



**HAL**  
open science

## Hessian-based covariance approximations in variational data assimilation

Igor Gejadze, Victor P. Shutyaev, François-Xavier Le Dimet

► **To cite this version:**

Igor Gejadze, Victor P. Shutyaev, François-Xavier Le Dimet. Hessian-based covariance approximations in variational data assimilation. *Russian Journal of Numerical Analysis and Mathematical Modelling*, 2018, 33 (1), pp.25-39. 10.1515/rnam-2018-0003 . hal-02068763

**HAL Id: hal-02068763**

**<https://hal.science/hal-02068763v1>**

Submitted on 15 Mar 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Hessian-based covariance approximations in variational data assimilation

Igor Yu. Gejadze<sup>1</sup>, Victor Shutyaev<sup>2</sup>, Francois-Xavier Le Dimet<sup>3</sup>

**Abstract.** The problem of variational data assimilation (estimation) for a nonlinear model is considered in general operator formulation. Hessian-based methods are presented to compute the estimation error covariances. The importance of dynamic formulation and the role of the Hessian and its inverse are discussed.

**Keywords:** variational data assimilation, optimality system, Hessian, estimation error covariance

**MSC:** 65K10, 86A22, 93B40

## 1 Introduction

Data assimilation (state or parameter estimation) for high-dimensional distributed parameter dynamical systems has become a powerful analysis tool in recent decades. It is a fusion process of incomplete and, possibly, indirect observations of state variables with a mathematical model governed by partial differential equations, complemented by *a priori* information. The applications include model initialization in meteorology and oceanography, air and water quality monitoring, 'calibration' of groundwater and reservoir models, discharge estimation and forecasting in river hydraulics and hydrology, flow estimation and control in aerospace engineering, process control in chemical and nuclear engineering, etc. In different applications these estimation problems are also referred as 'inverse problems', 'data assimilation'(DA), and 'calibration'. Variational methods were introduced in meteorology by Sasaki [27]. These methods consider the equations governing the flow as constraints and the problem is closed by using a variational principle, e.g. the minimization of the discrepancy between the model prediction and the observations. Optimal Control Approach (Lions [19], Marchuk [21]) for data assimilation was proposed by Marchuk and Penenko [23], Penenko and Obraztsov [25], Le Dimet [15], Le Dimet and Talagrand [18], Talagrand and Courtier [30]. Therefore, variational estimation/DA is a method based on the optimal control theory, which can also be understood as a special case of the maximum a-posteriori probability (MAP) estimator [8]. This method is preferred for weather and ocean forecasting in major operational centers around the globe, particularly in the form of the incremental 4D-Var [7], and in the form of the *ensemble 4D-Var* [6]. Variational estimation is widely used in other scientific and engineering applications, such as aerospace engineering [2] and astrophysics [4], to mention a few.

---

<sup>1</sup>UMR G-EAU, IRSTEA-Montpellier, 361 Rue J.F. Breton, BP 5095, 34196, Montpellier, France

<sup>2</sup>Institute of Numerical Mathematics, Russian Academy of Sciences, Moscow Institute for Physics and Technology, 119333 Gubkina 8, Moscow, Russia. E-mail: shutyaev@inm.ras.ru

<sup>3</sup>LJK, Université de Grenoble Alpes, BP 51, 38051 Grenoble Cedex 9, France

Uncertainty quantification is an important topic closely associated with estimation and data assimilation. An overview of the original work by the authors on the Hessian-based advanced methodology for computing the estimation error covariances in the framework of variational estimation is presented in the coming sections. The main results are given in a general operator formulation. The accent is made on feasibility of the suggested methods for the uncertainty quantification in high-dimensions, where the statistical methods (Monte Carlo involving associated tricks, e.g. localization, importance sampling, etc.) may not produce a sensible outcome due to a very small sample being available. Some of the approaches could equally be useful for improving (in terms of acceleration, memory savings and robustness) the estimation/DA technique itself.

## 2 Variational data assimilation: basic concept

Let us consider a mathematical (or numerical) model which describes behavior of a natural system in terms of its state variables  $X \in \mathcal{X}$ . Let  $U \in \mathcal{U}$  be the set of the model inputs (controls), then the model can be considered as a 'control-to-state' mapping  $\mathcal{M} : \mathcal{U} \rightarrow \mathcal{X}$ , such that

$$X = \mathcal{M}(U), \quad (2.1)$$

where  $\mathcal{U}$  and  $\mathcal{X}$  are the input and state spaces, correspondingly. For modeling the system behavior the true input vector  $\bar{U}$  must be specified. Under the 'perfect model' assumption the following can be postulated:  $\bar{X} = \mathcal{M}(\bar{U})$ . In reality, some components of  $\bar{U}$  contain uncertainties  $\varepsilon \in \mathcal{U}$ . Thus, instead of  $\bar{U}$  we use its best available approximation (background/prior)

$$U^* = \bar{U} + \varepsilon, \quad (2.2)$$

where  $\varepsilon$  is also called the background error. Because of the presence of  $\varepsilon$ , the predicted state  $X|U^* = \mathcal{M}(U^*)$ , that is,  $X$  evaluated (or conditioned) on  $U^*$ , also contains an error  $\delta X = \mathcal{M}(U^*) - \mathcal{M}(\bar{U})$ .

The state observing tools are represented by an observation operator  $C : \mathcal{X} \rightarrow \mathcal{Y}$  in the form

$$Y = C(X) = C(\mathcal{M}(U)) := G(U), \quad (2.3)$$

where  $G : \mathcal{U} \rightarrow \mathcal{Y}$  is a generalized input-to-observations mapping and  $\mathcal{Y}$  is the 'observation' space. The true observations would be  $\bar{Y} = G(\bar{U})$ , however the actual observations usually contain noise  $\xi$  (observation uncertainty), i.e.

$$Y^* = \bar{Y} + \xi. \quad (2.4)$$

The aim of data assimilation is to obtain  $\hat{U} = U|Y^*$ , i.e. an estimate of  $U$  conditioned on observations  $Y^*$ , which should be better than the prior  $U^*$  in the sense  $\|\hat{U} - \bar{U}\| < \|U^* - \bar{U}\|$ . In the Bayesian framework the posterior probability density of  $U$  conditioned on observations  $Y^*$  is given by the Bayes formula [29]

$$p(U|Y^*) = \frac{p(Y^*|U)p(U)}{p(Y^*)}. \quad (2.5)$$

Looking for the mode of the posterior density  $p(U|Y^*)$ , i.e. maximizing  $p(U|Y^*)$ , is the essence of variational data assimilation (estimation). Under the Gaussian assumption on the prior and observation uncertainties, i.e.  $\varepsilon \sim N(0, B)$ ,  $\xi \sim N(0, R)$ , where  $B$  is the background error covariance and  $R$  - the observation error covariance, maximizing  $p(U|Y^*)$  is equivalent to minimizing the cost-function [31]

$$J(U) = \frac{1}{2} \|R^{-1/2}(G(U) - Y^*)\|_{\mathcal{Y}}^2 + \frac{1}{2} \|B^{-1/2}(U - U^*)\|_{\mathcal{U}}^2. \quad (2.6)$$

Thus, the estimate  $\hat{U}$  is obtained from the necessary optimality condition

$$J'_U(\hat{U}) = 0. \quad (2.7)$$

For the operator  $G(U)$  we define the tangent linear operator  $G'(U)$  (Gateaux derivative) and its adjoint  $(G'(U))^*$  [22] as follows:

$$G'_U(U)w = \lim_{t \rightarrow 0} \frac{G(U + tw) - G(U)}{t}, \quad w \in \mathcal{U}, \quad (2.8)$$

$$(w, (G'_U(U))^*w^*)_{\mathcal{U}} = (G'_U(U)w, w^*)_{\mathcal{Y}}, \quad w^* \in \mathcal{Y}. \quad (2.9)$$

Given the above operator definitions, the full gradient of  $J(u)$  in (2.7) can be expressed in the form:

$$J'_U(U) = (G'_U(U))^*R^{-1}(G(U) - Y^*) + B^{-1}(U - U^*). \quad (2.10)$$

Thus, the estimate  $\hat{U}$  is the solution to the operator equation

$$(G'_U(\hat{U}))^*R^{-1}(G(\hat{U}) - Y^*) + B^{-1}(\hat{U} - U^*) = 0. \quad (2.11)$$

The above presented general operator formulation of DA/estimation problems is common in nonlinear regression and in the inverse problems theory.

