, then there exists  $\tilde{\varphi} = \bar{\varphi} + \tau(\varphi - \bar{\varphi})$ ,  $\tau \in [0, 1]$ , such that the Taylor-Lagrange formula [21] is valid:  $F(\varphi) - F(\bar{\varphi}) = F'(\tilde{\varphi})\delta\varphi$ . Then equation (3.4) is equivalent to the system:

$$(3.5) \quad \begin{cases} \frac{\partial \delta\varphi}{\partial t} - F'(\tilde{\varphi})\delta\varphi = 0, & t \in (0, T), \\ \delta\varphi|_{t=0} = \delta u, \end{cases}$$

$$(3.6) \quad \begin{cases} -\frac{\partial \varphi^*}{\partial t} - (F'(\varphi))^* \varphi^* = -C^* V_2(C\delta\varphi - \xi_2), \\ \varphi^*|_{t=T} = 0, \end{cases}$$

$$(3.7) \quad V_1(\delta u - \xi_1) - \varphi^*|_{t=0} = 0.$$

Let us introduce the operator  $\tilde{H}$  defined by the successive solutions of the following problems:

$$(3.8) \quad \begin{cases} \frac{\partial \psi}{\partial t} - F'(\tilde{\varphi})\psi = 0, & t \in (0, T), \\ \psi|_{t=0} = v, \end{cases}$$

$$(3.9) \quad \begin{cases} -\frac{\partial \psi^*}{\partial t} - (F'(\varphi))^* \psi^* &= -C^* V_2 C \psi, \quad t \in (0, T) \\ \psi^*|_{t=T} &= 0, \end{cases}$$

$$(3.10) \quad \tilde{H}v = V_1 v - \psi^*|_{t=0}.$$

Below we introduce two auxiliary operators  $R_1, \tilde{R}_2$ . Let  $R_1 = V_1$ . Let us introduce the operator  $\tilde{R}_2 : Y_{obs} \rightarrow X$  acting on the functions  $g \in Y_{obs}$  according to the formula

$$(3.11) \quad \tilde{R}_2 g = \theta^*|_{t=0},$$

where  $\theta^*$  is the solution to the adjoint problem

$$(3.12) \quad \begin{cases} -\frac{\partial \theta^*}{\partial t} - (F'(\varphi))^* \theta^* &= C^* V_2 g, \quad t \in (0, T) \\ \theta^*|_{t=T} &= 0. \end{cases}$$

From (3.8)–(3.12) we conclude that the system (3.5)–(3.7) is equivalent to the single equation for  $\delta u$ :

$$(3.13) \quad \tilde{H}\delta u = R_1 \xi_1 + \tilde{R}_2 \xi_2.$$

Under the hypothesis that  $\tilde{H}$  is invertible, we get

$$(3.14) \quad \delta u = \tilde{T}_1 \xi_1 + \tilde{T}_2 \xi_2,$$

where  $\tilde{T}_1 = \tilde{H}^{-1} R_1$ ,  $\tilde{T}_2 = \tilde{H}^{-1} \tilde{R}_2$ ,  $\tilde{T}_1 : X \rightarrow X$ ,  $\tilde{T}_2 : Y_{obs} \rightarrow X$ . We shall call  $\tilde{T}_1, \tilde{T}_2$  the error transfer operators.

Let us note that the functions  $\varphi, \tilde{\varphi}$  in (3.5)–(3.7) depend on  $\xi_1, \xi_2$ , so as a result, the error transfer operators  $\tilde{T}_1, \tilde{T}_2$  also depend on  $\xi_1, \xi_2$  (nonlinearly), and it is not possible to represent  $\delta u$  through  $\xi_1, \xi_2$  in an explicit form. To derive from (3.14) the covariance operator of  $\delta u$ , we need to introduce some approximations  $T_1, T_2$  of the operators  $\tilde{T}_1, \tilde{T}_2$ , which do not depend on  $\xi_1, \xi_2$ . Consider the functions  $\varphi, \tilde{\varphi}$  in (3.5)–(3.7). Since  $\tilde{\varphi} = \bar{\varphi} + \tau \delta \varphi$ ,  $\varphi = \bar{\varphi} + \delta \varphi$ , we assume that

$$(3.15) \quad F'(\tilde{\varphi}) \approx F'(\bar{\varphi}), \quad (F'(\varphi))^* \approx (F'(\bar{\varphi}))^*,$$

then (3.5)–(3.7) reduces to

$$(3.16) \quad \begin{cases} \frac{\partial \delta \varphi}{\partial t} - F'(\bar{\varphi}) \delta \varphi &= 0, \quad t \in (0, T), \\ \delta \varphi|_{t=0} &= \delta u, \end{cases}$$

$$(3.17) \quad \begin{cases} -\frac{\partial \varphi^*}{\partial t} - (F'(\bar{\varphi}))^* \varphi^* &= -C^* V_2 (C \delta \varphi - \xi_2), \\ \varphi^*|_{t=T} &= 0, \end{cases}$$

$$(3.18) \quad V_1(\delta u - \xi_1) - \varphi^*|_{t=0} = 0.$$

The assumption (3.15) is the first-order approximation of the Taylor-Lagrange formula for  $F'$  under the hypothesis that  $F$  is twice continuously Frechet differentiable

[21]. The problem (3.16)–(3.18) is a linear problem; with fixed  $\bar{\varphi}$  it is the necessary optimality condition to the following optimal control problem: find  $u$  and  $\varphi$  such that

$$(3.19) \quad \begin{cases} \frac{\partial \delta \varphi}{\partial t} - F'(\bar{\varphi})\delta \varphi = 0, & t \in (0, T) \\ \delta \varphi|_{t=0} = \delta u \\ S_1(\delta u) = \inf_v S_1(v), \end{cases}$$

where

$$(3.20) \quad S_1(\delta u) = \frac{1}{2}(V_1(\delta u - \xi_1), \delta u - \xi_1)_X + \frac{1}{2}(V_2(C\delta \varphi - \xi_2), C\delta \varphi - \xi_2)_{Y_{obs}}.$$

We shall refer to the problem (3.19)–(3.20) as 'auxiliary DA problem'.

With the choice (3.15), the operator  $\tilde{H}$  is the Hessian  $H$  of the problem (3.19)–(3.20). This Hessian is defined by the successive solutions of the following problems:

$$(3.21) \quad \begin{cases} \frac{\partial \psi}{\partial t} - F'(\bar{\varphi})\psi = 0, & t \in (0, T), \\ \psi|_{t=0} = v, \end{cases}$$

$$(3.22) \quad \begin{cases} -\frac{\partial \psi^*}{\partial t} - (F'(\bar{\varphi}))^*\psi^* = -C^*V_2C\psi, & t \in (0, T) \\ \psi^*|_{t=T} = 0, \end{cases}$$

$$(3.23) \quad Hv = V_1v - \psi^*|_{t=0}.$$

Note that for  $\xi_2 = 0$  the operator  $H$  coincides with the Hessian of the original nonlinear DA problem  $\mathcal{H}$  on the exact solution  $u = \bar{u}$ .

The Hessian  $H$  acts in  $X$  as a self-adjoint operator with domain of definition  $D(H)=X$ . Moreover, due to  $V_1, V_2$ , the operator  $H$  is positive definite, and hence invertible. Then,  $T_1=H^{-1}R_1$ ,  $T_2=H^{-1}R_2$ , where  $R_2$  is  $\tilde{R}_2$  with  $\varphi$  replaced by  $\bar{\varphi}$  in (3.12).

Note that if the tangent linear hypothesis is valid (e.g. [11]), then for small  $\delta \varphi$  we can choose (3.15), and so we can use  $T_1, T_2$  instead of  $\tilde{T}_1, \tilde{T}_2$ . However, the TLH is not the necessary condition, i.e. the transition from  $\tilde{T}_1, \tilde{T}_2$  to  $T_1, T_2$  may not require the TLH to be valid.

As follows from (3.14), the impact of the errors  $\xi_1, \xi_2$  on the value of the analysis error  $\delta u$  is determined by the error transfer operators. The norm of these operators may be considered as a sensitivity coefficient: the smaller the norm, the smaller the impact of  $\xi_i$  on  $\delta u$ . This approach was used for deterministic error analysis in [9, 15, 17]. Here, assuming the stochastic nature of the errors  $\xi_1, \xi_2$ , we will derive the covariance operator through the covariance operators of the input errors.

**4. Covariance operators.** Below we suppose that the data errors  $\xi_1, \xi_2$  are normally distributed, unbiased and mutually uncorrelated. By  $V_{\xi_i}$  we denote the covariance operator of the corresponding data error  $\xi_i$ ,  $i = 1, 2$ , i.e.  $V_{\xi_1} \cdot = E[(\cdot, \xi_1)_X \xi_1]$ ,  $V_{\xi_2} \cdot = E[(\cdot, \xi_2)_{Y_{obs}} \xi_2]$ , where  $E$  is the expectation. By  $V_{\delta u}$  we denote the covariance operator of the analysis error:  $V_{\delta u} \cdot = E[(\cdot, \delta u)_X \delta u]$ .

The covariance matrix estimated by the inverse Hessian in variational data assimilation has been considered by many authors (e.g. [29, 30, 28, 11, 32, 14]) for a

linearized evolution model (using the TLH). We will use the equation (3.14) to derive the formulas for estimating the covariance operator  $V_{\delta u}$  involving the Hessian of the auxiliary DA problem (3.19)–(3.20).

Consider the error equation (3.14) with  $\tilde{T}_1, \tilde{T}_2$  replaced by  $T_1, T_2$ :

$$(4.1) \quad \delta u = T_1 \xi_1 + T_2 \xi_2,$$

where  $T_i = H^{-1}R_i$ ,  $T_1 : X \rightarrow X$ ,  $T_2 : Y_{obs} \rightarrow X$ . Let us denote by  $V$  the covariance operator  $V \cdot = E[(\cdot, \delta u)_X \delta u]$  where  $\delta u$  satisfies the approximate error equation (4.1). From (4.1) we get

$$(4.2) \quad V = T_1 V_{\xi_1} T_1^* + T_2 V_{\xi_2} T_2^*.$$

To find the covariance operator  $V$ , 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

$$(4.3) \quad T_1 V_{\xi_1} T_1^* = H^{-1}V_1 H^{-1} = H^{-1}V_{\xi_1}^{-1}H^{-1}.$$

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  $(R_2 g, p)_X$ ,  $g \in Y_{obs}$ ,  $p \in X$ . From (3.11)–(3.12),

$$(R_2 g, p)_X = (\theta^*|_{t=0}, p)_X = (C^* V_2 g, \phi)_Y = (g, R_2^* p)_{Y_{obs}},$$

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

$$(4.4) \quad \begin{cases} \frac{\partial \phi}{\partial t} - F'(\bar{\varphi})\phi = 0, & t \in (0, T), \\ \phi|_{t=0} = p. \end{cases}$$

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

$$(4.5) \quad Hp = v,$$

$$(4.6) \quad \begin{cases} \frac{\partial \phi}{\partial t} - F'(\bar{\varphi})\phi = 0, & t \in (0, T), \\ \phi|_{t=0} = p, \end{cases}$$

$$(4.7) \quad \begin{cases} -\frac{\partial \theta^*}{\partial t} - (F'(\bar{\varphi}))^* \theta^* = C^* V_2 V_{\xi_2} V_2 C \phi, & t \in (0, T) \\ \theta^*|_{t=T} = 0, \end{cases}$$

$$(4.8) \quad Hw = \theta^*|_{t=0},$$

then

$$(4.9) \quad T_2 V_{\xi_2} T_2^* v = w.$$

If  $V_2 = V_{\xi_2}^{-1}$ , then  $C^*V_2V_{\xi_2}V_2C = C^*V_2C$  and from (4.6)–(4.7) we obtain that

$$\theta^*|_{t=0} = Hp - V_1p,$$

where  $H$  is the Hessian defined by (3.21)–(3.23).

