



Available online at www.sciencedirect.com

ScienceDirect

Comput. Methods Appl. Mech. Engrg. 290 (2015) 228–249

**Computer methods
in applied
mechanics and
engineering**

www.elsevier.com/locate/cma

Ensemble Kalman Filters and geometric characterization of sensitivity spaces for uncertainty quantification in optimization

Bijan Mohammadi

Montpellier University, Mathematics & Modelling Institute, CC51, 34095 Montpellier, France

Received 21 August 2014; received in revised form 16 March 2015; accepted 16 March 2015

Available online 23 March 2015

Abstract

We present an original framework for uncertainty quantification (UQ) in optimization. It is based on a cascade of ingredients with growing computational complexity for both forward and reverse uncertainty propagation. The approach is merely geometric. It starts with a complexity-based splitting of the independent variables and the definition of a parametric optimization problem. Geometric characterization of global sensitivity spaces through their dimensions and relative positions by the principal angles between global search subspaces bring a first set of information on the impact of uncertainties on the functioning parameters on the optimal solution. Joining the multi-point descent direction and the quantiles on the optimization parameters permits to define the notion of Directional Extreme Scenarios (DES) without sampling of large dimension design spaces. One goes beyond DES with Ensemble Kalman Filters (EnKF) after the multi-point optimization algorithm is cast into an ensemble simulation environment. This formulation accounts for the variability in large dimension. The UQ cascade ends with the joint application of the EnKF and DES leading to the concept of Ensemble Directional Extreme Scenarios (EDES) which provides more exhaustive possible extreme scenarios knowing the Probability Density Function of our optimization parameters. A final interest of the approach is that it provides an indication of the size of the ensemble which must be considered in the EnKF. These ingredients are illustrated on an history matching problem.

© 2015 Elsevier B.V. All rights reserved.

Keywords: EnKF; Uncertainty quantification; Principal angles; Inverse problems; History matching; Extreme scenarios

1. Introduction

Forward and backward uncertainty propagation [1–3] are obviously of great importance. For instance, taking into account manufacturing uncertainties is of utmost importance for a design procedure to be efficient as it is impossible to make sure that the final product will exactly correspond to the design specifications. In shape optimization, for instance, either in mono or multidisciplinary frameworks, this uncertainty is rarely accounted for [4,5].

This question can be formulated through both forward and backward uncertainty propagation. Indeed, manufacturing uncertainties can be prescribed through probability density functions of the design parameters (often coming as characteristics of the product manufacturing process) [6,7]. They can also be introduced through uncertainties on the performances of a design (or other observations).

E-mail address: bijan.mohammadi@um2.fr.

1.1. Context of the work

We consider a generic situation where the simulation aims at predicting a given quantity of interest j (e.g. the maximum value of a variable in a given area) and there are a few functioning \mathbf{u} and several control parameters \mathbf{x} involved. The ranges of the functioning parameters define the global operating/functioning conditions of a given design. This splitting of the independent variables in two sets is important for the rest of the paper.

The literature on uncertainty quantification (UQ) is huge. In short, in our situation, forward propagation aims at defining a probability density function for j knowing those of \mathbf{x} and \mathbf{u} [8–10]. This can be done, for instance, through Monte Carlo simulations or a separation between deterministic and stochastic features using Karhunen–Loeve theory (polynomial chaos theory belongs to this class) [11–15]. Examples of shape optimization with polynomial chaos and surrogate models during optimization to address the issue of functional evaluations are given in [16,17].

Backward propagation aims at reducing models bias or calibrating models parameters knowing the probability density function of j (or other constraints and observations) [18–20]. This can be seen as a minimization problem and Kalman filters [21] give, for instance, an elegant framework for this inversion assimilating the uncertainties on the observations.

Our aim is to propose a geometric framework to address, in our particular situation, the curse of dimensionality of existing approaches related to the explosion of their computational complexity due to the sampling necessary to access probabilistic information (momentum), even if this can be improved with intelligent sampling techniques [22,23]. The different ingredients presented here can be applied with either high-fidelity or reduced order models, when available [24–27]. Low-order models are often used instead of the full models to overcome the computational complexity of UQ.

After the splitting of the independent variables mentioned above, we define a multi-point formulation to account for the variability on \mathbf{u} . This is feasible because the size of \mathbf{u} is assumed small. We define a global sensitivity space using the sensitivities of j with respect to \mathbf{x} for the multi-point problem. Once this space built, we analyze the dimension of its free generator subspace. Previous works have shown how to perform this task and how to use this information for adaptive sampling and robust optimization [28,29].

The next step is to analyze the impact of different modeling or discretizations on the results. Different models or solution procedures lead to different sensitivity spaces. Beyond their respective dimensions, principal angles [30–33] between the respective sensitivity spaces permit to measure the deviation due such changes. The dimensions of the spaces and the angles are interesting measures for both the epistemic and aleatory uncertainties. Indeed, suppose that, at given modeling procedure, the dimensions of the sensitivity spaces remain unchanged when enriching the sampling of the functioning parameter range, this stability would be a first indication of a low level of sensitivity of the simulations with respect to this parameter. Once this is established, principal angles between subspaces permit to analyze both the impact of a given evolution of the modeling on the sensitivity spaces or an enrichment of our sampling. Eventually, constant dimension and low angles will clearly indicate a situation of low uncertainty.

These ingredients can be used in a context of multi-point robust analysis of a system to define worst-case scenarios for its functioning. To this end we combine a multi-point sensitivity with the probabilistic features of the control parameters through their quantiles [6,34]. These ingredients permit to define the concept of Directional Extreme Scenarios (DES) without a sampling of large dimension design spaces.

Ensemble Kalman filters (EnKF) [21,35–39] permit to go beyond the directional uncertainty quantification concept when accounting for the uncertainties in large dimension. It also permits backward uncertainty propagation assimilating the uncertainty on the functional and constraints during the design. We propose to cast our multi-point optimization problem into the ensemble formulation. The geometric uncertainty quantification construction ends with the joint application of the EnKF and DES leading to the concept of Ensemble Directional Extreme Scenarios (EDES) which provides exhaustive possible extreme scenarios knowing the probability density function of the optimization parameters and this without any sampling of a large dimensional parameter space. Finally, the geometric information on the global sensitivity spaces provides lower bounds on the size of the ensemble which must be considered in the EnKF.

1.2. Summary of this work ingredients and calculation complexity

To summarize this work will propose a cascade of ingredients to account for uncertainties avoiding any sampling of large dimensional spaces. We insist on the fact a sampling is only necessary for the functioning parameters \mathbf{u} range

(see Section 2) leading to a multi-point optimization problem with a linear complexity with respect of the size of the sampling. One originality is the definition of the direction d in Section 2.1. d is also used, together with quantiles from the PDF of \mathbf{x} , to define quantile-based extreme scenarios in Section 3. In Sections 3 and 4 geometric analysis of the sensitivity space S_m (space dimension, principal angles) gives some information on the impact of the variability of \mathbf{x} on the results without any extra functional evaluation. To go beyond directional extreme scenarios and better account for the variability of \mathbf{x} , the previous ingredients are cast into the ensemble Kalman filter framework in Section 5. Finally, the geometric analysis ingredients are used for each of the ensemble run to introduce Ensemble Directional Extreme Scenarios in Section 5.3.