In the theory of dynamical systems it is usual to formulate DA/estimation problems in terms of non-stationary partial differential equations (PDE) [19]. Let  $\mathcal{X}$  be the state space such that  $\mathcal{X} = L_2(0, T; \Omega)$ , with a norm  $\|\cdot\|_{\mathcal{X}} = (\cdot, \cdot)_{\mathcal{X}}^{1/2}$ , where  $\Omega$  is a bounded domain of the natural space  $R^d$  ( $d = 1, 2$  or  $3$ ). For the initial state control problem considered in [12] one can write

$$\begin{cases} \frac{\partial \varphi}{\partial t} = F(\varphi) + f, & t \in (0, T) \\ \varphi|_{t=0} = u, \end{cases} \quad (2.12)$$

where  $\varphi = \varphi(x, t)$  (analog of  $X$ ) is the unknown function belonging for any time instant  $t$  to  $L_2(\Omega)$ ,  $x \in R^d$  is the spatial variable,  $u(x) \in L_2(\Omega)$ ,  $f(x) \in \mathcal{X}$  are the initial state and the source term, respectively, and  $F(x)$  is a nonlinear operator mapping  $L_2(\Omega)$  into  $L_2(\Omega)$ . Suppose that for a given pair  $(u, f)$  there exists a unique solution  $\varphi \in \mathcal{X}$  to (2.12). Next, we introduce the functional

$$J(u) = \frac{1}{2} \|R^{-1/2}(C(\varphi) - Y^*)\|_{\mathcal{Y}}^2 + \frac{1}{2} \|B^{-1/2}(u - u^*)\|_{L_2(\Omega)}^2, \quad (2.13)$$

where  $u^*$  is a prior (background) of  $u$ . Thus, operator  $G$  is defined by the formula  $G(u) = C(\varphi)$ , where  $\varphi$  is the solution to equation (2.12). Now, consider the following data assimilation problem with the aim to identify the initial condition: find  $u \in L_2(\Omega)$  and  $\varphi \in \mathcal{X}$  such that

they satisfy (2.12), and on the set of solutions to (2.12), the functional  $J(u)$  takes the minimum value, i.e.

$$\begin{cases} \frac{\partial \varphi}{\partial t} = F(\varphi) + f, & t \in (0, T) \\ \varphi|_{t=0} = u, \\ J(u) = \inf_{v \in L_2(\Omega)} J(v). \end{cases} \quad (2.14)$$

The necessary optimality condition reduces the problem (2.14) to the following optimality system [19], [1],[24]:

$$\begin{cases} \frac{\partial \varphi}{\partial t} = F(\varphi) + f, & t \in (0, T) \\ \varphi|_{t=0} = \hat{u}, \end{cases} \quad (2.15)$$

$$\begin{cases} -\frac{\partial \varphi^*}{\partial t} - (F'(\varphi))^* \varphi^* = -(C'(\varphi))^* R^{-1}(C(\varphi) - Y^*), & t \in (0, T) \\ \varphi^*|_{t=T} = 0, \end{cases} \quad (2.16)$$

$$B^{-1}(\hat{u} - u^*) - \varphi^*|_{t=0} = 0 \quad (2.17)$$

with the unknowns  $\varphi, \varphi^*, \hat{u}$ , where  $(F'(\varphi))^*$  is the adjoint to the Frechet derivative of  $F$ , and  $(C'(\varphi))^*$  is the adjoint to the Frechet derivative of  $C$  defined by

$$(C'(\varphi)\eta, \psi)_Y = (\eta, (C'(\varphi))^*\psi)_X, \quad \eta \in \mathcal{X}, \psi \in \mathcal{Y}.$$

We assume that the system (2.15)–(2.17) has a unique solution. It is easy to see that this system is a detailed representation of the estimator equation (2.11) in the case the model  $\mathcal{M}$  in (2.1) is a dynamical model and the input  $U = \{u\}$  is the initial state  $\varphi(x, 0)$ . Obviously, the general results obtained for (2.11) are valid for the system (2.15)–(2.17). However, the latter is used for better understanding of the structure of adjoint operators, which is sometimes vital for practical implementation of the variational DA method.

### 3 Estimation error covariance and the inverse Hessian

#### 3.1 Clarification of the existing theory

Here, we use the main results from [12] and [13] using the general operator formulation, whereas in papers the dynamic formulation is used. The latter has its own value (see Section 3.3).

Let us consider an estimation error  $\delta U = \hat{U} - \bar{U}$  and equation (2.11). We notice that

$$G(\hat{U}) - Y^* = G(\hat{U}) - (G(\bar{U}) + \xi) = G'_U(\tilde{U})\delta U - \xi,$$

where  $\tilde{U} = \bar{U} + \tau\delta U$ ,  $\tau \in [0, 1]$ , and

$$\hat{U} - U^* = (\hat{U} - \bar{U}) - (U^* - \bar{U}) = \delta U - \varepsilon.$$

Then, equation (2.11) yields the error equation

$$(G'_U(\hat{U}))^* R^{-1}(G'_U(\tilde{U})\delta U - \xi) + B^{-1}(\delta U - \varepsilon) = 0, \quad (3.18)$$

from where the estimation error can be expressed:

$$\delta U = \left( (G'_U(\hat{U}))^* R^{-1} G'_U(\tilde{U}) + B^{-1} \right)^{-1} \left( (G'_U(\hat{U}))^* R^{-1} \xi + B^{-1} \varepsilon \right). \quad (3.19)$$

Using the first order approximations for operators  $G'_U(\hat{U}) \approx G'_U(\tilde{U}) \approx G'_U(\bar{U})$  we express  $\delta U$  as follows:

$$\delta U \simeq H^{-1}(\bar{U}) \left( (G'_U(\bar{U}))^* R^{-1} \xi + B^{-1} \varepsilon \right), \quad (3.20)$$

where

$$H(\cdot) = (G'_U(\cdot))^* R^{-1} G'_U(\cdot) + B^{-1}. \quad (3.21)$$

Here,  $(\cdot)$  denotes a placeholder for the argument of operators  $G$  and  $H$ , which shall be called the 'origin'.

It has been shown that  $H(\cdot)$  represents the first-order term in the Hessian of the cost-function  $J(U)$  in (2.6) or, otherwise, is the Hessian of the auxiliary cost-function as follows:

$$\mathcal{J}(\cdot, \delta U) = \frac{1}{2} \|R^{-1/2} (G'_U(\cdot) \delta U - f_1)\|_{\mathcal{Y}}^2 + \frac{1}{2} \|B^{-1/2} (\delta U - f_2)\|_{\mathcal{U}}^2, \quad (3.22)$$

where  $f_1 \in \mathcal{Y}$  and  $f_2 \in \mathcal{U}$  are trial functions, including  $f_1 = 0$  and  $f_2 = 0$ . We assume that  $H$  is positive definite and, hence, invertible. If the errors  $\xi$  and  $\varepsilon$  indeed satisfy the conditions  $\varepsilon \sim N(0, B)$ ,  $\xi \sim N(0, R)$ , then the estimation error covariance is

$$P = E[\delta U \delta U^T] \simeq H^{-1}(\bar{U}). \quad (3.23)$$

The above relationship is exact for linear  $G$ ; for nonlinear  $G$  it is valid if the tangent linear hypothesis is valid. Since in reality the true value  $\bar{U}$  may not be known, one uses the particular estimate (event)  $\hat{U}$  instead, i.e.

$$P \simeq H^{-1}(\hat{U}). \quad (3.24)$$

There are several papers presenting essentially the same result for dynamical models [26, 32]. However, there has also been some confusion. For example,  $H(\cdot)$  is often perceived as a Hessian of the cost-function  $J(U)$  in (2.6), whereas it is the Hessian of the auxiliary cost-function  $\mathcal{J}(U)$  in (3.22). Another important point is that  $H^{-1}(\bar{U})$  is the best possible approximation to  $P$  using  $H$ , whereas  $H(\hat{U})$  is only an approximation to  $H(\bar{U})$ . The latter fact has never been clearly underlined before.