Then we get

$$R_2V_{\xi_2}R_2^* = H - V_1$$

and

$$(4.10) \quad T_2V_{\xi_2}T_2^* = H^{-1}R_2V_{\xi_2}R_2^*H^{-1} = H^{-1}(H - V_1)H^{-1}.$$

From (4.3), (4.10) it follows the result for  $V$ :

$$(4.11) \quad V = T_1V_{\xi_1}T_1^* + T_2V_{\xi_2}T_2^* = H^{-1}HH^{-1} = H^{-1}.$$

The last formula gives the covariance operator through the Hessian  $H$  defined by (3.21)–(3.23). Note that  $H$  is the Hessian of the auxiliary DA problem (3.19)–(3.20) (i.e. the Hessian of the cost functional  $S_1(\delta u)$  in the auxiliary DA problem (3.19)–(3.20)). This result coincides with well-known results reported elsewhere (e.g. [29, 28]), which have been obtained by replacing the original nonlinear model with the TLM from the very beginning (assuming the TLH is valid). In our analysis we do not need the TLH, when deriving the system for errors (3.5)–(3.7). Later we require that the operators  $T_i$  must be good approximations to  $\tilde{T}_i$ . This condition could be satisfied even if the TLH is not valid. Note also that all previous results reported are given for a discretized problem, whereas we provide the proof for the continuous case.

*Remark.* The formulas (3.16)–(3.18) involve the exact solution  $\bar{u}, \bar{\varphi}$  which, in fact, is not known. However, often it is supposed to be known in the framework of the "identical twin experiment" used to tune data assimilation algorithms. On the other hand, instead of  $\bar{u}, \bar{\varphi}$ , one may use some optimal solution  $u_{opt}, \varphi_{opt}$  obtained in the course of a current data assimilation procedure. In this case, assuming  $u_0 = u_{opt} + \xi_1$ ,  $\varphi_{obs} = C\varphi_{opt} + \xi_2$ , we obtain the same system (3.16)–(3.18) for the deviation errors  $\delta u = u - u_{opt}$ ,  $\delta\varphi = \varphi - \varphi_{opt}$  with  $\bar{u} = u_{opt}$ ,  $\bar{\varphi} = \varphi_{opt}$ .

## 5. Algorithms for computing the covariance matrix.

**5.1. BFGS.** Consider the covariance operator  $V$  defined by (4.11):

$$(5.1) \quad V = H^{-1}.$$

There are different numerical methods to compute the covariance matrix. The method of finite differences described in [8] computes the Hessian of the original nonlinear problem  $\mathcal{H}$ , which remains to be inverted. We also know that in the nonlinear case the covariance is related to the inverse Hessian of the auxiliary DA problem  $H$ , not to  $\mathcal{H}$ . Besides, this method is not sufficiently accurate due to truncations used in a local Taylor expansion and is expensive for practical implementation. The sensitivity matrix method [29] computes directly  $H^{-1}$ . This method is computationally efficient if the dimension of the observation vector is much smaller than the dimension of the state vector, and so is feasible mainly for the 3D-VAR applications. It requires full storage of the resulting covariance matrix. The second order adjoint method allows the action  $Hv$  to be computed, thus does not require full storage of  $H$ . The inverse

Hessian may be constructed by solving the eigenvalue problem involving  $Hv$  using Lanczos-type methods [24] and taking into account that  $H$  is self-adjoint.

In order to find the inverse Hessian  $H^{-1}$ , the quasi-Newton BFGS algorithm may also be used [26, 7]. This algorithm generates  $H^{-1}$  in the course of a minimization process. Furthermore, it has been proved that for a linear-quadratic case (linear constraints, quadratic cost functional), the sequence of the BFGS updates leads to the exact value of  $H^{-1}$  in at most  $n$  iterations, where  $n$  is the state vector dimension [10].

The use of the BFGS in this context is attractive because of the possibility of getting the optimal solution and the covariance matrix in 'one shot'. Indeed, this is a tempting idea in the framework of the incremental approach widely used in practical DA. However, these are two different tasks and there are different criteria used to measure progress. For DA problems it is natural to start minimization from the closest initial guess available and stop iterations as soon as successive iterates become stuck. However, a set of updates produced during this particular minimization process may turn out to be insufficient to reveal (implicitly) the 'leading' eigenvalues, which are needed to approximate  $H^{-1}$  accurately. Since we deal with the nonlinear problem here, we cannot solve the DA problem and find the covariance matrix at the same time. Therefore, we intend to solve independently the auxiliary DA problem (3.19)-(3.20) by the BFGS algorithm, then retrieve  $H^{-1}$  as the main result of the minimization. Let us note that the problem (3.19)-(3.20) is defined by values  $\xi_1$  and  $\xi_2$ , which could be arbitrary non-trivial functions theoretically. We can specify these functions to produce the best estimate of the covariance matrix. We must also choose the initial guess for the minimization process.

Applied for solving the auxiliary DA problem (3.19)-(3.20), the BFGS algorithm has the form [26]:

$$(5.2) \quad d^k = H_k^{-1} S'_1(\delta u^k),$$

$$(5.3) \quad \delta u^{k+1} = \delta u^k - \alpha^k d^k,$$

$$(5.4) \quad H_{k+1}^{-1} = \left( I - \frac{sy^T}{y^T s} \right) H_k^{-1} \left( I - \frac{ys^T}{y^T s} \right) + \frac{ss^T}{y^T s},$$

where  $s = \delta u^{k+1} - \delta u^k$ ,  $y = S'_1(\delta u^{k+1}) - S'_1(\delta u^k)$ ,  $H_k^{-1}$  is the approximation to  $H^{-1}$  on the  $k$ -th iteration,  $S'_1(\delta u^k)$  is the value of the gradient of  $S_1$  in  $\delta u$  at the point  $\delta u^k$ ,  $\alpha^k$  are iterative parameters,  $I$  is the identity operator.

In the limited-memory version of the algorithm (so-called LBFGS, [20]) the matrices  $H_{k+1}^{-1}$  are not explicitly computed at each iteration. Instead, the algorithm computes the action  $H_{k+1}^{-1}v$  by the Broyden-Fletcher-Goldfarb-Shanno (BFGS) formula (5.4), using the last  $m$  pairs of vectors  $s$  and  $y$ , which are stored in memory in a cyclic order. Keeping in mind potential large scale applications, we use the LBFGS algorithm, but by taking  $m > n$  we retain all  $m$  pairs of vectors  $s, y$  produced in the course of minimization (for now). The covariance matrix itself can be obtained by applying successively (5.4) to the unity vectors.

Consider the auxiliary DA problem (3.19)-(3.20). There are three important points in tuning up the BFGS algorithm for computing the covariance matrix:

a) we must specify driving functions  $\xi_1, \xi_2$  entering (3.20). We suggest as follows:

$$(5.5) \quad \xi_1 = \tilde{u}, \quad \xi_2 = C\delta\tilde{\varphi},$$

where  $\delta\tilde{\varphi}$  satisfies the problem

$$(5.6) \quad \begin{cases} \frac{\partial \delta\tilde{\varphi}}{\partial t} - F'(\tilde{\varphi})\delta\tilde{\varphi} = 0, & t \in (0, T), \\ \delta\tilde{\varphi}|_{t=0} = \tilde{u}. \end{cases}$$

In this case, the solution of (3.19) is  $\delta u = \tilde{u}$ , and  $S_1(\tilde{u}) = 0$ ;

b) we start iterating from a most distant initial point, for example using a large constant value  $u^0 = 10^{35}$ . This is done to guarantee that the number of iterations required for minimization is sufficient to 'evoke' all leading eigenvalues of the inverse Hessian (we assume that the BFGS algorithm solves the eigenvalue problem implicitly). When the approximation of the inverse Hessian accumulated in the BFGS is sufficiently accurate, both the cost functional  $S_1(\delta u)$  and the norm of the gradient start to decrease sharply. When this acceleration is noticed, iterations must be stopped;

c) One needs the exact minimum along the direction of descent to be achieved. This is a necessary condition for the BFGS to be an exact method for computing the inverse Hessian [10]. We note that the auxiliary DA problem (3.19)-(3.20) is a linear problem, therefore the analytical step search is available. Let  $k$  denote the iteration index and  $d^k$  the direction of descent built by the BFGS algorithm, then the optimal step  $\alpha^k$  can be derived from the condition

$$(5.7) \quad \frac{\partial S_1(\delta u^{k-1} + \alpha d^k)}{\partial \alpha} = 0.$$

Applying this condition to (3.20) we obtain as follows

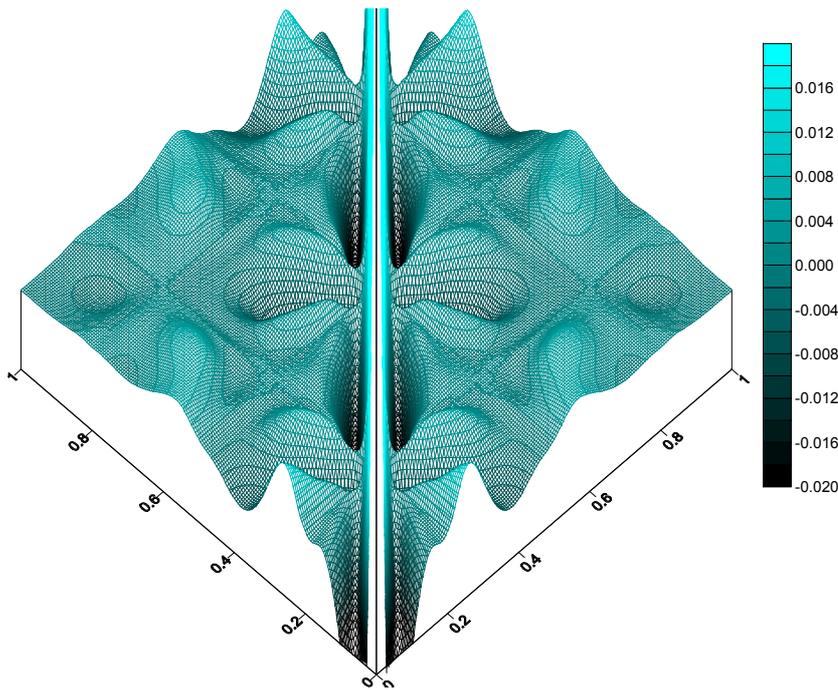
$$(5.8) \quad \alpha^k = - \frac{(V_1(\delta u^{k-1} - \xi_1), d^k)_X + (V_2(C\delta\varphi^{k-1} - \xi_2), C\varphi_d)_{Y_{obs}}}{(V_1 d^k, d^k)_X + (V_2 C\varphi_d, C\varphi_d)_{Y_{obs}}},$$

where  $\varphi_d$  and  $\delta\varphi^k$  are the solutions of the problem (5.6) for  $\tilde{u} = d^k$  and  $\tilde{u} = \delta u^k$ , respectively.