Concerning the computational cost of these analysis, one can say that, when using the same calculation ingredients that for a high-fidelity simulation (i.e. without calling for low-order models or cheaper discretizations), the best calculation complexity one might think of for a simulation under uncertainty is when its cost is comparable to the deterministic situation. This is clearly unreachable except if all the extra effort can be achieved in a fully parallel manner and parallel to the initial deterministic calculation. This permits for the time to solution to remain unchanged for a simulation when accounting for the presence of uncertainties. This is obviously the case with Monte Carlo approaches. But these are quite expensive and do not take advantage of available simulation environments. In particular, when an adjoint-based optimization environment exists. Hence, we propose an original way to upgrade existing platform without abandoning what has been built for the deterministic situations and with keeping the time to solution unchanged in the presence of uncertainties with two sources of parallelism coming from the multi-point formulation to account for the uncertainties on the functioning parameters and from the EnKF formulation for those on the optimization variables and observation data.

In this work, m being the size of the sampling for the range of \mathbf{u} introduced in Section 2 and q the ensemble size introduced in Section 5, the maximum calculation complexity (to build the Ensemble Directional Extreme Scenarios) at each iteration of optimization is: $2mq$ direct and adjoint solutions of the state equation. This is especially interesting if n , the size of the control parameter \mathbf{x} , is large compared to q .

2. Parametric optimization

We are interested in a class of optimization problems where the cost function involves a functioning parameter \mathbf{u} not considered as a design parameter:

$$\min_{\mathbf{x} \in \mathbf{O}_{ad}} j(\mathbf{x}, \mathbf{u}), \quad \mathbf{u} \in \mathbf{I} \subset \mathbb{R}^\alpha, \mathbf{O}_{ad} \subset \mathbb{R}^n \quad (1)$$

where \mathbf{x} is the design vector belonging to \mathbf{O}_{ad} the optimization admissible domain. Usually, the functioning parameters \mathbf{u} are just a few. On the other hand, the size n of \mathbf{x} is usually large. Together, \mathbf{x} and \mathbf{u} fully describe our system.

This splitting between functioning parameters (or operating conditions) and design variables is central to our discussion.

In [28,40] we showed how to use multi-point optimization to address such optimization problem. The aim is to remove, during the optimization, the dependency in \mathbf{u} . This is done minimizing a functional $J(\mathbf{x})$ encapsulating this dependency expressed through $\mathbf{A} = \{j(\mathbf{x}, \mathbf{u}_l), \mathbf{u}_l \in \mathbf{I}_m\}$ over \mathbf{I}_m a given sampling of \mathbf{I} :

$$J = \mathbf{J}(\mathbf{A}), \quad \text{such that } \mathbf{G}(\mathbf{A}) \leqslant 0. \quad (2)$$

Several choices are possible for \mathbf{J} and \mathbf{G} to address the issue of robust design. For instance, following Taguchi's definition, one can look for minimal-variance design or only a given level for the variance. Indeed, a classical approach to extend single point design and improve off-design points is to control mean performance and variance of the functional [41]. One can also look for information about the tails of the distributions which can be linked to the variance in the Gaussian framework (we will use this relationship in Section 3 in quantile-based extreme scenarios).

To control the first two moments one can choose, for instance, $\mathbf{J} = \mu$ and $\mathbf{G} = \sigma - \sigma_0$ with

$$\mu = \frac{1}{m} \sum_{\mathbf{u}_l \in \mathbf{I}_m} j(\mathbf{x}, \mathbf{u}_l). \quad (3)$$

For the sake of simplicity, we consider a uniform weighting in the sum but more sophisticated choices are possible and in particular choices which enforce the performance we eventually want for the design: constant performance

over the functioning parameters ranges for a robust design [28]. σ_0 is the level of variability of the initial design or of a reference configuration.

We proceed as in First-Order Second Moment (FOSM) methods [42]. σ measures the deviation between $j(\mathbf{x}, \mathbf{u}_l)$:

$$\sigma = \frac{1}{m-1} \sum_{\mathbf{u}_l \in \mathbf{I}_m} (j(\mathbf{x}, \mathbf{u}_l) - \mu)^2. \quad (4)$$

We use this definition when the number of functioning parameters α is large. However, as we want the final performance to be as regular as possible with respect to \mathbf{u} , we ask, when possible, for σ to also monitor the regularity of the performance defined for instance adding the following term to (4):

$$\frac{1}{2} \sum_{\mathbf{u}_l \in \mathbf{I}_m} \langle \nabla_{\mathbf{u}} j(\mathbf{x}, \mathbf{u}_l), \nabla_{\mathbf{u}} j(\mathbf{x}, \mathbf{u}_l) \rangle, \quad (5)$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean scalar product. When α is small, $\nabla_{\mathbf{u}} j$ is evaluated by finite differences on the finite difference grid \mathbf{I}_m . This is reasonable because \mathbf{u} is of small size and that we assume uniform PDF for \mathbf{u} giving a Cartesian grid for a uniform sampling of the ranges of variation of all the components of \mathbf{u} .

Such minimization problems have brought new interest to descent methods and this not only because of their lower computational complexity than the gradient free methods [43–45]. Indeed, we will see that gradients are useful to see what should actually be the search space in a context of robust multi-point design. More precisely, the definition of the suitable search space is only possible after identification of the maximal free generator subspace in the space S_m generated by the gradients of the functional at the sampling points \mathbf{I}_m (see Section 3). This is the subspace generated by the orthonormal basis identified by the Gram–Schmidt procedure at given precision [40,46]. Indeed, without other information and considering vector spaces theory, the size of the sampling should be larger than the dimension of the control space (i.e. $m = n + 1$). We noted, however, [29,28] that the dimension of the maximal free search subspace is usually much smaller than n .

2.1. A multi-point descent algorithm

We consider an iterative descent algorithm for our constrained minimization problem involving a direct simulation chain linking the parameters (\mathbf{x}, \mathbf{u}) to the state y solution of a state equation $y = F^{-1}(\mathbf{x}, \mathbf{u})$ and to a functional j . We assume the functional is enough regular with continuous gradient in both \mathbf{u} and \mathbf{x} . When cast into the Kalman filters framework in Section 5, y (or some part of it) will be called observation. More sophisticated minimization algorithms can be used instead of a simple descent method.

$$\left\{ \begin{array}{l} \text{Given } \mathbf{x}_0, 0 < \rho, 0 < \eta \ll 1, \mathbf{I}_m, k_{max}, 0 < \varepsilon \ll 1, \\ \text{optimization iterations } k = 1, \dots, k_{max} \\ \left\{ \begin{array}{l} -m \text{ parallel state equation solutions } y_l = F^{-1}(\mathbf{x}_k, \mathbf{u}_l), \mathbf{u}_l \in \mathbf{I}_m, \\ -m \text{ parallel evaluations of } j(\mathbf{x}_k, y(\mathbf{x}_k, \mathbf{u}_l)), \mathbf{u}_l \in \mathbf{I}_m, \\ -m \text{ parallel solutions of the adjoint state } V \text{ equation: } V^t F_y(y_l) = j_y^t, \mathbf{u}_l \in \mathbf{I}_m, \\ -m \text{ parallel evaluations of } \nabla_{\mathbf{x}} j(\mathbf{x}_k, \mathbf{u}_l) = j_{\mathbf{x}} + (V^t F_{\mathbf{x}})^t, \mathbf{u}_l \in \mathbf{I}_m, \\ \text{-define } d \text{ the descent direction: } d(\mu(\mathbf{x}_k), \sigma(\mathbf{x}_k)) = \overline{\nabla_{\mathbf{x}} \mu} - \langle \overline{\nabla_{\mathbf{x}} \mu}, \overline{\nabla_{\mathbf{x}} \sigma} \rangle \overline{\nabla_{\mathbf{x}} \sigma} + \eta \overline{\nabla_{\mathbf{x}} \sigma}, \\ \text{-control parameter variation, } \mathbf{x}_{k+1} = \mathbf{x}_k - \rho d, \\ \text{Stop if } \|d\| \leq \varepsilon, \end{array} \right. \end{array} \right.$$