## 3.2 Computing the inverse Hessian using the LBFSGS

First, we define the preconditioned (projected) Hessian in the form

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

which is usually much better conditioned than  $H$ . Having computed  $\tilde{H}^{-1}$  in the limited-memory form  $H^{-1}$  is recovered as follows:

$$H^{-1} = B^{1/2} \tilde{H}^{-1} (B^{1/2})^*.$$

It is easy to see that  $\tilde{H}$  is the Hessian of the following auxiliary cost-function

$$\tilde{\mathcal{J}}(\cdot, \delta U) = \frac{1}{2} \|R^{-1/2}(G'_U(\cdot)B^{1/2}\delta U - f_1)\|_y^2 + \frac{1}{2} \|\delta U - f_2\|_u^2. \quad (3.25)$$

We use the LBFGS method [20] for generating the limited-memory representation of  $\tilde{H}^{-1}$  as the by-product of minimization of  $\tilde{\mathcal{J}}(\cdot, \delta U)$ . A set of *secant* pairs of a given (prescribed) size is accumulated during minimization, then the product  $\tilde{H}^{-1}v$  is recovered by the BFGS recursion. The key improvement of the quality of approximation (for a given number of secant pairs) has been achieved by using the exact step search, as described in Section 5.1 in [12]. The optimized (in terms of computational expenses) version of this algorithm is presented in our later paper [11], §6.2.

The Lanczos method is routinely used for computing  $\tilde{H}^{-1}$ , whereas the BFGS or LBFGS are usually seen as inferior to the Lanczos. Indeed, for the same accuracy of approximation achieved, representation using the BFGS update formula is less compact than the eigenvalue representation. The eigenvalue analysis, however, could be noticeably more expensive to compute (in terms of function calls). Besides, control of computational costs in the LBFGS algorithm is easier than in the Lanczos method.

We can cite only [35] (preceding to [12, 13] in time) where the same technique is considered. The main focus in this paper is, however, on methods for computing the inverse of the analysis error covariance  $P$ , given the BFGS representation of  $H^{-1}$  obtained in the inner loop of the incremental approach. At the same time, matrix  $H^{-1}$  itself is not of interest and its accuracy is not assessed.

### 3.3 On the importance of dynamic formulation

Let us note that in papers [12] and [13] the dynamic formulation is used. For example in the initial state control problem, the Hessian-vector product is defined by the successive solution of the tangent linear (TL) and adjoint models as follows

$$\begin{cases} \frac{\partial \psi}{\partial t} - F'(\varphi)\psi &= 0, \quad t \in (0, T), \\ \psi|_{t=0} &= v, \end{cases} \quad (3.26)$$

$$\begin{cases} -\frac{\partial \psi^*}{\partial t} - (F'(\varphi))^*\psi^* &= -(C'(\varphi))^*R^{-1}C'(\varphi)\psi, \quad t \in (0, T) \\ \psi^*|_{t=T} &= 0, \end{cases} \quad (3.27)$$

$$H(u)v = B^{-1}v - \psi^*|_{t=0}, \quad (3.28)$$

where  $\varphi(u)$  is the solution of equation (2.12). These equations correspond to equation (3.21) with  $U = \{u\}$ . The dynamic formulation is important because it shows different implementation options. For example, if operator  $G$  is considered as a black box, then operators  $G'_U(\cdot)$  and  $(G'_U(\cdot))^*$  can be obtained by means of the Automatic Differentiation (AD) following the 'discretize-then-optimize' approach. In one hand, the forward, tangent linear and adjoint models obtained this way are mutually 'consistent'. This is a vital property required for minimization of realistic high-dimensional models. On the other hand, the models generated by means of the AD are often computationally cumbersome and should be optimized manually.

For computing  $H^{-1}$  one must compute the product  $Hv$  several times at the same origin point (model trajectory). Thus, the trajectory may be computed only once. In practice, computing the solution and the perturbation in the AD produced TL model is done 'simultaneously', i.e. one line of the TL solver follows the corresponding line of the direct solver. Moreover, in case of nonlinear iterations (to converge on  $\varphi$ ), the TL solver follows step by step the iterative procedure, computing the perturbation variable by increments. This is highly inefficient. For computing the Hessian, however, it might be sufficient to have full consistency between the TL and adjoint models only. The forward and TL models can be partially consistent, i.e. consistent in terms of the discretized operator  $F(\cdot)$ , however the time integration method could be different. In this case the model's trajectory once computed and stored in memory can be supplied into the TL solver. Since the TL model is linear, the nonlinear treatment involved with the forward solver is not involved with the TLM. Thus, one can obtain a computationally efficient TL model. By applying the AD to the auxiliary cost-function (3.22), one can get the code which computes its gradient. The latter defines the product  $Hv$  according to (3.25). The code involves the adjoint model consistent with the modified TLM. Let us note that this adjoint may not be suitable for solving the DA problem itself.

For the parameter estimation (control) problem considered in [13], the mathematical model is described by the evolution problem

$$\begin{cases} \frac{\partial \varphi}{\partial t} = F(\varphi, \lambda) + f, & t \in (0, T) \\ \varphi|_{t=0} = u, \end{cases} \quad (3.29)$$

where  $u$  is the known initial condition,  $\lambda \in \mathcal{X}_p$  is an unknown model parameter and  $\mathcal{X}_p$  is the parameter space.

Let us introduce the functional

$$J(\lambda) = \frac{1}{2} \|R^{-1/2}(C(\varphi) - Y^*)\|_y^2 + \frac{1}{2} \|B^{-1/2}(\lambda - \lambda^*)\|_{\mathcal{X}_p}^2, \quad (3.30)$$

where  $\lambda^* \in \mathcal{X}_p$  is a prior (background) of  $\lambda$ , and  $B$  is the covariance of  $\delta\lambda = \lambda^* - \bar{\lambda}$ , and consider the following DA problem with the aim to estimate the parameter  $\lambda$ : for given  $(u, f)$ , find  $\lambda \in \mathcal{X}_p$  and  $\varphi \in \mathcal{X}$  such that they satisfy (3.29), and on the set of solutions to (3.29), the functional  $J(\lambda)$  takes the minimum value, i.e.

$$\begin{cases} \frac{\partial \varphi}{\partial t} = F(\varphi, \lambda) + f, & t \in (0, T) \\ \varphi|_{t=0} = u, \\ J(\lambda) = \inf_{v \in \mathcal{X}_p} J(v). \end{cases} \quad (3.31)$$

The Hessian-vector product is defined by the successive solution of the tangent linear and adjoint models as follows:

$$\begin{cases} \frac{\partial \psi}{\partial t} - F'_\varphi(\bar{\varphi}, \bar{\lambda})\psi = F'_\lambda(\bar{\varphi}, \bar{\lambda})v, & t \in (0, T), \\ \psi|_{t=0} = 0, \end{cases} \quad (3.32)$$

$$\begin{cases} -\frac{\partial \psi^*}{\partial t} - (F'_\varphi(\bar{\varphi}, \bar{\lambda}))^* \psi^* = -(C'(\bar{\varphi}))^* R^{-1} C'(\bar{\varphi})\psi, & t \in (0, T) \\ \psi^*|_{t=T} = 0, \end{cases} \quad (3.33)$$



$$H(\hat{\lambda})v = B^{-1}v - (F'_\lambda(\bar{\varphi}, \bar{\lambda}))^* \psi^*. \quad (3.34)$$

These equations correspond to equation (3.21) with  $U = \{\lambda\}$ . From (3.26)-(3.28) and (3.32)-(3.34) one can understand the difference between  $G'_u(\cdot)$  and  $G'_\lambda(\cdot)$ . That is, there exists a basic TL equation (system of PDE-based equations) which describes the evolution of perturbations through the system. However, the perturbations are applied at different entries: at the initial condition (case of the initial value control), and as a source term (case of parameter control). The adjoint equations coincide, the expression for the gradient expressed via the adjoint variable  $\psi^*$  is different. This structure suggests an easy way of modifying the system (3.26)-(3.28) into (3.32)-(3.34). It is important to keep consistency between operator  $F'_\lambda(\cdot, \cdot)$  and its adjoint. Since  $F'_\lambda(\cdot, \cdot)$  is not acting on  $\psi$ , it may not necessarily be consistent with  $F'_\varphi(\cdot, \cdot)$ . All these details are important to design a computationally efficient TL and adjoint models for computing the Hessian-vector product (possibly not suitable for computing the gradient of  $J(U)$ ).