In order to compute the gradient of (3.20) given the constraints in (3.19), we use the automatic differentiation tool TAPENADE [12]. First we differentiate the algorithm which implements the nonlinear evolution problem (2.1) and computes the cost functional (2.3) afterwards. The output TLM and adjoint codes are verified using classical tests. Then we use the TLM code as a model in (3.19), (3.20), while the adjoint code is used to compute the gradient. It is worth saying here that our attempt to obtain the inverse Hessian using inconsistent TLM and adjoint models (i.e. obtained using the 'differentiate-then-discretize' approach) were unsuccessful. When the BFGS is properly tuned and the TLM and adjoint models are consistent the method produces the estimates of excellent quality. As an example we show the covariance matrix for the 1D linear diffusion evolution problem in Fig.5.1 (some additional details on this case are presented in Sect.7). Let us note that in [7] for example, the BFGS method is considered as inferior to the Lanczos method from the point of view of accuracy.

From now on we shall call the  $H$ -covariance an estimate of the covariance matrix  $V_{\delta u}$  via the inverse Hessian of the auxiliary DA problem (3.19)-(3.20) obtained by the BFGS algorithm. We also denote by  $\sigma^2 := \text{diag}(V)$  the  $H$ -variance.

*Remark 1.* The use of the LBFGS which keeps only a few latest pair-vectors is discussed in [32] in a slightly different context (preconditioning). This possibility cannot be ruled out in general, but it is not clear how to assess in advance the number of latest pair-vectors to be kept in memory. We will show in numerical experiments

FIG. 5.1. *Covariance: linear diffusion problem*

that in some cases it might be difficult to separate the 'leading' eigenvalues from the rest of the eigenvalue spectrum of the inverse Hessian. As compared to the Lanczos-type methods, the LBFGS is probably easier to implement.

*Remark 2.* For all numerical examples presented in Sec.7 we verify the covariance matrix produced by the BFGS algorithm against the exact value of  $H^{-1}$ . That is, we explicitly compute  $H$  as a matrix by successively applying (3.21)-(3.23) to the unity vectors, then invert  $H$  using the singular value decomposition (SVD) technique. In most cases the difference is small and cannot be visually noticed on the graphs.

**5.2. Fully nonlinear ensemble method.** In order to check the  $H$ -covariance matrix we estimate  $V_{\delta u}$  using an ensemble (statistical) approach. Let us assume that we know the 'true' state  $\bar{u}$ . Then we compute  $u_b = \bar{u} + \xi_1$  and  $\hat{\varphi} = C\bar{\varphi} + \xi_2$  where  $\xi_1, \xi_2$  are individual implementations of the background error and the measurement error. Using those values we solve the original nonlinear DA problem (2.2)-(2.3) and find the solution error  $\delta u = u - \bar{u}$ . This procedure is repeated  $n$  times for new values  $\xi_1, \xi_2$  each time to get the ensemble of errors  $\delta u$ , then the ensemble covariance matrix is

$$(5.9) \quad \hat{V} = \frac{1}{n} \sum_1^n \delta u \delta u^T.$$

Another way is to generate an ensemble of solutions  $u$ , then the ensemble covariance matrix is

$$(5.10) \quad \hat{V} = \frac{1}{n} \sum_1^n (u - \hat{E}[u])(u - \hat{E}[u])^T; \quad \hat{E}[u] = \frac{1}{n} \sum_1^n u.$$

We also define the ensemble variance as  $\hat{\sigma}^2 := \text{diag}(\hat{V})$ .

This ensemble method must be distinguished from ensemble methods used in filtering, e.g. [6], [33]. The main difference is that in filtering the analysis is the current state (the cost functional is formulated for a given time instant), while we are looking for a 'retrospective' analysis, i.e. the initial condition. Therefore, these ensemble methods are not suitable for the hind-cast problem. In filtering, the forecast of the ensemble relies on a linearized forecast operator, i.e. a linearization error is introduced into the covariance matrix at the forecast stage. Since we do not propagate the covariance matrix in time (and, therefore, do not use any linearization) we call this ensemble method fully nonlinear. Let us also recall that in the nonlinear case the expectation  $E[u]$  can be biased (i.e.  $\hat{E}[u] \neq \bar{u}$  as  $n \rightarrow \infty$ ), even though  $E[\xi_1] = E[\xi_2] = 0$ , and that the mean value  $E[\delta u]$  and the covariance matrix  $V_{\delta u}$  are not sufficient statistics to define the analysis error probability distribution function.

## 6. Details of numerical implementation.

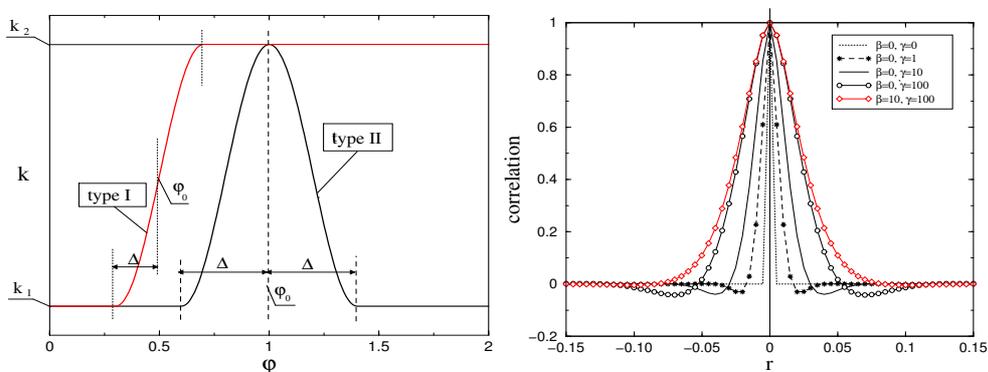


FIG. 6.1. Left - diffusion coefficient  $k(\varphi)$ . Right - the background error correlation function.

**6.1. Model.** As a model we use the 1D nonlinear convection-diffusion equation

$$(6.1) \quad \frac{\partial \varphi}{\partial t} + \frac{\partial(w\varphi)}{\partial x} = \frac{\partial}{\partial x} \left( k(\varphi) \frac{\partial \varphi}{\partial x} \right) + Q(\varphi), \quad t \in (0, T), \quad x \in (0, 1),$$

where  $\varphi = \varphi(t, x)$ ,  $w(t)$  is velocity,  $k(\varphi)$  is the diffusion coefficient,  $Q(\varphi)$  is the source term, with the Neumann boundary conditions

$$(6.2) \quad \frac{\partial \varphi}{\partial x} \Big|_{x=0} = \frac{\partial \varphi}{\partial x} \Big|_{x=1} = 0.$$

We use the implicit time discretization as follows

$$(6.3) \quad \frac{\varphi^i - \varphi^{i-1}}{h_t} + \frac{\partial(w\varphi^i)}{\partial x} - \frac{\partial}{\partial x} \left( k(\varphi^i) \frac{\partial \varphi^i}{\partial x} \right) - Q(\varphi^i) = 0, \quad i \in (1, \dots, N), \quad x \in (0, 1),$$

where  $i$  is the time integration index,  $h_t = T/N$  is a time step. The spatial operator is discretized on a uniform grid ( $h_x$  is the spatial discretization step,  $j = 1, \dots, m$  is the node number,  $m$  is the total number of mesh nodes) using the 'power law' first-order scheme as described in [25]. For each time step we perform nonlinear iterations,

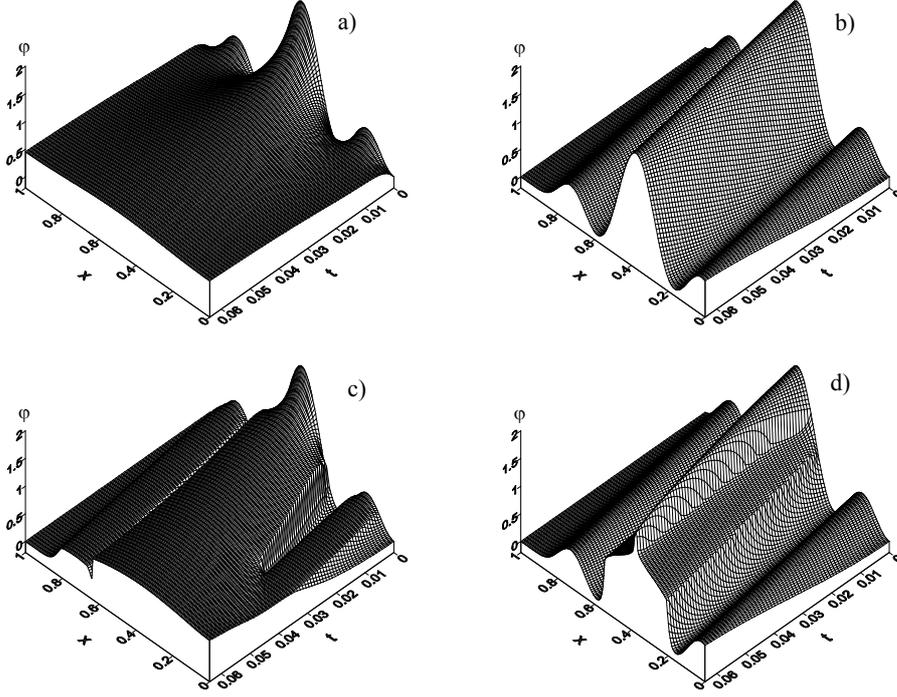


FIG. 6.2. Examples of field variable  $\varphi(t, x)$ : a) diffusion-dominated case, b) convection-dominated case, c)-d) nonlinear cases

assuming initially that  $k(\varphi^i) = k(\varphi^{i-1})$ ,  $Q(\varphi^i) = Q(\varphi^{i-1})$ , and keep iterating until (6.3) is satisfied (i.e. the norm of the left-hand side in (6.3) becomes smaller than a threshold  $\epsilon_1 = 10^{-12}\sqrt{m}$ ). In all computations presented in this paper we use the following parameters: observation period  $T = 0.064$ , discretization steps  $h_t = 0.001$ ,  $h_x = 0.005$ , state vector dimension  $m = 200$ .

From numerical experiments we found that the covariance matrix is most sensitive to the variation in  $k(\varphi)$  or, more exactly, in the Peclet number  $Pe = w/k(\varphi)$ . The source term  $Q(\varphi)$  affects the covariance matrix through the transport coefficients. For example, if  $k(\varphi) = \text{const}$ , then the independent influence of  $Q(\varphi)$  is negligible. Thus, we concentrate our attention on the nonlinearity in  $k(\varphi)$ . In order to set up tests we consider two types of  $k(\varphi)$  as shown in Fig.6.1(left). In the first case  $k(\varphi)$  varies smoothly from the level  $k_1$  to the level  $k_2$  within the interval  $[\varphi_0 - \Delta, \varphi_0 + \Delta]$  subjected to the rule (6.4)(type I), in the second case  $k(\varphi)$  varies between levels  $k_1$  and  $k_2$  and back to  $k_1$  within the interval  $[\varphi_0 - \Delta, \varphi_0 + \Delta]$  subjected to the rule (6.5)(type II):

$$(6.4) \quad k(\varphi) = \frac{1}{2}(k_1 + k_2) - \frac{1}{2}(k_1 - k_2) \sin(\pi(\varphi - \varphi_0)/(2\Delta))$$

$$(6.5) \quad k(\varphi) = k_1 + \frac{1}{2}(k_2 - k_1)(1 + \cos(\pi(\varphi - \varphi_0)/(\Delta)))$$

We can control the level of nonlinearity by changing parameters  $\Delta$  and  $k_1, k_2$ .