where $\bar{a} = a/\|a\|$ is the normalized vector a and

$$\nabla_{\mathbf{x}} \mu = \frac{1}{m} \sum_{\mathbf{u}_l \in \mathbf{I}_m} \nabla_{\mathbf{x}} j(\mathbf{x}_k, \mathbf{u}_l), \quad (6)$$

and following the definition chosen for σ we have

$$\nabla_{\mathbf{x}} \sigma = \frac{2}{m-1} \sum_{\mathbf{u}_l \in \mathbf{I}_m} (j(\mathbf{x}_k, \mathbf{u}_l) - \mu)(\nabla_{\mathbf{x}} j(\mathbf{x}_k, \mathbf{u}_l) - \nabla_{\mathbf{x}} \mu), \quad (7)$$

with possibly a second term if the regularity issue is included in the definition of σ :

$$\sum_{\mathbf{u}_l \in \mathbf{I}_m} \langle \nabla_{\mathbf{u}} j(\mathbf{x}_k, \mathbf{u}_l), \nabla_{\mathbf{u}\mathbf{x}} j(\mathbf{x}_k, \mathbf{u}_l) \rangle,$$

where $\nabla_{\mathbf{x}\mathbf{u}} j(\mathbf{x}_k, \mathbf{u}_l)$ is obtained from $\nabla_{\mathbf{x}} j(\mathbf{x}_k, \mathbf{u}_l)$ by finite differences on \mathbf{I}_m , component by component, following what has been done for $\nabla_{\mathbf{u}} j$.

The definition of the descent direction permits to make sure that both μ and σ decrease for small descent steps. Indeed, a first order development in \mathbf{x} gives:

$$\sigma(\mathbf{x}_{k+1}) - \sigma(\mathbf{x}_k) = \|\nabla_{\mathbf{x}} \sigma\| \overline{\nabla_{\mathbf{x}} \sigma} \cdot (\mathbf{x}_{k+1} - \mathbf{x}_k) = -\rho \eta \|\nabla_{\mathbf{x}} \sigma\| \leq 0,$$

and we have

$$\mu(\mathbf{x}_{k+1}) - \mu(\mathbf{x}_k) = \|\nabla_{\mathbf{x}} \mu\| \overline{\nabla_{\mathbf{x}} \mu} \cdot (\mathbf{x}_{k+1} - \mathbf{x}_k) = -\rho \|\nabla_{\mathbf{x}} \mu\| (1 - \zeta^2 + \eta \zeta),$$

where $\zeta = \langle \overline{\nabla_{\mathbf{x}} \mu}, \overline{\nabla_{\mathbf{x}} \sigma} \rangle$. Therefore, μ is also decreasing as $1 - \zeta^2 + \eta \zeta \geq 0$ for $|\zeta| \leq 1$ and $0 < \eta \ll 1$ as chosen in the algorithm.

The analysis above can be extended to other functions in (2), other than μ and σ , with

$$d(\mathbf{J}, \mathbf{G}) = \overline{\nabla_{\mathbf{x}} \mathbf{J}} - \langle \overline{\nabla_{\mathbf{x}} \mathbf{J}}, \overline{\nabla_{\mathbf{x}} \mathbf{G}} \rangle \overline{\nabla_{\mathbf{x}} \mathbf{G}} + \eta \overline{\nabla_{\mathbf{x}} \mathbf{G}},$$

as far as \mathbf{J} and \mathbf{G} are such that the quantities in this formula can be defined which implies that we need the functional and constraint to be differentiable.

3. Geometrical characterization of the global search space

During the iterations of the algorithm above a global sensitivity space can be defined using the gradients $\nabla_{\mathbf{x}} j(\mathbf{x}, \mathbf{u}_l) \in \mathbb{R}^n$ for $\mathbf{u}_l \in \mathbf{I}_m$ (\mathbf{I}_m being a m -point sampling of \mathbf{I}):

$$S_m = \text{Span}\{\nabla_{\mathbf{x}} j(\mathbf{x}, \mathbf{u}_l), \mathbf{u}_{l=1, \dots, m} \in \mathbf{I}_m\} \subset \mathbb{R}^{n \times m}.$$

A major quantity of interest is the dimensions p of this subspace as it can be used to identify if a design is robust or not with respect to \mathbf{u} . This dimension depends obviously on the sampling \mathbf{I}_m of the range of variation of \mathbf{u} . But it also depends on the modeling (the governing equations) and the way the state equations are solved (the numerical schemes). Each situation provides a different global sensitivity space. p is therefore also interesting to quantify the impact of a change in the modeling or discretization on a design, also called epistemic uncertainty. Indeed, $p \ll m$ means that \mathbf{u} will have low impact on the sensitivity space S_m for the design \mathbf{x}_k . Of course, this is interesting if the dependency between the functional and the parameters involves a state equation. To illustrate this issue, consider a scalar \mathbf{u} , a constant vector ψ and a functional $j(\mathbf{x}, \mathbf{u}) = \psi^t \mathbf{x} + \mathbf{u}$. Here, $p = 1$ is independent of m since $\nabla_{\mathbf{x}} j(\mathbf{x}, \mathbf{u})$ is independent of \mathbf{u} .

In the past we showed how to evaluate p using an incomplete Gram–Schmidt orthonormalization algorithm. This algorithm is parametric with a threshold to define the accuracy of the orthonormalization. One notice that new independent directions appear increasing this precision. This corresponds to the detection of smaller scales in the gradients which are filtered by a less accurate orthonormalization [28,40,46].

The information on the dimension of S_m for different modeling and/or discretizations can be completed by the respective positions of the different subspaces through the principal angles between them. These two sets of information permit to see uncertainty quantification from a geometric and quite deterministic point of view [40].

Let us briefly recall how to practically compute the principal angles $\theta_i \in [0, \pi/2]$ [31–33,40].

For simplicity, suppose A and B are two subspaces of dimension k of \mathbb{R}^n , $n \geq 2k$, although this is not a prerequisite to define the principal angles. Now, let $\{a_i, i = 1, \dots, k\}$ and $\{b_i, i = 1, \dots, k\}$ be two arbitrary orthonormal bases for A and B . Consider M being the matrix of the projection operator Pr_A of B onto A defined by:

$$Pr_A(b_i) = \sum_{j=1}^k \langle b_i, a_j \rangle a_j, \quad M = (\langle b_i, a_j \rangle)_{i,j}.$$

The principal angles can be linked to this operator through $M = G\Sigma H^t$, where G and H are orthogonal matrices and $\Sigma = \text{diag}(\cos(\theta_i))$.

As G and H are orthogonal matrices, this is a Singular Vector Decomposition (SVD) of M . G and H are unknown, but we do not need them to get the principal angles θ_i . We recall that the columns of G are the left-singular vectors of M and eigenvectors of MM^t and the columns of H are the right-singular vectors of M and eigenvectors of M^tM . And most important that $\cos^2(\theta_i)$ are the eigenvalues of $Pr_A^t Pr_A$ which writes in matrix form as: $M^t M = (G\Sigma H^t)^t (G\Sigma H^t) = H\Sigma^2 H^t$ with $\Sigma^2 = \text{diag}(\cos^2(\theta_i))$.