### 3.4 On the role of the Hessian and its inverse

The importance of the Hessian matrix and its inverse in variational DA for geophysical applications is underlined in [32], although this has been a well-known fact in statistics for decades (see, for example, [3]). The previous Section illustrates that for linear and moderately non-linear DA/estimation problems  $H^{-1}(\cdot)$  can serve as an approximation of the estimation (analysis) error covariance matrix. In particular, confidence intervals for the components of the estimate  $\hat{U}$  can be defined by the corresponding diagonal elements (variance) of  $H^{-1}(\hat{U})$ . A column  $c_i$  of  $\mathcal{H}^{-1}(\cdot)$  which includes the  $i^{\text{th}}$  diagonal element can be obtained by solving the equation  $H(\cdot)c_i = e_i$  (where  $e_i$  is a Euclidean unit vector). If the number of requested diagonal elements is significant, it would be much less expensive to evaluate  $H^{-1}(\cdot)$  once and keep it in some limited-memory form, than to retrieve necessary diagonal elements using the Hessian-vector product rule.

In addition,  $H^{-1}(\cdot)$  is involved in several other aspects of statistical pre- and post-processing and design of DA systems. Firstly, as the Hessian  $H(\cdot)$  is equivalent to the Fisher information matrix (up to a constant multiplier), the diagonal elements of the inverse Hessian can also be used in the context of optimal experimental design involving such optimality criteria as  $l$ -optimality [34], for example. Secondly, the analysis probability density function (pdf) is defined by the estimate  $\hat{U}$  and the estimation error covariance. Random functions from the Gaussian distribution  $\mathcal{N}(\hat{U}, H^{-1}(\hat{U}))$  can therefore be used as ‘particles’ of the ensemble of initial states, which may be useful for ensemble forecasting [33]. These functions can be generated using  $U = \hat{U} + H^{-1/2}(\hat{U})\xi$ , where  $\xi \sim \mathcal{N}(0, I)$ , or using the eigenvalues of  $H(\hat{U})$  [10]. However, in highly non-linear cases the ‘particles’ generated using this approach are unlikely to belong to the true posterior distribution, thus one must solve perturbed DA problems. This approach is referred to as the *fully non-linear ensemble method* [12], or randomised maximum likelihood method [5]. In these cases, an approximation of  $H^{-1/2}(\hat{U})$  can be used for preconditioning the non-linear minimization process to accelerate convergence, often with impressive results. Thirdly, the analysis error  $\delta U$  and the data errors  $\varepsilon$  and  $\xi$  are related via the approximate error equation (3.20), which can be used as a meta-model for investigating the effects of non-Gaussian data errors on the estimation error pdf. Lastly, if the model depends on parameters  $\theta \in \Theta$ , where  $\Theta$  is the parameter space, an important problem is to evaluate the sensitivity of the

estimation error to uncertainty  $\xi_\theta$  in these parameters. This can be done using the relationship

$$H(\cdot)\delta U = (G'_U(\cdot))^* R^{-1} G'_\theta(\cdot) \xi_\theta.$$

Once again,  $H$  must be inverted to obtain  $\delta U$ .

Another application is related to the use of  $H$  as a coefficient matrix (and preconditioner) in incremental 4D-Var [7]. Here each step of an outer iterative Gauss-Newton process is of the form  $U^{j+1} = U^j + \alpha^j \delta U^j$ , with a discrete approximation  $U^j$  of the unknown initial state at iteration  $j$ , descent step  $\alpha^j$  and update (descent direction)  $\delta U^j$ . As the update satisfies

$$H(U^j)\delta U^j = -G(U^j), \quad (3.35)$$

where  $G(U^j)$  is the gradient of the cost function, a system of linear equations involving  $H$  has to be solved at each step. Given a Hessian-vector product evaluation routine, the systems in (3.35) are usually solved iteratively using, for example, the conjugate gradient (CG) algorithm. An approximation of  $H^{-1}$  or  $H^{-1/2}$ , if available at a reasonable cost, can therefore be used to precondition equation (3.35) to accelerate convergence of this inner iteration.

## 4 Effective inverse Hessian method

### 4.1 Theory of the method

Here, we use the main results from [11] and [28]. As mentioned above, the DCA error contains two components: linearization error component and the origin error component. First, we will see how the linearization component can be reduced.

Let us first consider a related abstract problem. Let  $g$  be a random vector of dimension  $n$  defined as follows

$$g = A(u, \eta)\xi,$$

where  $\xi$  is a Gaussian random vector of dimension  $n$  with zero mean and the covariance matrix  $B_\xi$ ,  $\eta(\xi)$  is a random vector with elements which are linearly or nonlinearly dependent on  $\xi$ ,  $u$  is a given non-random vector and  $A$  is a  $n \times n$  matrix with elements  $a_{i,j}$  nonlinearly dependent on  $u$  and  $\eta$ . Let us assume that the matrix  $A$  can be presented in the form

$$A(u, \eta) = A(u, 0) + \delta A(u, \eta),$$

where  $A(u, 0)$  is the non-random part of  $A$  and  $\delta A(u, \eta)$  is the random part. The covariance matrix of the vector  $g$  is

$$P_g = E \left[ (A(u, \eta)\xi - E[g])(A(u, \eta)\xi - E[g])^T \right]. \quad (4.36)$$

One can see that if  $\eta$  does not depend on  $\xi$ , i.e. vectors  $\eta$  and  $\xi$  are not correlated, then  $E[g] = 0$ , and  $E \left[ A(u, \eta)\xi\xi^T A^T(u, \eta) \right] = E \left[ A(u, \eta)E[\xi\xi^T]A^T(u, \eta) \right]$ . For correlated vectors  $\eta$  and  $\xi$  (since we assume that  $\eta = \eta(\xi)$ ), the following proposition can be put forward. Matrix  $E \left[ A(u, \eta)E[\xi\xi^T]A^T(u, \eta) \right] = E \left[ A(u, \eta)B_\xi A^T(u, \eta) \right]$  is a better approximation to  $P_g$  in (4.36) than  $A(u, 0)B_\xi A^T(u, 0)$ , both in terms of the Frobenius and Riemann distances. The approximation accuracy increases when the correlation level between  $\eta$  and  $\xi$  decreases. The justification of this proposition is provided in [11], and its validity has been analyzed numerically.

Let us consider equation (3.19) in the form

$$\delta U = \left( \mathcal{H}(\bar{U}, \delta U, \tau) \right)^{-1} \left( (G'_U(\bar{U} + \delta U))^* R^{-1} \xi + B^{-1} \varepsilon \right), \quad (4.37)$$

where

$$\mathcal{H}(\bar{U}, \delta U, \tau) = (G'_U(\bar{U} + \delta U))^* R^{-1} G'_U(\bar{U} + \tau \delta U) + B^{-1}. \quad (4.38)$$

Since the nonlinear least-squares estimator is asymptotically unbiased, we assume that  $E[\delta U]$  is small and, therefore, for the covariance  $P$  we have an expression as follows:

$$\begin{aligned} P := E[\delta U \delta U^T] &= E \left[ \mathcal{H}^{-1} (G'_U)^* R^{-1} \xi \xi^T R^{-1} G'_U (\mathcal{H}^{-1})^* \right] \\ &+ E \left[ \mathcal{H}^{-1} B^{-1} \varepsilon \varepsilon^T B^{-1} (\mathcal{H}^{-1})^* \right] \\ &+ E \left[ \mathcal{H}^{-1} B^{-1} \varepsilon \xi^T R^{-1} G'_U (\mathcal{H}^{-1})^* \right] \\ &+ E \left[ \mathcal{H}^{-1} (G'_U)^* R^{-1} \xi \varepsilon^T B^{-1} (\mathcal{H}^{-1})^* \right]. \end{aligned} \quad (4.39)$$

As discussed in the beginning of this Section, we approximate the products  $\xi \xi^T$ ,  $\varepsilon \varepsilon^T$ ,  $\xi \varepsilon^T$  and  $\varepsilon \xi^T$  in (4.39) by  $E[\xi \xi^T] = R$ ,  $E[\varepsilon \varepsilon^T] = B$ ,  $E[\xi \varepsilon^T] = 0$  and  $E[\varepsilon \xi^T] = 0$ , respectively. Thus, we write an approximation of  $P$  as follows:

$$P = E \left[ \mathcal{H}^{-1}(\bar{U}, \delta U, \tau) \left( (G'_U(\bar{U} + \delta U))^* R^{-1} G'_U(\bar{U} + \delta U) + B^{-1} \right) (\mathcal{H}^{-1}(\bar{U}, \delta U, \tau))^* \right], \quad (4.40)$$

where the expression in the round brackets is the Hessian (3.21) at the point  $\bar{U} + \delta U$ . Therefore, the equation (4.40) can be written in the form

$$P = E \left[ \mathcal{H}^{-1}(\bar{U}, \delta U, \tau) H(\bar{U} + \delta U) (\mathcal{H}^{-1}(\bar{U}, \delta U, \tau))^* \right]. \quad (4.41)$$