In the following examples we compute the evolution of the initial condition given by the expression

$$(6.6) \quad \varphi(0, x) = \sum_{k=1}^3 a_k (1 - \cos 2k\pi x),$$

where the coefficients are:  $a_1 = 0.5$ ,  $a_2 = -0.5$ ,  $a_3 = 0.5$ . We consider four different cases: a) diffusion-dominated case  $k = 1.0$ ,  $w = 0$  (Fig.6.2(a)); b) convection-dominated case  $k = 0.001$ ,  $w = -2.0$  (Fig.6.2(b)); c) nonlinear case  $w = -2.0$ ,  $k(\varphi)$  is defined by (6.4)(type I) with parameters  $k_1 = 0.01$ ,  $k_2 = 1.0$ ,  $\Delta = 0.2$  and  $\varphi_0 = 0.5$  (Fig.6.2(c)); d) nonlinear case  $w = -2.0$ ,  $k(\varphi)$  is defined by (6.5)(type II) with parameters  $k_1 = 0.001$ ,  $k_2 = 0.2$ ,  $\Delta = 0.4$  and  $\varphi_0 = 1.0$  (Fig.6.2(d)).

It can be seen that in the diffusion-dominated case the initial state function  $\varphi(0, x)$  quickly degenerates approaching a uniform value, while in the convection-dominated case it is transported along characteristics (some dissipation due to numerical diffusion can still be noticed). In the third case the solution exhibits a mixed behavior with a presence of sharp local field gradients. In the fourth case one can see the most interesting example of evolution. Here diffusion occurs only within a narrow band of  $\varphi$  values. As a result, a smooth initial state function ends up as a polygon like shape, while the 'central peak' is consumed to keep sharp field gradients. We do not consider this behavior as a meaningful physical process (even though it could be). Our purpose is to set up such tests where nonlinear effects become dominant.

**6.2. Background error covariance matrix and noise generation.** In the numerical implementation we deal with a finite-dimensional problem, hence we will assume that all operators in the cost functional are matrices. In order to define (2.3), (3.20) one needs to specify the weights  $V_1 = V_{\xi_1}^{-1}$  and  $V_2 = V_{\xi_2}^{-1}$ , where  $V_{\xi_1}$  is the background error covariance matrix and  $V_{\xi_2}$  is the measurement error covariance matrix. Those two usually represent our a-priori knowledge on the stochastic properties of errors. If the problem (2.2) is solved repeatedly in time, then  $V_{\xi_1}$  can be estimated from the previous covariance matrix (or directly from the previous Hessian  $H$ ).

The simplest approach is to assume that  $V_{\xi_1}$ ,  $V_{\xi_2}$  are diagonal matrices with the elements equal to  $\sigma_b^2(j)$ ,  $j = 1, \dots, m$  and  $\sigma_{obs}^2$ . This could be a reasonable assumption about  $V_{\xi_2}$ , but it is too simplistic concerning  $V_{\xi_1}$ . We illustrate that if  $V_{\xi_1}$  is diagonal, then the variance looks almost trivial and can be easily predicted without computations. Therefore, the off-diagonal elements must be introduced into  $V_{\xi_1}$ .

In solving ill-posed inverse problems [31] it is often assumed that the solution is a smooth function which belongs to a Sobolev space of certain order, particularly  $W_2^2$  is defined by the norm

$$(6.7) \quad \|\cdot\|_{W_2^2(0,1)}^2 = \int_0^1 \left( (\cdot)^2 + \beta^* \left( \frac{\partial \cdot}{\partial x} \right)^2 + \gamma^* \left( \frac{\partial^2 \cdot}{\partial x^2} \right)^2 \right) dx.$$

Let us assume that the background error  $\xi_1$  is a smooth function, i.e. it belongs to  $W_2^2$ . In the finite-dimensional form this would be equivalent to the weight matrix

$$(6.8) \quad V_1 = I + \beta A_1 + \gamma A_2,$$

where  $\beta = \beta^*/h_x^2$ ,  $\gamma = \gamma^*/h_x^4$ ,  $I$ ,  $A_1$ ,  $A_2$  are  $m \times m$  matrices, such that  $I$  is the identity

matrix and  $A_1$  and  $A_2$  are as follows

$$A_1 = \begin{pmatrix} 1 & -1 & \cdot & \cdot & \cdot & 0 \\ -1 & 2 & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & -1 & 2 & -1 \\ 0 & \cdot & \cdot & \cdot & -1 & 1 \end{pmatrix}, A_2 = \begin{pmatrix} 1 & -2 & 1 & \cdot & \cdot & \cdot & \cdot & 0 \\ -2 & 5 & -4 & 1 & \cdot & \cdot & \cdot & \cdot \\ 1 & -4 & 6 & -4 & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & -4 & 6 & -4 & 1 \\ \cdot & \cdot & \cdot & \cdot & 1 & -4 & 5 & -2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 & -2 & 1 \end{pmatrix}.$$

Then the covariance matrix of the background error can be computed as follows

$$V_{\xi_1} = \frac{\sigma_b^2}{c} V_1^{-1},$$

where  $c$  is a diagonal element of the matrix  $V_1^{-1}$  taken at a node far enough from the domain boundaries. The row of  $V_1^{-1}$  which corresponds to the same node for different values  $\beta$  and  $\gamma$  is shown in Fig.6.1(right). This behavior remains unchanged for all nodes within the interior domain, but changes in the vicinity of the boundaries. From this figure we may conclude that when  $\gamma$  increases, the correlation radius increases,  $\beta$  affects the shape of the correlation function stretching it in the vertical direction. Here we can observe elements of similarity between Tikhonov regularization and the covariance matrix approach: the regularization parameter is related to  $c/\sigma_b^2$ , parameter  $\gamma$  controls the correlation radius, while  $\beta$  is responsible for the shape of correlation function. In numerical experiments we always assume  $\beta = 0$ .

Other methods exist to specify the background covariance matrix, see e.g. [13], [3].

In order to perform ensemble computations we have to generate the error  $\xi_1$  such that its covariance matrix is  $V_{\xi_1}$ . Let  $V_{\xi_1} = U\Sigma U^T$  be the SVD of  $V_{\xi_1}$ , then the individual error can be obtained as follows

$$(6.9) \quad \xi_1 = U\sqrt{\Sigma}\xi^{(I)},$$

where  $\xi^{(I)}$  is a normally distributed random series, uncorrelated, with zero mean and standard deviation equal to the unity. This series is produced by a pseudo-random generator, which uses the subroutines 'gasdev' and 'run2' as described in [27].

We assume that the observation error covariance matrix  $V_{\xi_2}$  is diagonal with the elements  $\sigma_{obs}^2$ , therefore

$$(6.10) \quad \xi_2 = \sigma_{obs}\xi^{(II)},$$

where  $\xi^{(II)}$  is produced by a pseudo-random generator as above. The superscripts  $(I)$ ,  $(II)$  in (6.9)-(6.10) stand to show that two separate random generators are used to produce mutually uncorrelated series  $\xi^{(I)}$ ,  $\xi^{(II)}$ .

We assume that the observations  $\hat{\varphi}$  are available every time step at the sensor locations  $\hat{x}_k$ , such that  $\hat{x}_1 = 0.2$ ,  $\hat{x}_2 = 0.5$  and  $\hat{x}_3 = 0.8$ . Thus, the observation operator  $C$  is defined. The same sensor configuration is used for all computations presented in this paper.

**6.3. Minimization algorithm.** In order to solve the original nonlinear DA problem (2.2)-(2.3) we also use the quasi-Newton BFGS algorithm. In ensemble computations the 'true' state  $u = \bar{u}$  is always used as a starting point for the minimization process. For step search we use the routine CSRCH [22], while the initial value of descent step is obtained by (5.8). The number of function evaluations allowed in the

step search routine is limited by 20. The BFGS algorithm is embedded into the outer loop. This allows us to restart the minimization process from the latest approximation of the sought control  $u$  if iterations are stopped with an error message. The outer loop stops if the restarted BFGS algorithm fails to produce any further minimization. The final result is discarded if the norm of the gradient remains larger than a certain threshold. The outer loop is introduced because we deal with a strongly nonlinear problem. As the nonlinearity (controlled mainly by  $\Delta$  in (6.4)–(6.5)) and the correlation radius (controlled by  $\gamma$  in (6.7)) grow, multiple points of zero variation of the cost functional (2.3) start to appear and the distance between the 'true' solution and 'ghost' solutions becomes noticeable. Generally, in order to search for a global minimum one must restart from a modified value of the latest approximation of  $u$ . However, in this paper we do not use global optimization.

## 7. Numerical examples.

**7.1. Verification of the methods for computing covariance.** In the linear case the  $H$ -covariance is the exact estimate of the actual covariance  $V_{\delta u}$ . Therefore, we check that the  $H$ -covariance matrix and the ensemble covariance matrix are equivalent. Let  $n$  be the ensemble size. We assume that the 'true' solution  $\bar{u}$  is given by (6.6) and  $\xi_1, \xi_2$  are individual implementations of the background error given by (6.9) and the observation error given by (6.10). The following parameters are used in this test:  $\sigma_b^2(j) = 0.1$ ,  $j = 1, \dots, m$ ,  $\sigma_{obs}^2 = 10^{-3}$ ;  $\gamma = 10^2$ . We compute the  $H$ -covariance matrix as explained in Sect.5.1, and the ensemble covariance matrix  $\hat{V}$  for different ensemble sizes  $n = 100, 400, 1600$ .

The results for the convection-dominated case (Fig.6.2(b)) are presented in Fig.7.1 and Fig.7.2. The bold solid line in Fig.7.1(left) shows the  $H$ -variance, other lines show the ensemble variance (5.9) for different  $n$ . One can see that as  $n$  increases,  $\hat{\sigma}^2$  converges to  $\sigma^2$ . The same behavior holds if  $\hat{\sigma}^2$  is computed using (5.10), though for small  $n$  the ensemble variance deviates from the  $H$ -variance more strongly. In Fig.7.2 we show the  $H$ -covariance (upper/left), and the ensemble covariance for  $n = 100$  (upper/right),  $n = 400$  (lower/left) and  $n = 1600$  (lower/right). It can be seen again that when  $n$  increases,  $\hat{V}$  converges to  $V$ . Let us note that even though the ensemble variance looks quite satisfactory (in relative terms) for  $n = 100$ , the off-diagonal elements show the presence of 'ghost' correlations of the same magnitude as true correlations. That is, in order to obtain a good estimation of the whole covariance matrix (rather than only its diagonal elements) one apparently needs an ensemble size much larger than  $n = 100$ . In this respect it looks somewhat surprising that the value  $n = 100$  is considered as a sufficient ensemble size in [6], while  $n = 10$  is used in [33]. It is worth mentioning that there are special methods (e.g., so-called 'localization') for dealing with 'ghost' (spurious) correlations arising from small samples in ensemble data assimilation [2].

We also compute the variance for different spatial discretization steps. These results are presented in Fig.7.1(right). Here, the case  $m = 200$  corresponds to the case presented in the left panel. We can notice that as the mesh size  $h_x = 1/m$  decreases, the solution behaves consistently, converging to a limit. This behavior makes us assured that the discretization of the convection-diffusion operator is correct.