Therefore, to find the principal angles between subspaces A and B , knowing an orthonormal basis in each subspace, one should calculate M and find the eigenvalues of $M^t M$ and take the square root of them. This last operation is valid as the angles are between 0 and $\pi/2$, and their cosine therefore always positive. In our case, an orthonormal basis is already available after the Gram–Schmidt orthonormalization has been used to identify the dimension of global search space S_m .

4. Directional extreme scenarios

We showed how to introduce some geometric characterizations of the global sensitivity spaces built with the gradients of a functional $\nabla_{\mathbf{x}} j(\mathbf{x}, \mathbf{u}_l)$ for $\mathbf{u}_l \in \mathbf{I}_m$ a sampling of a functioning parameter variation range. This was feasible as we assumed there are only a few of such parameters in a system. On the other hand, it is not possible to account for the variability in the design parameter $\mathbf{x} \in \mathbb{R}^n$ as $1 \ll n$.

In [6] we showed how to introduce quantiles such as the Value at Risk (VaR) in optimization algorithms with the aim of quantifying our confidence on the optimal solution at low complexity without a sampling of the control space. This concept can be used in combination with the direction d of Section 2.1 to define the notion of Directional Extreme Scenarios (DES).

We consider that a given control parameter $\mathbf{x}_{i=1,\dots,n}$ represents, with a confidence level of a , the interval $[\mathbf{x}_i + \text{VaR}_a^-, \mathbf{x}_i + \text{VaR}_a^+]$ around \mathbf{x}_i and $\text{VaR}_a^- \leq 0 \leq \text{VaR}_a^+$. One expresses then possible deviations from \mathbf{x} via two probability density functions (PDF) and the corresponding values at risk. One can make the hypothesis that the upper and lower bounds of the variations are symmetric, in which case $\text{VaR}_a^- = -\text{VaR}_a^+$. This is typically the case when uncertainties on a parameter follow a Gaussian distribution. Gaussian distributions are interesting as their Values at Risk are explicitly known: $\text{VaR}_{0.99} = 2.33$ and $\text{VaR}_{0.95} = 1.65$ for $N(0, 1)$ and $\text{VaR}_a(N(0, \sigma)) = \sigma \text{VaR}_a(N(0, 1))$.

To summarize $\mathbf{x} + \text{VaR}_a(\mathbf{x})$ is a closed domain in \mathbb{R}^n around \mathbf{x} :

$$\mathbf{B}_a(\mathbf{x}) = \{z \in \mathbb{R}^n : \mathbf{x}_i + \text{VaR}_a^-(\mathbf{x}_i) \leq z_i \leq \mathbf{x}_i + \text{VaR}_a^+(\mathbf{x}_i), i = 1, \dots, n\}. \quad (8)$$

The knowledge of the level of uncertainties on \mathbf{x} is by itself a problem which we do not discuss here. Manufacturing uncertainties are usually known through probability density functions coming as characteristics of the product manufacturing process: realization within margin errors. These PDFs can also be identified by assimilation observation data.

Now, we would like to use this information together with the descent direction d given in Algorithm 2.1 to define extreme values for the local variability of the functional along the direction d . We make the assumption that the functional is monotonic over $\mathbf{B}_a(\mathbf{x})$, continuous and bounded. It is therefore uniformly continuous. This guarantees that a small perturbation in the design parameters will have small effect on the performance, something which is very reasonable to assume for any robust design. For the sake of simplicity, one can suppose the functional is convex in $\mathbf{B}_a(\mathbf{x})$. If the functional does not behave as such it would be very hazardous to pretend providing any robust design.

These assumptions imply that extreme values of the functional over $\mathbf{B}_a(\mathbf{x})$ are reached on $\partial\mathbf{B}_a(\mathbf{x})$, the boundary of $\mathbf{B}_a(\mathbf{x})$.

The intersection of $\partial\mathbf{B}_a(\mathbf{x})$ and d permits to define two Directional Extreme Scenarios for \mathbf{x} as worst-case scenarios:

$$(DES) : \mathbf{x}^\pm = d \cap \partial\mathbf{B}_a(\mathbf{x}). \quad (9)$$

A sketch of this construction is given in Fig. 1. The uncertainty domain is of rectangular shape because we have assumed the components of \mathbf{x} are independent (i.e. the correlation matrix is diagonal, see Section 5).

In [6] we showed how this can be linked to the notion of over-solving where it becomes useless to solve accurately near an optimum when the variation in a control parameter falls below the level of the variability for this parameter.

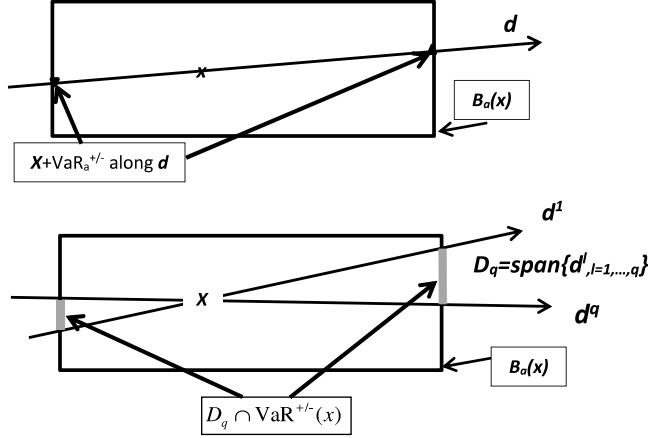


Fig. 1. Sketch of Directional Extreme Scenarios (DES) given by $\mathbf{x}^\pm = d \cap \partial B_a(\mathbf{x})$ and Ensemble Directional Extreme Scenarios (EDES) $D_q \cap \partial B_a(\bar{\mathbf{x}})$ given (16). The gray zone is not necessarily connected.

In other words, in the presence of uncertainties, one should not distinguish between the points in $B_a(\mathbf{x})$. This interpretation should also be used to define what we mean by uniqueness of the solution: one cannot distinguish between the solutions in $B_a(\bar{\mathbf{x}})$. This construction meets the ingredients developed in the Unscented Kalman filter approaches [47] and the extreme directional scenarios can be seen as two of the σ -points there.

The next step is to apply the geometric analysis of Section 3 after the definition of two global sensitivity spaces S_m^\pm at \mathbf{x}^\pm using $\nabla_{\mathbf{x}} j(\mathbf{x}^\pm, \mathbf{u}_l)$ for $\mathbf{u}_l \in \mathbf{I}_m$ following the same procedure than for S_m . The respective dimensions p^\pm of these spaces and their respective position through the principal angles between them give precious information on the impact of local variability on \mathbf{x} on the global search space [40,46]. For instance, if the dimensions are the same and angles small, this would mean that variability in \mathbf{x} will have very little impact on the design. Also, from the calculation complexity point of view, it is important that these have been built without a sampling of our parameter space of dimension n for \mathbf{x} .

In some situations, we might be interested in going beyond just extreme scenarios and include the treatment of the uncertainties on the control parameters and observation data in the optimization solution. Kalman filters provide an elegant framework for such developments. Also, in Section 5.3 we go beyond the assumptions above for j over $B_a(\mathbf{x})$ using the Ensemble Directional Extreme Scenarios (EDES) where the construction will be applied with the different directions $\{d^l, l = 1, \dots, q\}$ built for the ensemble simulation. This propose an ensemble of worst case scenarios as a local sampling of $\partial B_a(\mathbf{x})$. The history matching example in Section 6 will show a situation where the DES construction is not enough, indeed, and that the ensemble simulation is necessary for a more exhaustive sampling. Still, $2q$ (q being the ensemble size) the size of this sampling remains negligible compared to n the dimension of the global sensitivity space D_q .