The value  $\tau = 1/2$  is the optimal one to achieve the best approximation of a difference  $E[G(\bar{U} + \delta U) - G(\bar{U})]$  by  $E[G'_U(\bar{U} + \tau \delta U) \delta U]$ , for  $\delta U \sim N(0, P)$ . In this case, however, one must deal with operator  $\mathcal{H}(\bar{U}, \delta U, \tau = 1/2)$ , which is neither symmetric, nor positive definite, which may seriously complicate its eigenvalue analysis and the subsequent limited-memory representation. Besides, the double product formula is sensitive to the errors which result from the already accepted approximations. By assuming  $\tau = 1$  in  $\mathcal{H}(\bar{U}, \delta U, \tau)$  and keeping in mind that  $\mathcal{H}(\bar{U}, \delta U, \tau = 1) := H(\bar{U} + \delta U)$  this formula can be further simplified by the following expression:

$$P = E \left[ H^{-1}(\bar{U} + \delta U) \right]. \quad (4.42)$$

The right-hand side of (4.42) may be called the 'effective inverse Hessian', hence the name of the suggested method. In order to compute  $P$  directly using this equation, the expectation must be substituted by the ensemble mean:

$$P = \frac{1}{L} \sum_{l=1}^L H^{-1}(\hat{U}_l), \quad (4.43)$$

where  $\hat{U}_l$  are elements from the ensemble of estimates  $\{\hat{U}_l\}$ ,  $l = 1, \dots, L$ , being obtained using perturbed data. Obviously, having such ensemble evaluated, the covariance matrix can be computed according to its definition as follows:

$$\hat{P} = \frac{1}{L_s} \sum_{l=1}^{L_s} (\hat{U}_l - \bar{U})(\hat{U}_l - \bar{U})^T. \quad (4.44)$$

The advantage of (4.43) is, however, that for  $L = L_s$  it gives much better approximation of  $P$  than (4.44), particularly for small  $L_s$ .

## 4.2 Key implementation details

### 4.2.1 Preconditioning

Implementation of formula (4.43) implies that  $L$  inverse Hessians have to be computed. This looks like an extremely laborious task, however, computational costs can be drastically reduced by preconditioning. We notice that the inverse Hessians are evaluated at the origin points perturbed around the true input value  $\bar{U}$ , which means that the difference between  $H^{-1}(\bar{U} + \delta U_l)$  and  $H^{-1}(\bar{U})$  can be described by just a few low-rank updates. Thus, we consider a projected Hessian in the form

$$\tilde{H}(\cdot) = (B^{1/2})^* H(\cdot) B^{1/2} = (B^{1/2})^* (G'_U(\cdot))^* R^{-1} G'_U(\cdot) B^{1/2} + I, \quad (4.45)$$

and use  $\tilde{H}^{-1/2}(\bar{U})$  for the second-level preconditioning of  $\tilde{H}(\bar{U} + \delta U_l)$  as follows:

$$\tilde{H}(\bar{U} + \delta U_l) = \tilde{H}^{-1/2}(\bar{U}) \tilde{H}(\bar{U} + \delta U_l) \tilde{H}^{-1/2}(\bar{U}). \quad (4.46)$$

One can see that  $\tilde{H}(\cdot)$  is the Hessian of an auxiliary cost-function

$$\tilde{\mathcal{J}}(\cdot, \delta U) = \frac{1}{2} \|R^{-1/2} (G'_U(\cdot) \tilde{H}^{-1/2}(\bar{U}) B^{1/2} \delta U - f_1)\|_Y^2 + \frac{1}{2} \|\delta U - f_2\|_U^2. \quad (4.47)$$

The limited memory approximation of  $\tilde{H}(\cdot)$  is evaluated as a by-product of minimization of this cost-function by the LBFGS method, then,  $H^{-1}(\bar{U} + \delta U_l)$  is recovered by the formula

$$H^{-1}(\bar{U} + \delta U_l) = B^{1/2} \tilde{H}^{-1/2}(\bar{U}) \tilde{H}^{-1}(\bar{U} + \delta U_l) \tilde{H}^{-1/2}(\bar{U}) (B^{1/2})^*. \quad (4.48)$$

Taking into account (4.43) and (4.48), the expression for the optimal solution error covariance  $P$  reads

$$P = B^{1/2} \tilde{H}^{-1/2}(\bar{U}) \left( \frac{1}{L} \sum_{l=1}^L \tilde{H}^{-1}(\bar{U} + \delta U_l) \right) \tilde{H}^{-1/2}(\bar{U}) (B^{1/2})^*. \quad (4.49)$$

The first step is, therefore, to find a limited-memory approximation of  $\tilde{H}^{-1}(\bar{U})$ . This is done by minimizing the auxiliary cost-function (3.25) using the LBFGS algorithm. Next, having the product  $\tilde{H}^{-1}(\bar{U}) \cdot v$  defined, one can evaluate the leading eigenpairs  $(\{\lambda_k^{(0)}, W_k^{(0)}\}, k = 1, \dots, K_0)$  of  $\tilde{H}^{-1}(\bar{U})$  using the Lanczos algorithm.

Given the leading (maximum magnitude) eigenpairs  $(\{\lambda_k, W_k\}, k = 1, \dots, K)$  of any symmetric operator  $A$ , the limited-memory representation of  $A$  in power  $\beta$  can be constructed as follows:

$$A^\beta \cdot v = I \cdot v + \sum_{k=1}^K (\lambda_k^\beta - 1) W_k (W_k)^* \cdot v, \quad (4.50)$$

In particular, for the square-root  $\tilde{H}^{-1/2}(\bar{U})$  we get:

$$\tilde{H}^{-1/2}(\bar{U}) \cdot v = I \cdot v + \sum_{k=1}^{K_0} ((\lambda_k^{(0)})^{1/2} - 1) W_k^{(0)} (W_k^{(0)})^* \cdot v, \quad (4.51)$$

Let us note that at this factorization stage we do not run PDE models, so the Lanczos method is affordable. The product  $\tilde{H}^{-1/2}(\bar{U}) \cdot v$  is involved in (4.47).

## 4.2.2 Quasi-random approach

The implementation of formulas (4.43) or (4.49) requires a set of estimates  $\hat{U}_l = \bar{U} + \delta U_l$ ,  $l = 1, \dots, L$  to be computed. Each estimate is a solution to the original data assimilation problem for the cost-function (2.6) with perturbed data  $U_l^* = \bar{U} + \xi_l$  and  $Y_l^* = \bar{Y} + \varepsilon_l$ , where  $\xi_l$  and  $\varepsilon_l$  are random events from  $\xi \sim N(0, R)$  and  $\varepsilon \sim N(0, B)$ , respectively. Evaluating such a set could be fairly expensive.

Here, an alternative approach is suggested. If we denote by  $f_{\delta U}$  the multivariate probability density of the estimation error  $\delta U$ , then the equation (4.42) can be re-written in the form

$$P = \int_{-\infty}^{+\infty} H^{-1}(\bar{U} + v) f_{\delta U}(v) dv. \quad (4.52)$$

Since we assume that the estimation error is approximately Gaussian with zero mean and covariance  $P$  we obtain

$$P = \frac{1}{(2\pi)^{M/2} |P|^{1/2}} \int_{-\infty}^{+\infty} H^{-1}(\bar{U} + v) \exp\left(-\frac{1}{2} v^T P^{-1} v\right) dv. \quad (4.53)$$

In contrast to formula (4.43), the above expression is a nonlinear matrix integral equation (deterministic) with respect to  $P$ , while  $v$  is a dummy variable. This equation can be solved by the fixed point iterative process

$$P^{k+1} = \frac{1}{(2\pi)^{M/2} |P^k|^{1/2}} \int_{-\infty}^{+\infty} H^{-1}(\bar{U} + v) \exp\left(-\frac{1}{2} v^T (P^k)^{-1} v\right) dv, \quad (4.54)$$

for  $k = 0, 1, \dots$ , starting with  $P^0 = H^{-1}(\bar{u})$ .