## 7.2. Some properties of the variance.

**7.2.1. Variance and transport coefficients.** Here we discuss qualitative properties of the variance  $\sigma^2$ , which can be considered as an optimal solution accuracy. All results presented here are obtained for a linear problem, thus the  $H$ -variance is the

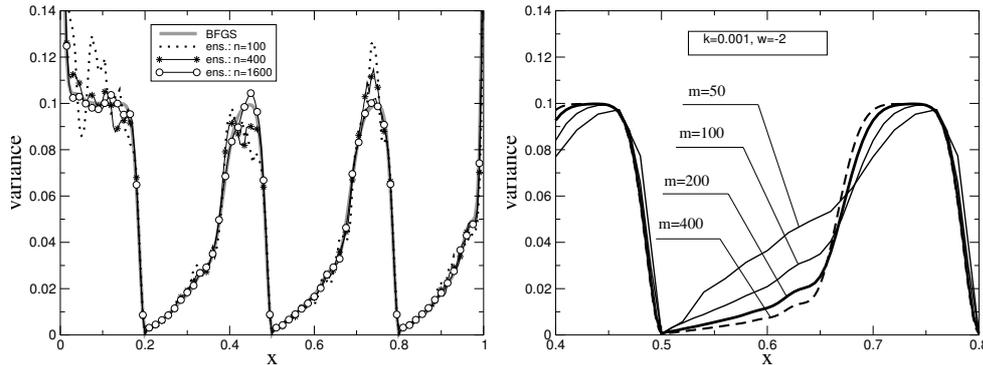


FIG. 7.1. *Left: H-variance and ensemble variance for different ensemble sizes. Right: H-variance for different spatial discretization steps  $h_x = 1/m$ .*

exact estimate of the actual variance. A basic feature of the variance (which could be predicted without computations) is that it varies from the level of observation error variance  $\sigma_{obs}^2$  at the observation points to the level of background error variance  $\sigma_b^2$ . This change happens within the 'region of influence' [5], the size of which depends mainly on the transport phenomena supported by the model and on the correlation radius of the background error.

In Fig.7.3(left) we present the variance as a function of convection rate  $w$  in (6.1) for the convection-dominated problem ( $k = 0.001$ , presented in Fig.6.2,b) and  $\gamma = 10^2$ . We define a 'convective range' as a domain of size  $L_c = wT$  measured from the observation point in the upstream direction, which in our case coincides with the direction of the  $x$ -axis. Apparently, perturbations in  $\varphi$  introduced outside this domain do not reach the sensor during the observation period.

Let us look first at the case  $w = -0.5$ ,  $L_c = 0.032$  (dashed line). We can see that at the observation points the variance  $\sigma^2$  is close to  $\sigma_{obs}^2$ , it sharply rises up to the background level  $\sigma_b^2$  in the downstream direction (to the left). In the upstream direction (to the right) there is a gentle slope within the convective range, then it rises up sharply to the background level  $\sigma_b^2$ . In the case  $w = -2$ ,  $L_c = 0.128$  the size of a gentle slope in the upstream direction is larger, though not sufficient to cover the distance between two observation points. One can also see a small transition area between the gentle and sharp slopes, which appears due to the influence of correlations in the background error. In the last case we consider  $w = -5$ ,  $L_c = 0.32$ , i.e. the whole area between two sensors is within the convective range. In this case the variance  $\sigma^2$  is dominated by measurements. It is clear that no information can propagate from the downstream direction, therefore  $\sigma^2 \approx \sigma_b^2$  for  $0 \leq x \leq 0.2$ .

Another factor which strongly affects the variance is the diffusion coefficient  $k(\varphi)$ . The influence of diffusion (dissipation) will be the focus of our investigation in Subsect.7.3 when we consider the nonlinear problem. Here we show the results for the linear problem in Fig.7.3(right), where the variance is computed for  $w = -2$ ,  $\gamma = 10^2$ , but for different values of the diffusion coefficient  $k = 10^{-3}, 10^{-2}, 10^{-1}, 10^0$ . The results show that diffusion acts in two contradictory ways. On one hand, it dissipates information (shape) by smoothing the field gradients. Therefore, the variance within the convective range increases when  $k(\varphi)$  grows (i.e. the optimal solution accuracy deteriorates). On the other hand, diffusion is a transport phenomena by itself, so

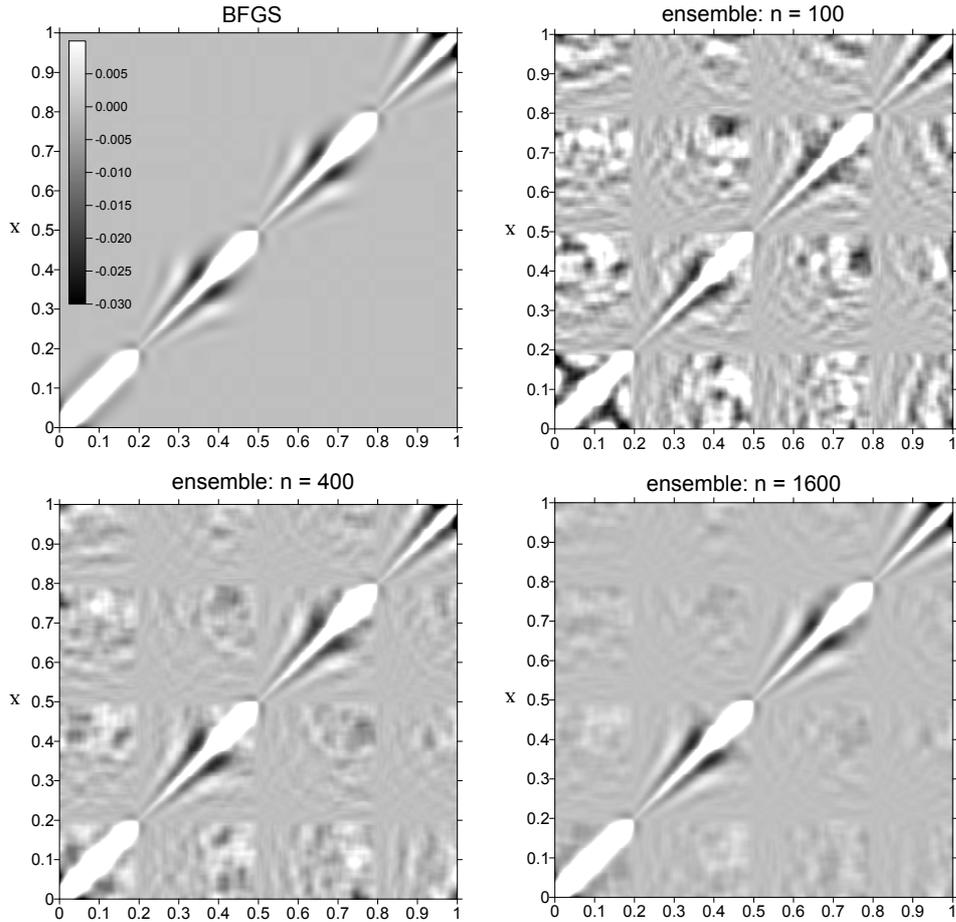


FIG. 7.2.  $H$ -covariance matrix and ensemble covariance matrix for different ensemble sizes (linear case)

outside the convective range the variance decreases (i.e. the optimal solution accuracy improves). Let us compare, for example, the variance computed for  $k = 10^{-1}$  (dashed line) and for  $k = 10^{-3}$  (bold solid line). We can see that at the observation points the variance has the expected value  $\sigma_{obs}^2$ , but it grows rapidly as we move upstream (to the right). Next we can observe a gentle slope shifted upwards compared to the convection-dominated case. This slope remains moderate as we move outside the convective range and, eventually, in the areas  $0.4 < x < 0.5$  and  $0.7 < x < 0.8$ , the variance becomes smaller than in the convection-dominated case.

**7.2.2. Variance and correlation radius of the background error.** An important factor which affects the variance is the correlation radius of the background error, in our case controlled by  $\gamma$ . In Fig.7.4 we show the convection-dominated case for  $w = -2$ , but for different values  $\gamma = 0, 1, 10, 100$ . Here we can see that as  $\gamma$  decreases the 'region of influence' quickly diminishes. In the case  $\gamma = 0$  which corresponds to the uncorrelated background error, we get an improvement of  $\sigma^2$  (in respect to  $\sigma_b^2$ ) very close to the observation point only (despite being within the convective

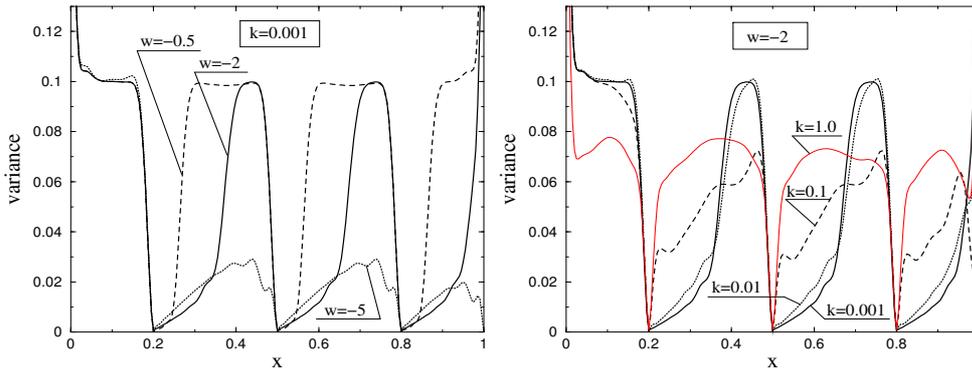


FIG. 7.3.  $H$ -variance. Left - as a function of convection rate  $w$ . Right - as a function of diffusion coefficient  $k$ .

range). This example shows that the off-diagonal elements of the background error covariance matrix are very important. If these elements are specified incorrectly, it may affect the  $H$ -covariance in a much more stronger way than the errors due to linearization of the error transfer operators.

Another important observation made in these experiments is that as the correlation radius increases the number of iterations required to obtain the  $H$ -covariance increases correspondingly. For example, we needed 31 iterations for  $\gamma = 0$ , (compare to the number of unknowns  $m = 200$ ), 120 iterations for  $\gamma = 1$  and about 200 iterations for  $\gamma \geq 10$ , i.e. essentially  $m$ . This can be explained by the fact that as  $\gamma$  increases the eigenvalue spectrum of the inverse Hessian shrinks, i.e. the eigenvalues of the inverse Hessian become badly separated and the number of required iterations naturally rises.

*Remark.* Let us note that due to (3.23) the Hessian  $H$  involves the matrix  $V_1$ . For  $\beta = 0$  (as always taken in numerical experiments), it follows from (6.8) that the eigenvalues  $\mu_k$  of  $V_1$  are defined by  $\mu_k = 1 + \gamma\lambda_k$ , where  $\lambda_k > 0$  are the eigenvalues of the matrix  $A_2$ . The difference between the maximal and minimal eigenvalues is determined by  $\mu_{max} - \mu_{min} = \gamma(\lambda_{max} - \lambda_{min})$ . It tends to the infinity as  $\gamma \rightarrow \infty$  for  $H$ . A similar difference for  $H^{-1}$  tends to zero, therefore the spectrum of  $H^{-1}$  shrinks.