5. Extended and Ensemble Kalman filters

The Kalman filter (KF) [21] provides solution for state estimation of linear systems under Gaussian noise and permits to sequentially assimilate observation data. Extended Kalman filters (EKF) based on the use of the Jacobians of the forecast and observation operators with respect to the state permit to extend the Kalman filter analysis to the case of nonlinear systems [35–37]. To avoid the introduction of these Jacobians, Ensemble Kalman filters (EnKF) [38,39] introduce an approximation of the error matrices using statistics generated through scenarios run in parallel. Our aim is to cast our multi-point minimization into an EnKF formulation.

Let us recall the EKF and EnKF algorithms for the sequential solution \mathbf{x}_k of the following nonlinear dynamical system under uncertainties and assimilating observations y_k .

$$\mathbf{x}_{k+1} = f(\mathbf{x}_k, \mathcal{U}_k) + q_k, \quad y_{k+1} = g(\mathbf{x}_k) + r_k, \quad (10)$$

where $\mathbf{x}_k, q_k \in \mathbb{R}^n$, $\mathcal{U}_k \in \mathbb{R}^v$ and $y_k, r_k \in \mathbb{R}^\pi$. We usually have $n > v$ and $n > \pi$ as we would like to be able to control a system with less control than the size of the system and also there are usually less available observations than this size. Kalman filters are for Gaussian noises and we assume that q_k and r_k are stationary zero-mean Gaussian perturbations with covariance matrices Q and R . The initial condition \mathbf{x}_0 and q_k and r_k are also assumed uncorrelated.

Defining the Jacobians $A_k = \nabla_{\mathbf{x}} f(\mathbf{x}_k, \mathcal{U}_k)$ and $C_k = \nabla_{\mathbf{x}} g(\mathbf{x}_k)$, the EKF algorithms is made of an assimilation (upperscript a) and a forecast (upperscript f) step:

$$\text{EKF : } \begin{cases} \text{Assimilation: } \mathbf{x}_k^a = \mathbf{x}_k^f + K_k(y_k - g(\mathbf{x}_k^f)), & P_k^a = (I - K_k C_k) P_k^f, \\ \text{Forecast: } \mathbf{x}_{k+1}^f = f(\mathbf{x}_k^a, \mathcal{U}_k), & P_{k+1}^f = A_k P_k^a A_k^t + Q, \end{cases} \quad (11)$$

where $K_k = P_{\mathbf{x}y_k}^f (P_{yy_k}^f)^{-1}$, $P_{\mathbf{x}y_k}^f = P_k^f C_k^t$ and $P_{yy_k}^f = C_k P_k^f C_k^t + R$. We recover the original Kalman filter algorithm if f and g are linear ($f(\mathbf{x}_k, \mathcal{U}_k) = A_k \mathbf{x}_k + B_k \mathcal{U}_k$, $g(\mathbf{x}_k) = C_k \mathbf{x}_k$).

The EnKF algorithm intends to avoid the progression steps for P_k^a and P_k^f requiring the Jacobians of f and g . Instead, taking advantage of q parallel forecast for an ensemble of states $(\mathbf{x}_k^{f_1}, \dots, \mathbf{x}_k^{f_q})$, the progression steps above are replaced by the following estimation of the forecast and assimilation states error covariance matrices:

$$P_k^f \sim \frac{1}{q-1} E_k^f (E_k^f)^t, \quad P_k^a \sim \frac{1}{q-1} E_k^a (E_k^a)^t,$$

where E_k^f and E_k^a are vectors of the deviation of each of the ensemble member from the mean previsions $\bar{\mathbf{x}}_k^f = (1/q) \sum_{i=1}^q \mathbf{x}_k^{f_i}$ and $\bar{\mathbf{x}}_k^a = (1/q) \sum_{i=1}^q \mathbf{x}_k^{a_i}$:

$$E_k^f = (\mathbf{x}_k^{f_1} - \bar{\mathbf{x}}_k^f, \dots, \mathbf{x}_k^{f_q} - \bar{\mathbf{x}}_k^f), \quad E_k^a = (\mathbf{x}_k^{a_1} - \bar{\mathbf{x}}_k^a, \dots, \mathbf{x}_k^{a_q} - \bar{\mathbf{x}}_k^a).$$

The same procedure is applied to generate from forecast observations ($y_k^{f_i} = g(\mathbf{x}_k^{f_i})$) the error output vector $E_{y_k}^f = (y_k^{f_1} - \bar{y}_k, \dots, y_k^{f_q} - \bar{y}_k)$ which permits to define the following approximation for the forecast observation error covariance matrices:

$$P_{\mathbf{x}y_k}^f \sim \frac{1}{q-1} E_k^f (E_{y_k}^f)^t \quad \text{and} \quad P_{yy_k}^f \sim \frac{1}{q-1} E_{y_k}^f (E_{y_k}^f)^t.$$

The ensemble initialization is based on the probability density functions for \mathbf{x} and y :

$$\mathbf{x}_0^{f_i} = \mathbf{x}_0 + q_0^i, \quad y_0^i = y_0 + r_0^i, \quad i = 1, \dots, q.$$

The assimilation and the forecast steps are then the same for the EKF (and also the KF) and each of the ensemble members in EnKF except that the approximations above are used in the definition of the Kalman gain matrix K_k :

$$\text{EnKF : } \begin{cases} \text{Ensemble Assimilation: } \mathbf{x}_k^{a_i} = \mathbf{x}_k^{f_i} + K_k(y_k^i - g(\mathbf{x}_k^{f_i})), \\ \text{Ensemble Forecast: } \begin{cases} \mathbf{x}_{k+1}^{f_i} = f(\mathbf{x}_k^{a_i}, \mathcal{U}_k) + q_k^i, \\ y_{k+1}^i = y_{k+1} + r_{k+1}^i. \end{cases} \end{cases} \quad (12)$$

The ensemble calculations are independent and can therefore be carried out in a fully parallel manner. The algorithm is in particular interesting when it succeeds with an ensemble of size $q \ll n$.

5.1. Optimization and Kalman filters

Our aim is to cast the optimization algorithm given in Section 2.1 into the Kalman filters formulation. We will be using the ensemble Kalman filter because the systems are nonlinear in our context. We would like to use this approach for robust optimal design accounting for uncertainties in the optimization parameters \mathbf{x} and observations y . These gather physical state variables and the constraints which will be seen as observations y . These can be, for instance, global or partial information on \mathbf{x} with associated uncertainties. The choice of the ensemble Kalman filter over the extended Kalman filter is because EKF would require the Hessian of j and the Jacobian of the governing equations and the constraints. These are not always available or easy to get. Our aim is to be able to account for the uncertainties using existing platforms.

The multi-point descent algorithm given in Section 2.1 controlling the first and second momentum of a functional can be seen as iterations of:

$$\mathbf{x}_{k+1} = f(\mathbf{x}_k, \mathcal{U}_k) + q_k, \quad y_{k+1} = g(\mathbf{x}_k) + r_k,$$

where the solution of the governing equations F (and other possible equality constraints) are gathered as observations in y_k . $\mathcal{U}_k \in \mathbb{R}^m$ represents the sampling of the range of the functioning parameter \mathbf{u} and is given by $\mathcal{U}_k = (\mathbf{u}_1, \dots, \mathbf{u}_m) \in \mathbb{R}^{m \times \alpha}$ (i.e. $v = m \times \alpha$). Recall that we have $\mathbf{u} \in \mathbf{I}_m \subset \mathbf{I} \subset \mathbb{R}^\alpha$ with m the size of the sampling. q_k and r_k represent the uncertainties on \mathbf{x}_k and y_k . These are zero-mean Gaussian with covariance matrices Q and R . In the sequel the approach is therefore limited to the Gaussian framework.