The evaluation of the multi-dimensional integral in (4.54) using the quasi-Monte Carlo method means returning to formula (4.43). Taking into account (4.54), the iterative process takes the form

$$\begin{cases} P^{k+1} &= \frac{1}{L} \sum_{l=1}^L H^{-1}(\bar{U} + \delta U_l^k), \\ P^0 &= H^{-1}(\bar{u}), \quad k = 0, 1, \dots \end{cases} \quad (4.55)$$

where  $\delta U_l^k \sim N(0, P^k)$ .

One can see that for each  $k$  the last formula looks similar to (4.43) with one key difference:  $\delta U_l^k$  in (4.55) is not an estimation error itself, but a vector having the statistical properties of this error. It is generated as follows:

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

where  $\eta_l$  is a quasi-random sequence from  $\mathcal{N}(0, I)$ . This approach is similar to the inflation approach used in statistical estimation methods (ensemble Kalman and particle filtering). If the preconditioning (4.45)-(4.46) is involved, the process (4.55) reads as follows:

$$\begin{cases} \tilde{P}^{k+1} &= \frac{1}{L} \sum_{l=1}^L \tilde{H}^{-1}(\bar{U} + \delta U_l^k), \\ \tilde{P}^0 &= I, \quad k = 0, 1, \dots, \end{cases} \quad (4.56)$$

where  $\delta U_i^k \sim N(0, P^k)$ , and the perturbation is defined as follows:

$$\delta U_i^k = (\tilde{P}^k)^{1/2} \tilde{H}^{-1/2}(\bar{U})(B^{1/2})^* \eta_i, \quad \eta_i \sim N(0, I).$$

Finally, the covariance  $P$  in the original space is recovered by the formula

$$P = B^{1/2} \tilde{H}^{-1/2}(\bar{U}) \tilde{P} \tilde{H}^{-1/2}(\bar{U})(B^{1/2})^*. \quad (4.57)$$

## 5 Estimation error covariance versus Bayesian posterior covariance

### 5.1 General theory

Here, we use the main results from [14], which point out the difference between the classical estimation (analysis) error covariance and the Bayesian posterior covariances, both in theoretical and computational aspects.

The estimation error covariance is associated with trying to find an approximation around the truth  $\bar{U}$ , whereas the data is also assumed to come from the truth:  $Y^* = G(\bar{U}) + \xi$ ,  $U^* = \bar{U} + \varepsilon$ , where  $\xi \sim \mathcal{N}(0, R)$  and  $\varepsilon \sim N(0, B)$  are the observation and background error, respectively. The estimation error is defined as  $\delta U = U - \bar{U}$  and its covariance is given by

$$P = E_a[(U - \bar{U})(U - \bar{U})^T] = E_a[\delta U \delta U^T],$$

where  $E_a$  denotes averaging (expectation) with respect to the probability density function (pdf)  $\rho_a$  which, taking into account the definitions of the data  $Y^*$  and  $U^*$ , can be defined as follows:

$$\rho_a(U) = \text{const} \cdot \exp\left(-\frac{1}{2}\|R^{-1/2}(G(U) - G(\bar{U}))\|_y^2 - \frac{1}{2}\|B^{-1/2}(U - \bar{U})\|_u^2\right). \quad (5.58)$$

The Bayesian posterior covariance is given by

$$\mathcal{P} = E_p[(U - E_p[U])(U - E_p[U])^T], \quad (5.59)$$

where  $E_p$  denotes averaging (expectation) with respect to the pdf

$$\rho_p(U) = \text{const} \cdot \exp\left(-\frac{1}{2}\|R^{-1/2}(G(U) - Y_0^*)\|_y^2 - \frac{1}{2}\|B^{-1/2}(U - U_0^*)\|_u^2\right). \quad (5.60)$$

In the above formula  $Y_0^*$  stands for the actual observations made by an instrument. It could be, in principle, seen in the form  $Y_0^* = G(\bar{U}) + \xi_0$ , where  $\xi_0$  is no longer a random variable, but a particular event (number) from  $\mathcal{N}(0, R)$ . However, this representation is not used in the Bayesian approach. Similarly,  $U_0^*$  stands for the actual background. Then, the estimate  $\hat{U}_0$  satisfies the equation

$$(G'_U(\hat{U}_0))^* R^{-1}(G(\hat{U}_0) - Y_0^*) + B^{-1}(\hat{U}_0 - U_0^*) = 0. \quad (5.61)$$

The posterior covariance is often approximated by trying to find the second moment of  $\rho_p$  centered around  $\hat{U}$  instead of  $E_p[U]$ , which is natural because  $\hat{U}$  is the output of 4D-Var. In the

linear Gaussian setup  $E_p[U]$  and  $\hat{U}$  coincide. This is not true in general, but can be expected to be a good approximation if the volume of data is large and /or noise is small (due to the asymptotic properties of the nonlinear least-square estimator, see [3], Vol. 1, Ch.6). Most importantly, due to different centering of Gaussian data, the Bayesian posterior covariance and the estimation error covariance are different objects and should not be confused. However, they are equal in the linear case.

In order to simulate the distribution (5.60) we define perturbed data as follows:  $Y^* = Y_0^* + \xi$  and  $U^* = U_0^* + \varepsilon$ . Then, the perturbed estimate  $\hat{U}$  satisfies equation

$$(G'_U(\hat{U}))^* R^{-1}(G(\hat{U}) - Y^*) + B^{-1}(\hat{U} - U^*) = 0. \quad (5.62)$$

One way to derive the error equation is to subtract (5.61) from (5.62). For the difference in the residual terms we obtain

$$\begin{aligned} & (G'_U(\hat{U}))^* R^{-1}(G(\hat{U}) - Y^*) - (G'_U(\hat{U}_0))^* R^{-1}(G(\hat{U}_0) - Y_0^*) = \\ & = \{(G'_U(\hat{U}))^* R^{-1}G(\hat{U}) - (G'_U(\hat{U}_0))^* R^{-1}G(\hat{U}_0)\} - \{(G'_U(\hat{U}))^* R^{-1}Y^* - (G'_U(\hat{U}_0))^* R^{-1}Y_0^*\} \\ & = \{(G'_U(\hat{U}))^* R^{-1}G(\hat{U}) - (G'_U(\hat{U}))^* R^{-1}G(\hat{U}_0)\}_1 \\ & + \{(G'_U(\hat{U}))^* R^{-1}(G(\hat{U}_0) - Y_0^*) - (G'_U(\hat{U}_0))^* R^{-1}(G(\hat{U}_0) - Y_0^*)\}_2 \\ & - (G'_U(\hat{U}))^* R^{-1}\xi. \end{aligned} \quad (5.63)$$

Using the Taylor-Lagrange formula [22] the expressions in the brackets take the form

$$\{\cdot\}_1 = (G'_U(\hat{U}))^* R^{-1}G'_U(\tilde{U}_1)\delta U,$$

$$\{\cdot\}_2 = [(G'(\tilde{U}_2))^*]'_U \delta U R^{-1}(G(\hat{U}_0) - Y_0^*),$$

where  $\tilde{U}_i = \hat{U}_0 + \tau_i(\hat{U} - \hat{U}_0)$ ,  $\tau_i \in [0, 1]$ ,  $i = 1, 2$ . For the difference in the penalty terms we obtain

$$B^{-1}(\hat{U} - U^*) - B^{-1}(\hat{U}_0 - U_0^*) = B^{-1}(\delta U - \varepsilon). \quad (5.64)$$

Combining (5.63) and (5.64) we obtain the error equation as follows:

$$(G'_U(\hat{U}))^* R^{-1}(G'_U(\tilde{U}_1)\delta U - \xi) + [(G'(\tilde{U}_2))^*]'_U \delta U R^{-1}(G(\hat{U}_0) - Y_0^*) + B^{-1}(\delta U - \varepsilon) = 0. \quad (5.65)$$

In the operator form this equation can be rewritten as

$$\mathcal{H}(\hat{U}, \tilde{U}_1, \tilde{U}_2) \cdot \delta U = (G'_U(\hat{U}))^* R^{-1}\xi + B^{-1}\varepsilon, \quad (5.66)$$

where

$$\mathcal{H}(\hat{U}, \tilde{U}_1, \tilde{U}_2) \cdot v = ((G'_U(\hat{U}))^* R^{-1}G'_U(\tilde{U}_1) + B^{-1}) \cdot v + [(G'(\tilde{U}_2))^*]'_U \cdot v R^{-1}(G(\hat{U}_0) - Y_0^*). \quad (5.67)$$

Let us note that, generally,  $\mathcal{H}$  is neither symmetric, nor positive definite.