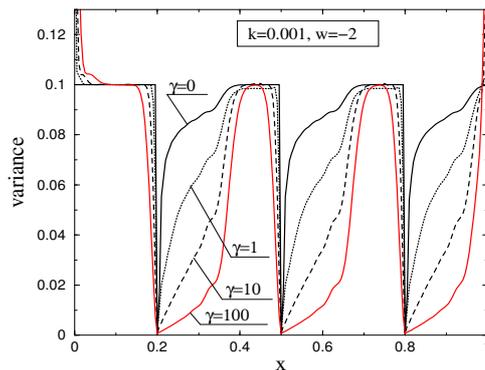


FIG. 7.4.  $H$ -variance as a function of correlation radius (related to  $\gamma$ ).

**7.3. Non-linear problem.** In the linear case the relationship (5.1) holds exactly, in the nonlinear case there could still be some confusion. First, there exists no exact relationship. It must be clearly stated that the covariance  $V_{\delta u}$  can be approximated by the  $H$ -covariance. Hence, there is no need to mention the Hessian of the original nonlinear DA problem  $\mathcal{H}$  in this respect. Apparently, the best approximation can be achieved when  $\bar{\varphi}$  in (3.19) is the solution to the problem (2.1) with  $u = \bar{u}$ , where  $\bar{u}$  is a 'true' initial condition. At a glance this looks senseless since we do not usually know the true value of controls (if we knew them, there would be no need to solve any DA problem). However, a classical methodology for creating and tuning DA algorithms is a so-called 'identical twin experiment', where bogus 'true' controls are specified. Thus, we can compute the covariance on a bogus 'true' initial condition and use this covariance as a reference to investigate its sensitivity to the variation  $u_{opt} - \bar{u}$ , where  $u_{opt}$  is an optimal solution actually available after solving the DA problem using real data.

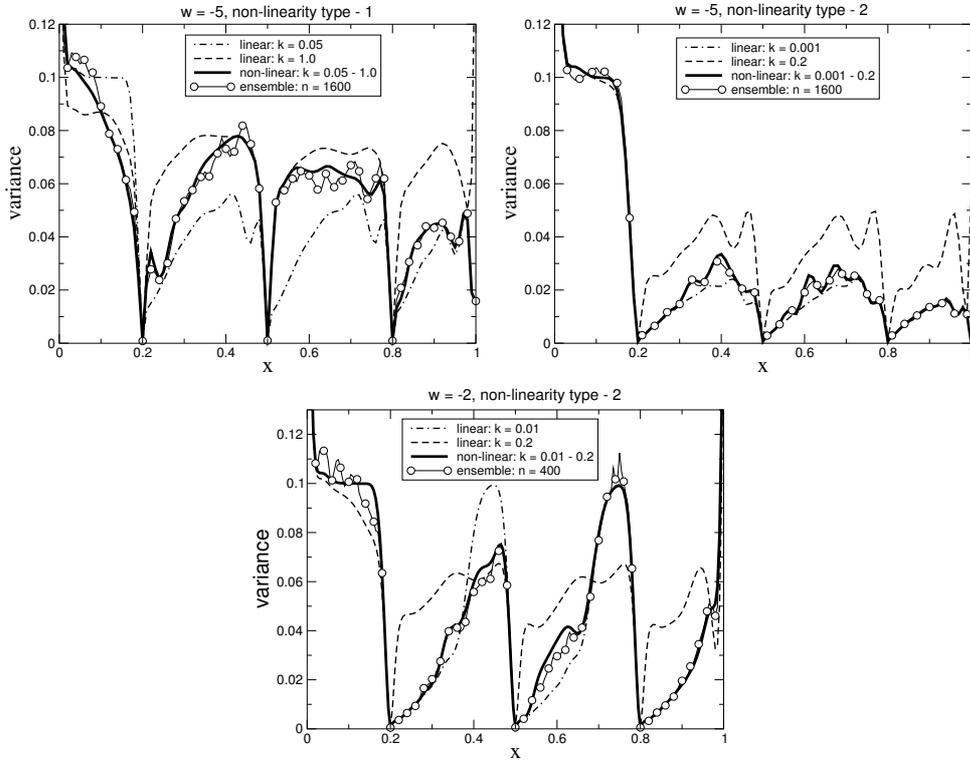


FIG. 7.5.  $H$ -variance and ensemble variance. Upper(left) - case A. Upper(right) - case B, Lower - case C.

Let us assume that we know a 'true' initial condition. Even in this case we can get only an approximation of the covariance  $V_{\delta u}$ . It is usually said that this must be a good approximation if the TLH is valid. However, the TLH is a sufficient 'local' condition. From our consideration we may conclude that we actually need the operators  $T_1, T_2$  to be a good approximation of the error transfer operators  $\tilde{T}_1, \tilde{T}_2$ . This consideration is different because  $\tilde{T}_1, \tilde{T}_2$  are operators which involve time integration and, therefore, can be stable to local perturbations. Even though this statement is

difficult to prove rigorously, we will try to illustrate it by numerical examples. Here for a nonlinear case, we compute the  $H$ -covariance and compare it to the ensemble covariance  $\hat{V}$ . We present three examples: *case A* - in Fig.7.5(upper/left), *case B* - in Fig.7.5 (upper/right), *case C* - in Fig.7.5(lower). In *case A* we use the constitutive formula (6.4) with parameters  $k_1 = 0.05$ ,  $k_2 = 1.0$ ,  $\varphi_0 = 0.5$ ,  $\Delta = 0.2$ ; the convection rate  $w = -5$ . In *case B* we use (6.5) with parameters  $k_1 = 10^{-3}$ ,  $k_2 = 1.0$ ,  $\varphi_0 = 1.0$ ,  $\Delta = 0.4$ ; the convection rate  $w = -5$ . In *case C* we use (6.5) with parameters  $k_1 = 10^{-2}$ ,  $k_2 = 0.2$ ,  $\varphi_0 = 1.0$ ,  $\Delta = 0.4$ ; the convection rate  $w = -2$ . In *cases A,B* the ensemble size is  $n = 1600$ , in *case C* we take  $n = 400$ . In all figures we show the  $H$ -variance (in solid line) and the ensemble variance computed with (5.9) (in line marked with circles). We also present the linear bounds, i.e. values of  $\sigma^2$  for a linear problem solved for  $k(\varphi) = k_1$  (in dash-dotted line) and for  $k(\varphi) = k_2$  (in dashed line). Naturally, in the nonlinear case the variance is situated between the linear bounds. The  $H$ -covariance matrix which corresponds to *case C* is presented in Fig.7.6(left), the ensemble covariance matrix for  $n = 400$  - in Fig.7.6(right). A good match between the  $H$ -variance and the ensemble variance can be noticed. It is satisfactory for off-diagonal elements given the ensemble size  $n = 400$ . We can see, however, that some distinctive details of the covariance presented by the  $H$ -covariance matrix are not properly resolved in the ensemble covariance.

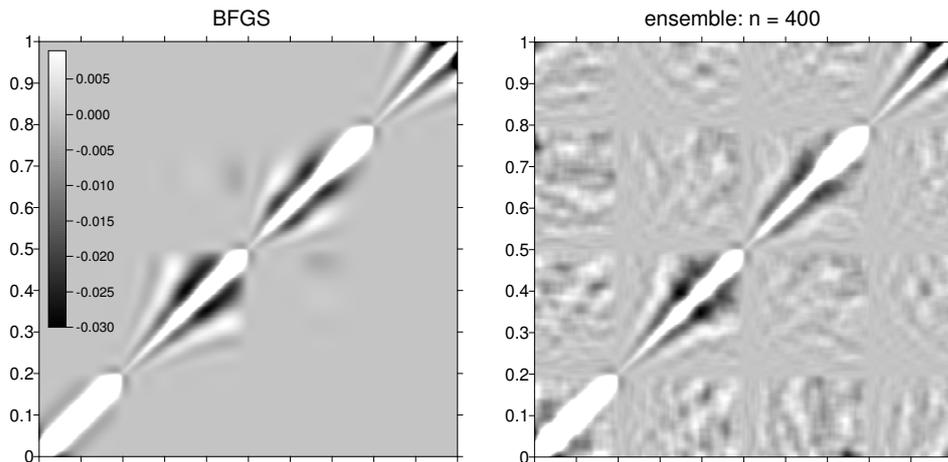


FIG. 7.6.  $H$ -covariance (left) and ensemble covariance (right), for nonlinear case  $C$

This result should be appreciated by taking into account the following fact: parameters in (6.4), (6.5) are chosen such that the nonlinear DA problem (2.2)-(2.3) is on the edge of solvability (as a single-extremum problem, at least). In all cases we had to discard about 15 – 20% of ensemble runs, because the solution process was not properly converging in terms of the gradient norm. A 'critical' set of parameters in (6.4)-(6.5) had been defined beforehand by the 'trial and error' approach. When a certain constitutive model for  $k(\varphi)$  is chosen, we solve the nonlinear DA problem starting from different initial points. Sometimes the iterations either converge to a 'ghost' minima, while the estimate of the unknown initial condition remains far from the 'true' state, or stick in places where the numerical gradient becomes false. We should be reminded at this point that we use a consistent adjoint model, which has passed successfully the gradient test (with the minimum of the classical V-shape curve

reaching about  $10^{-16}$ ).

Another difficulty with the ensemble approach is a rapidly increasing computational cost of a single nonlinear minimization procedure as the nonlinearity increases. First, the forward model itself requires nonlinear iterations at every time step. In *case C* for example, the average number of nonlinear iterations is about 15 – 20 for a 'true' initial condition which is smooth, but this number could be much larger when the current approximation of solution (obtained in the course of minimization) contains sharp field gradients. When the time step is reduced, the number of nonlinear iterations decreases accordingly, so the overall computational cost remains nearly the same. The number of iterations in the nonlinear control loop is 200 – 250. In the same case for the auxiliary DA problem (3.19)-(3.20) the BFGS usually needs about 200 iterations (the same number as the size of the problem), while the nonlinear evolution problem is solved only once and its solution  $\bar{\varphi}$  is saved in memory. In *case C* one run of the BFGS algorithm for the auxiliary (linearized) DA problem is noticeably less expensive (up to  $10^2$  times) than a run for the nonlinear DA problem (which must be done  $n$  times to get an ensemble). For the explicit time discretization scheme, which means a small time integration step, the computational cost of the two methods could

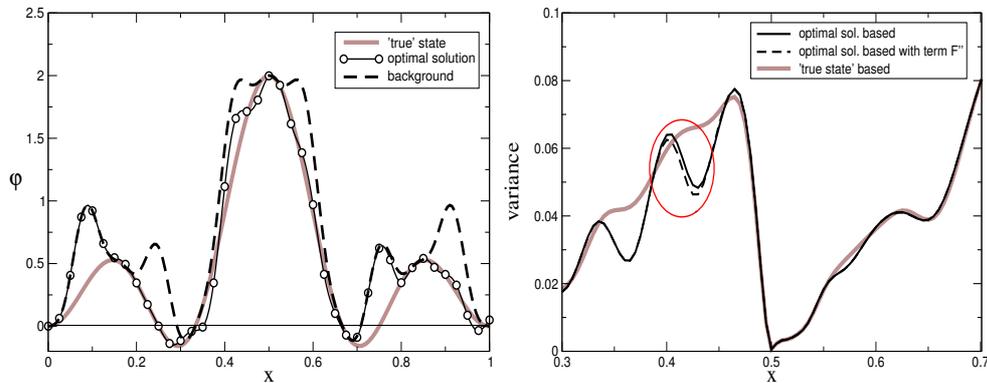


FIG. 7.7. Left - optimal solution. Right -  $H$ -variance based on this optimal solution.