5.2. Illustration on quadratic minimization under equality constraints

Let us start with a simple optimization problem under equality constraints and without any governing equation involved. We consider the minimization of a quadratic functional under two equality constraints:

$$j_0(\mathbf{x}, \mathbf{u}) = \frac{1}{2} \sum_{i=1}^n \left(\mathbf{x}_i - \mathbf{u} \frac{i}{n} \right)^2, \quad c_1(\mathbf{x}) = \sum_{i=1}^{n/2} \mathbf{x}_i - \frac{1}{2} = 0, \quad c_2(\mathbf{x}) = \sum_{i=n/2+1}^n \mathbf{x}_i - 1 = 0.$$

The size of the optimization problem is $n = 100$. Let us start with $\mathbf{u} = 1$ being a constant and not given over an interval $\mathbf{I} \subset \mathbb{R}$. Without constraints the optimal solution is obviously $\mathbf{x}_i = i/100$. With constraints and in the absence of noise this can be solved considering a functional of the form:

$$j(\mathbf{x}, \mathbf{u}) = j_0(\mathbf{x}, \mathbf{u}) + \frac{1}{2} c_1^2(\mathbf{x}) + \frac{1}{2} c_2^2(\mathbf{x}). \quad (13)$$

First order optimality condition for j leads to a $n \times n$ linear system $A\mathbf{x} = b(\mathbf{u})$ where $A_{i,i} = 2$, $A_{i,j} = 1$ when $i & j \leq n/2$ or $i & j \geq n/2$ and $A_{i,j} = 0$ elsewhere and $b_i = \mathbf{u} i/n - 1/2$ for $i \leq n/2$ and $b_i(\mathbf{u}) = \mathbf{u} i/n - 1$ for $i > n/2$. The solution of this system is shown in Fig. 2.

In the presence of uncertainties on the optimization parameter \mathbf{x} and the constraints c_i , this can be cast into the Kalman filter formalism (classical here as we are in a linear dependency situation) with c_i seen as two observations because these constraints can be considered as a priori global information on \mathbf{x}_k , and only partially satisfied (noise is the error one commits on the equality constraints):

$$f(\mathbf{x}_k, \mathbf{u}_k) = \mathbf{x}_k - \rho \nabla_{\mathbf{x}} j(\mathbf{x}_k, \mathbf{u}_k) \quad y_k^1 = c_1(\mathbf{x}_k), \quad y_k^2 = c_2(\mathbf{x}_k), \quad (14)$$

where $\nabla_{\mathbf{x}} j(\mathbf{x}_k, \mathbf{u}_k) = A\mathbf{x}_k - b(\mathbf{u}_k)$ with $\mathcal{U}_k = \mathbf{u}_k = \mathbf{u}$ and $\rho = 0.01$. The Gaussian uncertainties on the optimization parameter \mathbf{x}_k and constraints $y_k^{1,2}$ are given through their correlation matrices $Q = 0.02I_{100}$ and $R = 0.05I_2$. Fig. 2 shows comparisons between the unperturbed solution and when uncertainties on the optimization parameters and the constraints realizations have been reduced by the EnKF algorithm. Ensembles of size 5 to 20 have been tested and the convergence histories for the deviations between the assimilation and forecast ensemble solutions and the deviation between ensemble solutions at given time k -index.

5.2.1. Parametric quadratic minimization under equality constraints

Now consider the situation where \mathbf{u} is no longer a constant but belongs to an interval $\mathbf{I} = [0.8, 1.8]$. As we discussed in Section 2, this is quite representative of what we often have in practical situation with \mathbf{u} being a functioning parameter and we would like the optimal solution to be robust with respect to \mathbf{u} : the functional to be as much as possible flat on its domain of variation, for instance. We consider \mathbf{I}_m a $m = 20$ points uniform sampling of \mathbf{I} and $J(\mathbf{x}) = \mu$ defined in (3) and with $j(\mathbf{x}, \mathbf{u}_l)$ given by (13). With the notations above, the first order optimality condition leads this time to $A\mathbf{x} = \sum_{\mathbf{u} \in \mathbf{I}_m} b(\mathbf{u})$. As for the constant \mathbf{u} case, this can be cast into the Kalman filter formalism in the presence of uncertainties on the optimization parameter \mathbf{x} and the constraints c_i with $\nabla_{\mathbf{x}} J(\mathbf{x}_k, \mathbf{u}_k) = (A\mathbf{x}_k - \sum_{\mathbf{u} \in \mathbf{I}_m} b(\mathbf{u}))/m$ and the size of the control $\mathcal{U}_k = (\mathbf{u}_1^k, \dots, \mathbf{u}_m^k)$ being now m . The multi-point descent direction $d(\mu, \sigma)$ can be built as in Section 2.1 and the multi-point descent algorithm can be cast into the Kalman filter framework replacing $\nabla_{\mathbf{x}} j(\mathbf{x}_k, \mathbf{u}_k)$