The estimation error is expressed from (5.66) as follows:

$$\delta U = \mathcal{H}^{-1}(\hat{U}, \tilde{U}_1, \tilde{U}_2)((G'_U(\hat{U}))^* R^{-1}\xi + B^{-1}\varepsilon). \quad (5.68)$$

Next, we approximate (5.59) by the formula

$$\mathcal{P} = E_p [(\delta U \delta U^T)]. \quad (5.69)$$

Since operators  $\mathcal{H}(\hat{U}, \tilde{U}_1, \tilde{U}_2)$  and  $G'_U(\hat{U})$  nonlinearly depend on  $\delta U$ , it is not possible to consider them as constant multipliers when applying the expectation operator. One way is to fix  $\hat{U}, \tilde{U}_1, \tilde{U}_2$  at the value  $\hat{U}_0$ . In this case  $\mathcal{H}(\cdot)$  becomes the Hessian of the cost-function (2.6) at the origin point  $\hat{U}_0$ . Next, by applying (5.69) we obtain

$$\mathcal{P} = \mathcal{H}^{-1}(\hat{U}_0) H(\hat{U}_0) \mathcal{H}^{-1}(\hat{U}_0), \quad (5.70)$$

(compare this equation to (3.24)).

Since the condition number of the double product  $\mathcal{H}^{-1} H \mathcal{H}^{-1}$  is usually larger than the condition number of its components  $\mathcal{H}$  and  $H$ , the above formula could be quite sensitive to the approximation errors. One way is to simplify this formula assuming  $\mathcal{H}$  and  $H$  are not very different. These simplification is

$$\mathcal{P} = \mathcal{H}^{-1}(\hat{U}_0), \quad (5.71)$$

and, subsequently,

$$\mathcal{P} = H^{-1}(\hat{U}_0). \quad (5.72)$$

Another approach is to use the 'effective inverse Hessian' method presented in Sect. 4. In this case equation (5.70) becomes

$$\mathcal{P} = \frac{1}{L} \sum_{l=1}^L \mathcal{H}^{-1}(\hat{U}_l) H(\hat{U}_l) \mathcal{H}^{-1}(\hat{U}_l). \quad (5.73)$$

## 5.2 Implementation

As before, the computationally efficient and, therefore, feasible implementation is achieved using preconditioning. The first-level preconditioning yields

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

and the second-level preconditioning yields

$$\tilde{\tilde{H}}(\cdot) = \tilde{H}^{-1/2}(\hat{U}_0) \tilde{H}(\cdot) \tilde{H}^{-1/2}(\hat{U}_0), \quad (5.74)$$

$$\tilde{\tilde{\mathcal{H}}}(\cdot) = \tilde{H}^{-1/2}(\cdot) \tilde{\mathcal{H}}(\cdot) \tilde{H}^{-1/2}(\cdot). \quad (5.75)$$

At the first step we compute the leading eigenpairs  $(\{\lambda_k^{(0)}, W_k^{(0)}\}, k = 1, \dots, K_0)$  of  $\tilde{H}(\hat{U}_0)$ . This allows  $\tilde{H}^{-1/2}(\hat{U}_0)$  to be defined in the limited-memory form. Then, inside the ensemble loop (index  $l$ ) we compute:

a) the leading eigenpairs  $(\{\lambda_k^{(l)}, W_k^{(l)}\}, k = 1, \dots, K_l)$  of  $\tilde{H}(\hat{U}_l)$ , which allows us to define  $\tilde{H}^{-1/2}(\hat{U}_l)$  in (5.75);

b) the leading eigenpairs  $(\{\tilde{\lambda}_k^{(l)}, \tilde{W}_k^{(l)}\}, k = 1, \dots, \tilde{K}_l)$  of  $\tilde{\mathcal{H}}(\hat{U}_l)$ .

For a reasonably small  $\delta U_l = \hat{U}_l - \hat{U}_0$ , the number  $K_l$  of eigenpairs describing the Riemann distance between  $\tilde{H}(\hat{U}_l)$  and  $\tilde{H}(\hat{U}_0)$  has to be small as compared to  $K_0$ . Similarly,  $\tilde{K}_l$



has to be small as compared to  $K_l$  because the difference between  $\tilde{\mathcal{H}}(\hat{U}_l)$  and  $\tilde{H}(\hat{U}_l)$  is only due to the presence of the second-order term. Thus, the expenses of computing the eigenpairs  $\{\lambda_k^{(l)}, W_k^{(l)}\}$  and  $\{\tilde{\lambda}_k^{(l)}, \tilde{W}_k^{(l)}\}$  for given  $l$  can make only a fraction of those associated with computing  $\{\lambda_k^{(0)}, W_k^{(0)}\}$ .

Given all eigenpairs are computed, the posterior covariance in (5.73) is evaluated as follows:

$$\mathcal{P} = B^{1/2} \tilde{H}^{-1/2}(\hat{U}_0) A(\hat{U}_l) \tilde{H}^{-1/2}(\hat{U}_0) (B^{1/2})^*, \quad (5.76)$$

where

$$A(\hat{U}_l) = \frac{1}{L} \sum_{l=1}^L \tilde{H}^{-1/2}(\hat{U}_l) \tilde{\mathcal{H}}^{-\alpha}(\hat{U}_l) \tilde{H}^{-1/2}(\hat{U}_l). \quad (5.77)$$

In (5.77), cases  $\alpha = 2$ ,  $\alpha = 1$  and  $\alpha = 0$  correspond to approximations (5.70), (5.71) and (5.72), respectively. The quasi-random implementation described in Sect.4.2.2 is applied to substitute  $\hat{U}_l$  by  $\hat{U}_0 + \delta U_l$ .

In the dynamic formulation the Hessian-vector product  $\mathcal{H}(\cdot)v$  is defined by the successive solutions of the following problems [12]:

$$\begin{cases} \frac{\partial \psi}{\partial t} - F'(\varphi)\psi = 0, & t \in (0, T), \\ \psi|_{t=0} = v, \end{cases} \quad (5.78)$$

$$\begin{cases} -\frac{\partial \psi^*}{\partial t} - (F'(\varphi))^* \psi^* = (F''(\varphi)\psi)^* \varphi^* - C^* V_2 C \psi, & t \in (0, T) \\ \psi^*|_{t=T} = 0, \end{cases} \quad (5.79)$$

$$\mathcal{H}(u)v = V_1 v - \psi^*|_{t=0}. \quad (5.80)$$

Here  $\varphi$  and  $\varphi^*$  are involved, being taken from (2.15)–(2.16). The problem (5.79) is the so-called second-order adjoint problem [16]. One can see that it is a standard adjoint problem with a specially defined source term  $(F''(\varphi)\psi)^* \varphi^*$ . For certain  $F$ , this term is relatively expensive to compute as compared to  $(F'(\varphi))^* \psi^*$ , which could be a difficulty if the explicit time integration scheme is used. This is not an issue, however, when the implicit or semi-implicit schemes are used.

## 6 Conclusions

The main issue considered in this paper is the relationship between the Hessian of the cost-function and the estimation error covariance matrix in variational DA, whereas the discussion is focused on the essentially nonlinear case. In the classical (frequentist) statistical approach the estimation error  $\delta U$  is considered as a difference between the estimate  $\hat{U}$  and the 'truth'  $\bar{U}$ . The corresponding covariance matrix can be approximated by the inverse of the Hessian  $H$  of an auxiliary cost-function computed at the 'truth' point. If the 'truth' is known (the identical twin experiment framework only), this is the first-order approximation with respect to  $\delta U$ . In practice, the Hessian is computed at the estimate  $\hat{U}_0$ , which is obtained after assimilating the real data. The difference between  $\hat{U}_0$  and  $\bar{U}$  yields an additional error in the covariance (the origin error), which can not be removed in principle. Evaluating the inverse Hessian using the LBFGS algorithm with exact step search has been a novel computational approach.