**7.4. On the validity of the tangent linear hypothesis.** We mentioned before that we would like to check how well the TLH is actually satisfied. It requires that the approximation (3.15) is valid for any possible optimal solution  $\varphi_{opt}$ . The solution space is formed by those optimal solutions which correspond to all possible individual implementations of errors  $\xi_1, \xi_2$  given their stochastic properties. We generate an optimal solution (Fig.7.7(left)) using a certain arbitrary implementation of the background error  $\xi_1$ , assuming  $\xi_2 = 0$  (all parameters correspond to *case C*). Here we show the 'true' state (in faint solid line), the background (in dashed line) and the corresponding optimal solution (in solid line marked by circles). For this optimal solutions we compute

$$dF^{(1)} = F(\varphi_{opt}) - F(\bar{\varphi}),$$

where  $F(\cdot)$  is a spatial differential operator in (6.1) and

$$dF^{(2)} = F'(\bar{\varphi})(\varphi_{opt} - \bar{\varphi}),$$

where  $F'(\bar{\varphi})$  is the derivative of  $F$ , or its TLM. If the TLH is valid, then  $dF^{(1)} \approx dF^{(2)}$ .

Functions  $dF^{(1)}$  and  $dF^{(2)}$  for different times are presented in Fig.7.8 in solid and dashed lines correspondingly. These pictures show that the TLH is not actually valid in this particular case. Nevertheless, the  $H$ -covariance and the ensemble covariance (Fig.7.5(lower)) are in good agreement.

It has been mentioned before that we need the operators  $T_1, T_2$  to be a good approximation of the error transfer operators  $\tilde{T}_1, \tilde{T}_2$ . For example, one could impose a condition

$$(7.1) \quad \text{diag}(\eta_1 \eta_1^T + \eta_2 \eta_2^T) \ll \text{diag}(T_1 \xi_1 \xi_1^T T_1^T + T_2 \xi_2 \xi_2^T T_2^T),$$

where

$$\eta_1 = (\tilde{T}_1 - T_1) \xi_1, \quad \eta_2 = (\tilde{T}_2 - T_2) \xi_2,$$

Since  $\tilde{T}_1, \tilde{T}_2, T_1, T_2$  are operators which include time integration, the condition (7.1) might be satisfied even though the TLH is locally (time, space) violated. Furthermore, one actually needs to satisfy (7.1) in an average sense on a set of all possible implementations of  $\xi_1, \xi_2$ . Therefore, instead of (7.1) one may consider

$$(7.2) \quad \text{diag}(E[(\eta_1 \eta_1^T)] + E[\eta_2 \eta_2^T]) \ll \text{diag}(T_1 V_{\xi_1} T_1^T + T_2 V_{\xi_2} T_2^T).$$

The condition (7.2) should be significantly less demanding than (7.1).

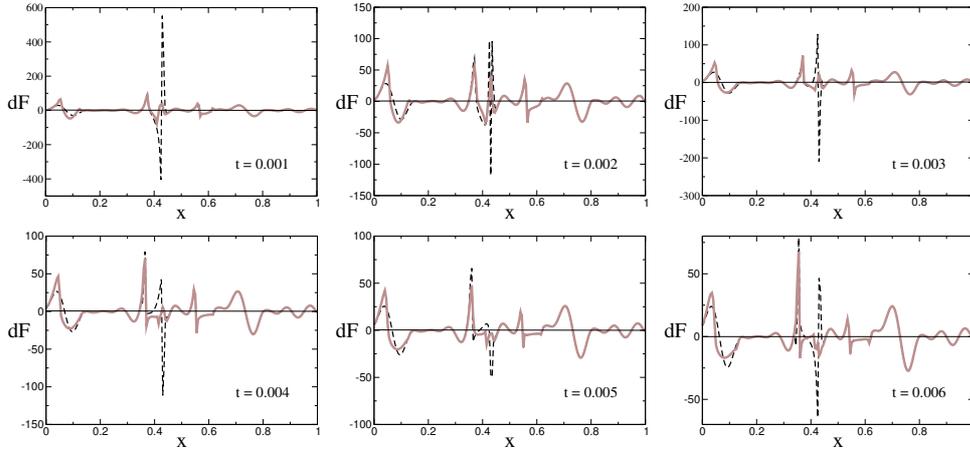


FIG. 7.8. Verification of the tangent linear hypothesis (TLH). Solid lines - nonlinear error, dashed lines - the TLM predicted error.

**7.5. On a 'bias' in the  $H$ -covariance.** For all computations presented above we have assumed that  $\bar{\varphi}$  is a 'true' state evolution. In practice the true initial condition is unknown, but we know a certain optimal solution  $u_{opt}$  which approximates it. Therefore, there is no other option but to use  $\varphi_{opt}$  as  $\bar{\varphi}$  in (3.19)-(3.20). Let us consider the  $H$ -variance computed on a bogus 'true' state as a reference value. We compare it to the  $H$ -variance computed on an optimal solution (we consider the optimal solution presented in Fig.7.7(left), case C).

The result of this comparison is shown in Fig.7.7(right). The Figure presents only the central part of the domain, where the nonlinear effects are strong. We can note that the  $H$ -variance based on the optimal solution (in bold solid line) deviates from the reference value (in faint solid line). This (individual) deviation is essentially larger than the difference between the  $H$ -variance and the ensemble variance (Fig.7.5(lower)), but it is, of course, much smaller than the local differences between  $dF^{(1)}$  and  $dF^{(2)}$  presented in Fig.7.8.

From this example we conclude that the  $H$ -covariance matrix computed on an available optimal solution could be more sensitive to the individual error in this optimal solution ( $u_{opt} - \bar{u}$ ), than to the errors caused by linearization the error transfer operators. This individual error causes a bias in the  $H$ -covariance matrix. A similar bias is also presented in the ensemble covariance matrix.

Finally, by dashed line in Fig.7.7(right) we show the diagonal elements of  $\mathcal{H}^{-1}$ , where  $\mathcal{H}$  is the Hessian of the original nonlinear DA problem (2.2)-(2.3) defined in the Appendix. We compute  $\mathcal{H}$  explicitly as a matrix, then invert this matrix using the SVD technique. As expected,  $diag(\mathcal{H}^{-1})$  is different from the  $H$ -variance (see the area confined by the ellipse) when computed on the optimal solution  $\varphi_{opt}$ . The difference between  $\mathcal{H}$  and  $H$  is the presence of the second order term in  $\mathcal{H}$  (A.4). If  $\mathcal{H}$  is computed on a 'true' state (and  $\xi_2 = 0$ ), then  $\varphi^* = 0$  and this term vanishes, otherwise it is not zero. We notice that in the presence of the second order term the variance computed on  $\varphi_{opt}$  deviates from the reference value even further. This example supports our point of view that the covariance matrix must be approximated by  $H^{-1}$ , not by  $\mathcal{H}^{-1}$ .

**7.6.  $H$ -variance for a simple 'benchmark' non-linear model.** In order to provide an extra check of our method we consider a simple nonlinear evolution model

$$(7.3) \quad x_{i+1} = x_i^{1+\alpha}, \quad i = 0, \dots, N,$$

where  $x$  is a time dependent scalar variable,  $i$  is the time integration index,  $N$  is the number of time steps and  $\alpha > 0$  is a constant parameter. We formulate the initial-value DA problem for the cost functional

$$(7.4) \quad S(x_0) = \frac{1}{2} \sum_{i=0}^N (x_i - \hat{x}_i)^2$$

where  $\hat{x}$  is the observation data. This model had been suggested in [34] to investigate the impact of validity of the TLH on the accuracy of  $\mathcal{H}$  estimated via  $H$ , where  $\mathcal{H}$  is the Hessian of the DA problem related to (7.3)-(7.4),  $H$  is the Hessian of the DA problem for a model obtained by linearizing the equation (7.3). As a measure of impact the authors consider the relative error  $R_H = |1 - \mathcal{H}/H|$ .

We define the relative error in the covariance estimate as  $R_V = |1 - \hat{V}/V|$ , where  $\hat{V}$  is obtained by the fully nonlinear ensemble method for a large ensemble size ( $n = 10^4$ ), and  $V = H^{-1}$  is the  $H$ -covariance. For the parameters used in [34] ( $x_0 = 100$ ,  $N = 144$ ,  $\alpha = 0.0048$ ,  $\sigma_m = 0.15x_N$ ) we have obtained  $R_V \approx 10\%$ , while  $R_H \approx 150\%$  is reported in [34]. This result shows that  $R_V$  cannot be correctly assessed by considering  $R_H$ . It also confirms that  $H^{-1}$  could approximate the actual covariance matrix  $V_{\delta u}$  much better than  $H$  approximates  $\mathcal{H}$ .

**Conclusions.** In this paper we consider the initial-value optimal control problem for a nonlinear evolution model. This is a typical DA problem often considered in

numerical weather prediction and oceanographic applications. The analysis can be used either for now-cast or hind-cast purposes. Correspondingly, the analysis error covariance matrix can be used as a basis for computing the background covariances or as a hind-cast solution quality estimate. The analysis error may be related to the errors of the input data via the error transfer operators without direct use of the tangent linear hypothesis. The approximation (linearization) of the error transfer operators allows the covariance operator to be derived which turns out to be the inverse Hessian of the auxiliary DA problem. This Hessian does not coincide in general with the Hessian of the original nonlinear DA problem. With the use of the quasi-Newton BFGS unconstrained minimization method, a numerical algorithm is constructed to compute an estimate of the analysis error covariance matrix as the inverse Hessian. We refer to this estimate as the ' $H$ -covariance'. The algorithm needs a special choice of initial guess, input functions in the auxiliary DA problem and the exact analytical step search for minimization along the direction of descent. We develop a fully nonlinear ensemble method, which allows us to check the accuracy of the  $H$ -covariance matrix in the nonlinear case. We also suggest a method for generating the background error covariance matrix based on the assumption that the state variable (this is a uni-variate case) belongs to the Sobolev space of a certain power. As a model we use the 1D nonlinear convection-diffusion equation with a strong nonlinearity in the diffusion coefficient. A series of numerical experiments have been performed. Finally, the following conclusions are formulated:

1. In a nonlinear case the covariance operator can be approximated by the inverse Hessian of the auxiliary DA problem based on the TLM constraints ( $H$ -covariance). It is our belief that this result is novel and has not been formulated in an operator form before.
2. In certain circumstances the  $H$ -covariance can be used as an estimate of the covariance operator/matrix far beyond the validity of the tangent linear hypothesis. This is possible because the analysis error and data errors are related by the error transfer operators, which might be fairly robust to the local TLH violations. Although, we accept that our results could still be related to the choice of models.
3. In practice, the exact solution is unknown. Instead, the auxiliary DA problem must be defined using an optimal solution (obtained as a result of DA). This substitution distorts (biases) the  $H$ -covariance matrix. A similar bias is present in the ensemble covariance matrix, i.e. this is a general difficulty common to both methods. The magnitude of the bias could be significantly larger than the linearization error.
4. The correlation radius of the background error proves to be one of the most influential parameters affecting the accuracy of the  $H$ -covariance matrix. A difficulty is that this is, in fact, an a-priori parameter. Thus, a small error in specifying this parameter could yield an error in the  $H$ -covariance much larger than the linearization error.
5. We suggest a fully nonlinear ensemble method as the ultimate test to check the  $H$ -covariance. Even though the idea of this method is very simple, it has not been used before (in this context, at least).
6. The fully nonlinear ensemble method could be very expensive computationally, even in parallel implementation. Besides, in order to resolve properly the off-diagonal elements of  $H^{-1}$ , one needs an ensemble of significant size. The gain of accuracy due to considering the nonlinear error evolution model could be marginal by reasons suggested above (bias, errors in the background error covariance matrix). Therefore, the  $H$ -covariance matrix approach remains a feasible and sufficiently accurate method

when applied to nonlinear problems.