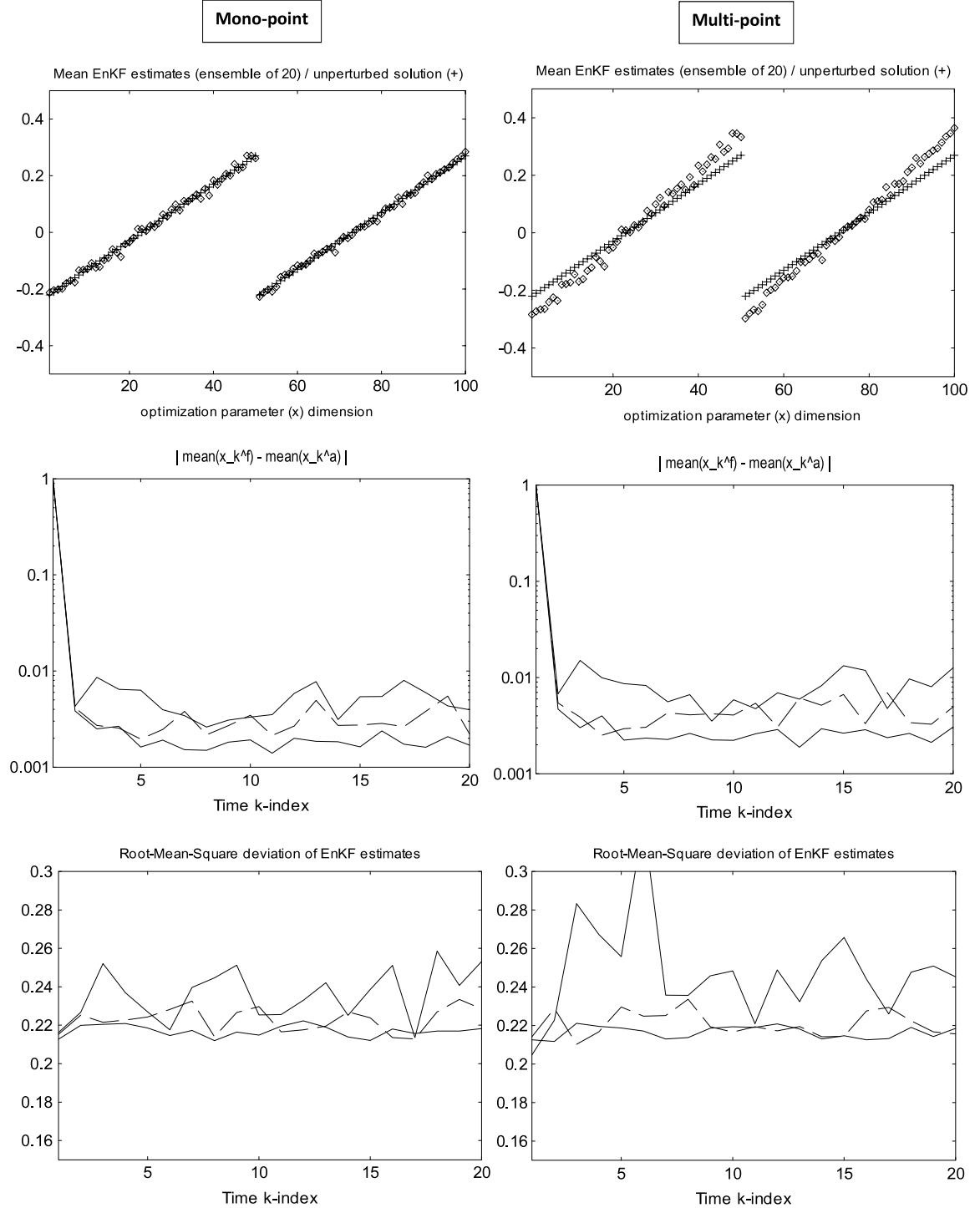


Fig. 2. Mono-point (left column) and parametric (right column) quadratic minimization under equality constraints: comparison between the unperturbed solution and when uncertainties on the optimization parameters and the constraints realizations have been reduced by the EnKF algorithm.

by d in (14). Fig. 2 shows comparisons between the mono-point and this parametric minimization for ensembles of size 5 to 20. The parametric solution proposed by the multi-point minimization deviates quite from the deterministic

solution. Even if this might appear being counter-intuitive, this means that it should be more difficult to fit the data in a parametric inverse problem. We will recover this behavior in Section 6 for a history matching problem.

5.3. Geometry of the global EnKF search space and ensemble size

At this point one can take advantage of the descent directions $\nabla_{\mathbf{x}} j(\mathbf{x}_k^l, \mathbf{u})$ or $d(\mu(\mathbf{x}_k^l), \sigma(\mathbf{x}_k^l))$ built during our ensemble simulations and proceed as in Section 3 to analyze the geometric characteristics of the related global search space D_q in the context of single point optimization (for a single \mathbf{u} value):

$$D_q = \text{Span}\{\nabla_{\mathbf{x}} j(\mathbf{x}_k^l, \mathbf{u}), l = 1, \dots, q\} \subset \mathbb{R}^{n \times q},$$

or multi-point (when $\mathbf{u} \in \mathbf{I}$) optimization:

$$D_q = \text{Span}\{d(\mu(\mathbf{x}_k^l), \sigma(\mathbf{x}_k^l)), l = 1, \dots, q\} \subset \mathbb{R}^{n \times q}.$$

An important interest of these spaces is to provide information on what should be the size of the ensemble in EnKF. This question is treated empirically in practice increasing the size q until the EnKF iterations succeed. An ensemble can be considered as enough rich if $\dim(D_q) < q$.

In addition to their respective dimensions, principal angles between two spaces D_q and $D_{q'}$ show how much deviation appears between subspaces for different ensemble sizes. This permits to define a new concept of convergence for the ensemble simulation.

This geometric construction can be used to define possible extreme scenarios introducing quantiles such as the value at risk (VaR) as presented in Section 4. More precisely, ensemble directions $d_k^l = d(\mu(\mathbf{x}_k^l), \sigma(\mathbf{x}_k^l))$, $l = 1, \dots, q$ permits to define a sampling of what one might call an extreme scenarios region around a mean inversion solution $\bar{\mathbf{x}}_k$:

$$D_q \cap \partial \mathbf{B}_a(\bar{\mathbf{x}}_k), \quad (15)$$

where $\bar{\mathbf{x}}_k = (1/q) \sum_{l=1}^q \mathbf{x}_k^l$. Hence, applying the Directional Extreme Scenarios (DES) concept to the each of the ensemble directions provides a sampling of the extreme scenarios region through the following Ensemble Directional Extreme Scenarios:

$$(\text{EDES}) : \{(\mathbf{x}_k^l)^{\pm}, l = 1, \dots, q\} = \{d_k^l \cap \partial \mathbf{B}_a(\bar{\mathbf{x}}_k), l = 1, \dots, q\}. \quad (16)$$

A sketch of this construction is given in Fig. 1 and compared to the situation presented in Section 4. These extreme scenarios will be illustrated in the examples of Section 6 and compared to the results by simple DES when the Ensemble Kalman filter is not used and the uncertainty of \mathbf{x} only accounted for through two directional VaR-based extreme scenarios (in mono and multi-point situations for \mathbf{u}) permitting for low-complexity analysis of the impact of variability in \mathbf{x} on the inversion.

6. Application to a history matching problem

We consider an inverse problem appearing in several science and engineering domains such as electroencephalography (EEG) inverse source localization and history matching in reservoir modeling [2]. We discuss this latter problem.

Reservoir simulation is a major activity for many oil companies. It starts with building the initial geological reservoir and fluid models and their discrete versions. These can be either PDE-based or relying on reduced order models, for instance involving response surfaces. Our reservoir model is PDE based but other solutions can be envisaged and this choice is not central to our discussion. In all cases, and as for many models in science and engineering, one should be aware that reservoir models are not reality and have inevitable errors also called reducible or epistemic uncertainties.

After the reservoir model is established, it must be improved through history matching which must include a quantification of uncertainties. This step consists of adjusting the model of a reservoir until it closely reproduces its past behavior available through the historical production and pressures data which are both noisy. These are not reducible and are called aleatory uncertainties.

In addition to these data, observations and constraints (which can be considered as a priori information and somehow as observations) are also available. Examples of such are well logs and information there on the core, seismic data, geological and production scenarios, etc.

Previous works exist on the application of ensemble Kalman filter methods to history matching problems (see [48–50] and references therein). EnKF being in the class of Monte Carlo simulation methods, it is usually stated that EnKF is suitable as it avoids gradient based minimization algorithms which can be trapped by local minima. We think that one should take advantage of the gradients information when available and rather than dropping gradient-based minimization algorithms cast them into the EnKF framework. Gradient information also reduces the risk for a pure EnKF estimates to deviate strongly from petrophysical properties. In addition, thanks to the knowledge of the dynamic global sensitivity spaces built during EnKF iterations, one have a good estimation of what the ensemble size should be.

To illustrate our purposes we consider a synthetic case with two injection and two production wells in a $100 \text{ m} \times 100 \text{ m}$ domain. We consider a simple time dependent Darcy-like model [51] over a domain Ω linking a single phase pressure \mathbf{p} to the flow velocity \mathbf{v} (gravity is neglected):

$$\partial_t \mathbf{p} + \nabla \cdot \mathbf{v} = \sum_{i=1,2} \mathbf{u}_i \delta_{s_i}, \quad \mathbf{v} = -\mathbf{x} \nabla \mathbf{p}, \quad (17)$$

which is solved on a 100×100 cartesian mesh following [52] and using a second order time integration. The literature on the solution of this equation is very rich. A particular choice of a numerical scheme is not central to our discussion. The algorithm given in Section 2.1 requires the adjoint of the governing equation in order to make the cost of sensitivity evaluation independent of the size n of the optimization space for \mathbf{x} . To make the code development as automatic as possible, the adjoint solver has been generated using automatic differentiation by Tapenade [53,54].