In the Bayesian approach the estimation error  $\delta U$  is considered as a difference between the estimates  $\hat{U}$  and  $\hat{U}_0$ , conditioned on a perturbed and unperturbed data, respectively. Due to different centering of data, the first-order (with respect to  $\delta U$ ) approximations of the estimation (analysis) error covariance and the Bayesian posterior covariance are different and should not be confused. The latter is computed via the Hessian-product formula (5.70), which involves both the Hessian of the original cost-function  $\mathcal{H}$ , and the Hessian of the auxiliary cost-function  $H$  (often called the Jacobian). In the dynamic formulation the Hessian-vector product  $\mathcal{H}(\cdot)v$  is defined by the second-order adjoint problem. The formulas of the type (5.70) occasionally appear in the literature on statistics and nonlinear regression in the general context of nonlinear least-squares (see e.g. [9]).

Introducing the 'effective' Hessian-based covariance approximations (4.43) and (5.73), as well as the relatively feasible methods for their evaluation, is our main contribution in terms of novelty. The concept of such approximations stems from the exact error equations (3.18) and (5.65), the idea of which, in turn, has been first presented in [17]. In practice, evaluating (4.43) or (5.73) requires an ensemble of the inverse Hessians  $H^{-1}(\hat{U}_l)$ , or the products  $\mathcal{H}^{-1}(\hat{U}_l)H(\hat{U}_l)\mathcal{H}^{-1}(\hat{U}_l)$  to be computed, however the size of such ensemble can be small enough ( $L = 25 - 50$ ). The suggested method can be regarded as an alternative to the standard sampling-based covariance evaluation method (involving localization and other possible tricks). The latter requires much larger ensemble size to achieve a similar approximation quality. The ways of improving the accuracy of the 'effective' Hessian-based covariance approximations are the matter of future research.

## Acknowledgements

This work was carried out within the Russian Science Foundation project 17-77-30001 (studies in Section 3), and the project 15-01-01583 of the Russian Foundation for the Basic Research.

## References

- [1] V.I. Agoshkov, G.I. Marchuk. On solvability and numerical solution of data assimilation problems, *Russ. J. Numer. Anal. Math. Modelling*, 1993, v.8, pp.1–16.
- [2] O.M. Alifanov, E.A. Artyukhin, and S.V. Rumyantsev. *Extreme Methods for Solving Ill-Posed Problems with Applications to Inverse Heat Transfer Problems*. Begel House Publishers, New York - Wallingford (U.K.), 1996.
- [3] T. Amemiya. *Handbook of Econometrics*. North-Holland Publishing Company, Amsterdam, 1983.
- [4] J.M. Bardsley and A. Luttmann. Total variation-penalized poisson likelihood estimation for ill-posed problems. *Advances in Computational Mathematics*, 31(1):35–59, 2009.
- [5] Y. Chen and D. Oliver. Ensemble randomized maximum likelihood method as an iterative ensemble smoother. *Math. Geosci.*, 44:1–26, 2012.

- [6] A.M. Clayton, A.C. Lorenc, and D.M. Barker. Operational implementation of a hybrid ensemble/4d-var global data assimilation system at the met office. *Quart. J. Roy. Meteor. Soc.*, 139:1445–1461, 2013.
- [7] P. Courtier, J.-N. Thepaut, and A. Hollingsworth. A strategy for operational implementation of 4d-var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, 120:1367–1387, 1994.
- [8] M. Dashti, K.J.H. Law, A.M. Stuart, and J. Voss. Map estimators and their consistency in bayesian nonparametric inverse problems. *Inverse Problems*, 2013, 29:17–44.
- [9] J.R. Donaldson and R.B. Schnabel. Computational experience with confidence regions and confidence intervals for nonlinear least squares. *Technometrics*, 29(1):67–82, 1987.
- [10] M. Ehrendorfer and J.J. Tribbia. Optimal prediction of forecast error covariances through singular vectors. *J. Atmos. Sci.*, 54:286–313, 1997.
- [11] I. Gejadze, G.J.M. Copeland, F.-X. Le Dimet, and V. Shutyaev. Computation of the analysis error covariance in variational data assimilation problems with nonlinear dynamics. *Journal of Computational Physics*, 230(22):7923–7943, 2011.
- [12] I. Gejadze, F.-X. Le Dimet, and V. Shutyaev. On analysis error covariances in variational data assimilation. *SIAM J. Sci. Computing*, 30(4):1847–1874, 2008.
- [13] I. Gejadze, F.-X. Le Dimet, and V. Shutyaev. On optimal solution error covariances in variational data assimilation problems. *Journal of Computational Physics*, 229(6):2159–2178, 2010.
- [14] I. Gejadze, V. Shutyaev, and F.-X. Le Dimet. Analysis error covariance versus posterior covariance in variational data assimilation. *Q.J.R. Meteorol. Soc.*, 139:1826–1841, 2013.
- [15] Le Dimet F.X. A general formalism of variational analysis, *Report OK-73091-22-1*, CIMMS, Norman, 1982.
- [16] F.-X. Le Dimet, I.M. Navon, and D. Daescu. Second-order information in data assimilation. *Monthly Weather Review*, 130(3):629–648, 2002.
- [17] F.-X. Le Dimet and V. Shutyaev. On deterministic error analysis in variational data assimilation. *Nonlinear Processes in Geophysics*, 14:1–10, 2005.
- [18] Le Dimet F.X., Talagrand O. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus*, 1986, v.38A, pp.97–110.
- [19] Lions J.L. *Contrôle optimal des systèmes gouvernés par des équations aux dérivées partielles*. – Paris: Dunod, 1968.
- [20] D. C. Liu and J. Nocedal, On the limited memory BFGS method for large scale minimization, *Math. Program.*, 1989, v.45, pp. 503–528.

- [21] Marchuk G.I. Perturbation theory and the statement of inverse problems. *Lect. Notes Comput. Sci.*, 1973, v.4, pp.159–166.
- [22] G.I. Marchuk, V.I. Agoshkov, and V.P. Shutyaev. *Adjoint Equations and Perturbation Algorithms in Nonlinear Problems*. CRC Press Inc., New York, 1996.
- [23] Marchuk G.I., Penenko V.V. Application of optimization methods to the problem of mathematical simulation of atmospheric processes and environment // *Modelling and Optimization of Complex Systems*. Proc. of the IFIP-TC7 conf. – New York: Springer, 1978, pp.240–252.
- [24] G.Marchuk, V.Shutyaev, and V.Zalesny. Approaches to the solution of data assimilation problems. In: *Optimal Control and Partial Differential Equations*. Amsterdam: IOS Press, 2001, 489–497.
- [25] Penenko V., Obraztsov N.N. A variational initialization method for the fields of the meteorological elements. *Soviet Meteorol. Hydrol.*,1976,, v.11, pp.1–11.
- [26] F. Rabier and P. Courtier. Four-dimensional assimilation in the presence of baroclinic instability. *Quart. J. Roy. Meteorol. Soc.*, 118:649–672, 1992.
- [27] Sasaki, Y. An objective analysis based on the variational method. *J. Meteor. Soc. Japan*, 1958, v.36, pp.77-88.
- [28] V. Shutyaev, I. Gejadze, G.J.M. Copeland, and F.-X. Le Dimet. Optimal solution error covariance in highly nonlinear problems of variational data assimilation. *Nonlinear Processes in Geophysics*, 19:177–184, 2012.
- [29] Stuart A.M. Inverse problems: a Bayesian perspective. *Acta Numerica*, 2010, vol.19, pp. 451–559.
- [30] Talagrand, O. and P. Courtier. Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Quart. J. Roy. Meteor. Soc.*, 1987, v.113, pp.1311-1328.
- [31] A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, Philadelphia, 2005.
- [32] W.C. Thacker. The role of the hessian matrix in fitting models to measurements. *J. Geophys. Res.*, 94(C5):6177–6196, 1989.
- [33] Z. Toth and E. Kalnay. Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, pp.2317–2330, 1993.
- [34] D. Ucinski. *Optimal Measurement Methods for Distributed-Parameter System Identification*. CRC Press, Boca Raton, FL, 2005.
- [35] F. Veerse. Variable-storage quasi-newton operators as inverse forecast/analysis error covariance matrices in variational data assimilation. *INRIA Technical Report 3685*, pages 1–28, 1999.