The BFGS algorithm (as presented) can be directly used in small and medium size problems. However, additional progress is needed to develop methods for computing the action  $H^{-1}(u)v$  or the scalar product  $v^T H^{-1}(u)v$  ( $v$  is an arbitrary vector), which are feasible for large scale applications. The scalar product is important in uncertainty analysis, goal-oriented control and for adaptive measurement techniques. In this case one needs to solve the system of equations in the form  $H(u)p = v$ . The following directions to construct a specialized solver for the equation  $H(u)p = v$  could be considered:

- a) the use of a multi-grid strategy;
- b) the use of reduced order models (Proper Orthogonal Decomposition) or local approximations (splines, wavelets);
- c) decomposition of the spatial domain by the 'region of influence' principle, hence decomposition of a global DA problem into a set of local open boundary DA problems.

A particular aim of this paper is to encourage future research on methods for computing actions and scalar products involving  $H$ -covariance matrix or target elements of this matrix that would be feasible for large scale applications.

**Appendix: Hessian via second-order adjoint.** Consider the Hessian  $\mathcal{H}(u)$  of the functional (2.3); it depends on  $u \in X$  (which may be the exact solution, the optimal solution, or some arbitrary function  $u \in X$ ). For a fixed  $u \in X$  the Hessian  $\mathcal{H}(u)$  is defined by the successive solutions of the below-formulated problems. First we find  $\varphi$  and  $\varphi^*$  by solving the direct and adjoint problems (like in the optimality system):

$$(A.1) \quad \begin{cases} \frac{\partial \varphi}{\partial t} = F(\varphi) + f, & t \in (0, T) \\ \varphi|_{t=0} = u, \end{cases}$$

$$(A.2) \quad \begin{cases} -\frac{\partial \varphi^*}{\partial t} - (F'(\varphi))^* \varphi^* = -C^* V_2 (C\varphi - \varphi_{obs}), & t \in (0, T) \\ \varphi^*|_{t=T} = 0. \end{cases}$$

Note that here  $u$  is not necessarily the optimal solution from the optimality system (2.4)–(2.6); it is just some fixed function at which we would like to compute the Hessian. (Hence, in general, the functions  $\varphi$  and  $\varphi^*$  do not satisfy the optimality system). Note also that (A.1)–(A.2) are usual two steps when we compute the gradient of the functional  $S(u)$  (at the point  $u$ ) using the adjoint problem.

Then, the action of the Hessian  $\mathcal{H}(u)$  on the function  $v \in X$  is defined by the successive solutions of the following problems [14]:

$$(A.3) \quad \begin{cases} \frac{\partial \psi}{\partial t} - F'(\varphi)\psi = 0, & t \in (0, T), \\ \psi|_{t=0} = v, \end{cases}$$

$$(A.4) \quad \begin{cases} -\frac{\partial \psi^*}{\partial t} - (F'(\varphi))^* \psi^* = (F''(\varphi)\psi)^* \varphi^* - C^* V_2 C\psi, & t \in (0, T) \\ \psi^*|_{t=T} = 0, \end{cases}$$

$$(A.5) \quad \mathcal{H}(u)v = V_1 v - \psi^*|_{t=0}.$$

Here  $\varphi$  and  $\varphi^*$  are involved, being taken from (A.1)–(A.2). The problem (A.4) is a so-called second-order adjoint problem [14].

*Remark 1.* If  $u$  is the optimal solution, then  $\varphi$  and  $\varphi^*$  are exactly the functions from the optimality system (2.4)–(2.6).

If  $u$  is the exact solution, and  $\xi_2 = 0$ , then  $\varphi$  is the exact solution of (2.4),  $\varphi^* = 0$ , and the second-order term  $(F''(\varphi)\psi)^*\varphi^*$  in (A.4) is omitted.

*Remark 2.* In the finite-dimensional space,  $\mathcal{H}(u)$  is a matrix. To obtain the first column of this matrix, we can choose  $v$  in (A.3)–(A.5) to be the first basis vector  $v = (1, 0, \dots, 0)$ . To obtain the second column of this matrix, we can choose  $v$  in (A.3)–(A.5) to be the second basis vector  $v = (0, 1, 0, \dots, 0)$ , and so on.

**Acknowledgments.** All authors gratefully acknowledge the contribution of the following: Prof. I.M. Navon, SCS, Florida State University for useful discussions, Dr. Graham Copeland, Civil Engineering Department, University of Strathclyde, for his help with the English language, three anonymous Reviewers for their fruitful remarks and suggestions.

#### REFERENCES

- [1] V. I. AGOSHKOV, G. I. MARCHUK, *On solvability and numerical solution of data assimilation problems*, Russ. J. Numer. Anal. Math. Modelling, 8/1 (1993), pp. 1–16.
- [2] J. L. ANDERSON, *Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter*, Physica D: Nonlinear Phenomena, 230 (2007), pp. 99–111.
- [3] R. N. BANNISTER, *How do observations affect background error covariance lengthscales?*, DARC Report, Reading, 2005. (<http://www.met.rdg.ac.uk/ross/DARC/Lengths.pdf>)
- [4] G. CHAVENT, *Local stability of the output least square parameter estimation technique*, Math. Appl. Comp., 2 (1983), pp. 3–22.
- [5] R. DALEY, *Atmospheric Data Analysis*, Cambridge Atmospheric and Space Science Series, Cambridge University Press, 1991.
- [6] G. EVENSEN, *Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics*, J. Geophys. Res., 99/c5 (1994), pp. 10143–10162.
- [7] M. FISHER, P. COURTIER, *Estimating the covariance matrices of analysis and forecast error in variational data assimilation*, ECMWF Research Department Techn. Memo. 220, Reading, 1995.
- [8] P. E. GILL, W. MURRAY, M. H. WRIGHT, *Practical Optimization*, Academic, San Diego, California, 1981.
- [9] I. YU. GEJADZE, V. P. SHUTYAEV, *An optimal control problem of initial data restoration*, Comput. Math. Math. Phys., 39/9 (1999), pp. 1416–1425.
- [10] GREENSTADT J., *Variations on variable-metric methods*, Math. Comp., 24 (1970), pp. 1–22 (Appendix by J. Bard).
- [11] GUNSON J.R., MALANOTTE-RIZZOLI P., *Assimilation studies of open-ocean flows. 2. Error measures with strongly nonlinear dynamics*, J. Geophys. Res., 101/C12 (1996), pp. 28473–28488.
- [12] L. HASCOËT, V. PASCUAL, TAPENADE 2.1 *user's guide*, INRIA Technical Report, 0300 (2004).
- [13] N. B. INGLEBY, *The statistical structure of forecast errors and its representation in the Met. Office global 3-D variational data assimilation scheme*, Quart. J. Roy. Meteorol. Soc., 127 (2001), pp. 209–231.
- [14] F.-X. LE DIMET, I. M. NAVON, D. N. DAESCU, *Second-order information in data assimilation*, Month. Weather Rev., 130/3 (2002), pp. 629–648.
- [15] F.-X. LE DIMET, P. NGNEPIEBA, V. P. SHUTYAEV, *On error analysis in data assimilation problems*, Russ. J. Numer. Anal. Math. Modelling, 17 (2002), pp. 71–97.
- [16] F.-X. LE DIMET, O. TALAGRAND, *Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects*, Tellus, 38A (1986), pp. 97–110.
- [17] F.-X. LE DIMET, V. P. SHUTYAEV, *On deterministic error analysis in variational data assimilation*, Nonlinear Processes in Geophysics, 14 (2005), pp. 1–10.

- [18] F.-X. LE DIMET, V. P. SHUTYAEV, I. YU. GEJADZE, *Analysis error via Hessian in variational data assimilation*, CARI'06 (Colloque Africain sur la Recherche en Informatique): 8th African Conference on Research in Computer Science, Nov. 6-9, 2006, Cotonou, Benin.
- [19] J.-L. LIONS, *Contrôle Optimal des Systèmes Gouvernés par des Équations aux Dérivées Partielles*, Dunod, Paris, 1968.
- [20] D. C. LIU, J. NOCEDAL, *On the limited memory BFGS method for large scale minimization*, Math. Prog. 45 (1989), pp 503–528.
- [21] G. I. MARCHUK, V. I. AGOSHKOV, V. P. SHUTYAEV, *Adjoint Equations and Perturbation Algorithms in Nonlinear Problems*. CRC Press Inc., New York, 1996.
- [22] J. J. MORE, D. J. THUENTE *Line search algorithms with guaranteed sufficient decrease*, ACM Transactions on Mathematical Software, 20/3 (1994), pp 286–307
- [23] I.M. NAVON, *Variational data assimilation, optimal parameter estimation and sensitivity analysis for environmental problems*. In the book: Computational Mechanics'95. Springer, New York, 1995, v.1, pp.740–746.
- [24] B. N. PARLETT, D. S. SCOTT, *The Lanczos algorithm with selective orthogonalization*, Math. Comp., 33/145 (1979), pp .217–238.
- [25] S. V. PATANKAR S.V., *Numerical Heat Transfer and Fluid Flow*, Hemisphere Publishing Corporation, New York, 1980.
- [26] E. POLAK, *Optimization : Algorithms and Consistent Approximations*, (Applied Mathematical Sciences Vol. 124), Springer, New York, 1997
- [27] W. H. PRESS, S. A. TEUKOLSKY, W. T. VETTERLING, B. P. FLANNERY, *Numerical Recipes in Fortran 77: the Art of Scientific Computing*, Second Edition, Cambridge Press, 1992.
- [28] F. RABIER, P. COURTIER *Four-dimensional assimilation in the presence of baroclinic instability*, Quart. J. Roy. Meteorol. Soc., 118 (1992), pp. 649–672.
- [29] W. C. THACKER, *The role of the Hessian matrix in fitting models to measurements*, J. Geophys. Res., 94/C5 (1989), pp. 6177–6196.
- [30] J.-N. THEPAUT, P. COURTIER, *Four-dimensional variational assimilation using the adjoint of a multilevel primitive equation model*, Quart. J. Roy. Meteorol. Soc., 117 (1991), pp. 1225–1254.
- [31] A. N. TIKHONOV, *Solution of incorrectly formulated problems and the regularization method*, English translation of Dokl. Akad. Nauk SSSR, 151 (1963), 501-504
- [32] F. VEERSÉ, *Variable-storage quasi-Newton operators as inverse forecast/analysis error covariance matrices in variational data assimilation*, INRIA Technical Report, 3685 (1999), 28 pp.
- [33] M. ZUPANSKI, *Maximum likelihood ensemble filter: theoretical aspects*, Monthly Weather Review, 133 (2005), pp. 1710–1726
- [34] Y. YANG, I. M. NAVON, P. COURTIER, *A new Hessian preconditioning method applied to variational data assimilation experiments using NASA general circulation model*, Monthly Weather Review, 124 (1996), pp. 1000–1017