In Eq. (17) \mathbf{x} describes the soil characteristics and our aim is its identification. More precisely, $\mathbf{x} = k(\mu\phi)^{-1}$ with $k(m^2)$ the permeability, $\mu(\text{bar} \cdot \text{hour})$ the fluid viscosity and $0 < \phi \leq 1$ is the porosity of the soil. An a priori available information on \mathbf{x} is that it must be piecewise constant over a number of plateau ω_i . This is accounted for through a projection step on \mathbf{x} . We have here an example of information provided by geophysicists from seismic observations. Hence, the optimization parameter K over the computation mesh is projected over this representation defining a set of couples $\{(K_i, \omega_i), i = 1, \dots, n_p\}$ where

$$K_i = \frac{1}{|\omega_i|} \int_{\omega_i} \mathbf{x}(s) ds, \quad \cup_{i=1}^n \omega_i = \Omega.$$

In the examples presented here, ω_i describes transverse strips as shown in Fig. 3 and we have a number of plateau $n_p \leq n$ with $n = 100$ the size of the optimization space. This a priori information is very useful for an efficient ensemble simulation permitting to keep the ensemble size low as shown in Fig. 5.

In addition to the information above we have well logs and \mathbf{x} is known at wells and this information defines the observation equations:

$$y_k^{i=1,\dots,4} = \mathbf{x}_k(s_i) - \mathbf{x}^*(s_i) = 0, \quad \text{at } s_i \text{ the locations of the 4 wells.}$$

The control variables $\mathbf{u} \in \mathbb{R}_+^2$ represents the injection pressure at the two injection wells $s_{1,2}$. This is an example where $\mathbf{u} \in \mathbb{R}_+^2$ in $j(\mathbf{x}, \mathbf{u})$ of Section 2.

We aim at recovering \mathbf{x} minimizing:

$$j(\mathbf{x}, \mathbf{u}) = \sum_{i=3,4} \int_0^T (\mathbf{p}_i(t, \mathbf{x}, \mathbf{u}) - \mathbf{p}_i^*(t))^2 dt,$$

from histories $\mathbf{p}_{i=3,4}^*(t)$ recorded at production wells $s_{3,4}$. $\mathbf{p}_i^*(t)$ is uncertain but this uncertainty is assumed small in comparison to what expected on the permeability and the observations. The physical time (100 h) over which these histories have been recorded and used for history matching is different from the k -index time in the EnKF iterations. These correspond to optimization iterations.

The uncertainties on the optimization parameters \mathbf{x}_k and constraints $y_k^{1,\dots,4}$ are both assumed Gaussian with correlation matrices $Q = 0.02I_{100}$ and $R = 0.05I_4$.

Our aim is to identify \mathbf{x} by our gradient-based+EnKF inversion approach accounting for the different uncertainties in a mono-point situation and when the control \mathbf{u} is defined through its range of variation.

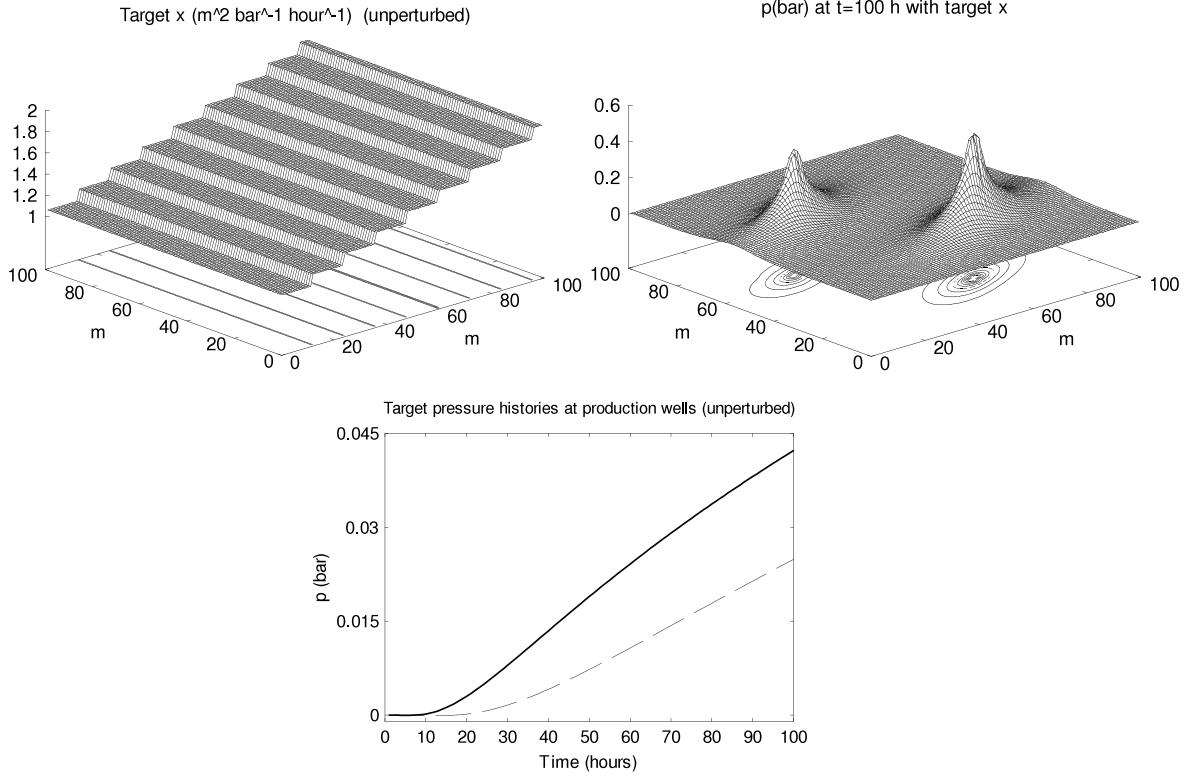


Fig. 3. Target \mathbf{x} after projection and definition of the set $\{(K_i, \omega_i), i = 1, \dots, 10\}$ where $K (= k(\mu\phi)^{-1})$, field pressure at $t = 100h$ and pressure histories $\mathbf{p}_{i=3,4}^*(t)$ recorded at production wells $s_{3,4}$ over 100 h and used as target for inversion.

6.1. Mono-point inversion under uncertainties

Before considering the geometric analysis of the global sensitivity spaces in a multi-point situation, let us show an example of the results one gets with the EnKF-based inversion in a mono-point configuration defined with constant values over time for the two components of the control \mathbf{u} at 1 bar and 1.3 bar at injection wells. These are not realistic values and are just illustrative. The governing equation being linear in \mathbf{p} , there is no difficulty in considering larger values for the injection pressure and the inversion will succeed as far as the target pressure histories are compatible.

To be sure the problem is admissible, an unperturbed target solution has been generated with a piecewise constant permeability distribution with $n_p = 10$ plateau (shown in Fig. 3).

We consider three ensembles of size $q = 10, 20$ and 40 . Fig. 4 shows the difference between the target \mathbf{x} and \mathbf{p} and the mean ensemble estimates for the ensemble of size 40. It also shows the deviation between ensemble solutions. These results permit to see the regions of maximum deviation between the EnKF estimates and also between these EnKF estimates and the solution by a deterministic inversion (i.e. when the inversion does not account for the uncertainties on the data and observations). It is also interesting to see that the target pressure histories at production wells $s_{3,4}$ are well recovered by the ensemble mean solution. But that the uncertainty on the production at the two wells are different.

Fig. 5 shows the histories over time k -index of the deviation between the ensemble analysis and forecast EnKF estimates and the root-mean-square deviation of the forecast estimates for three ensembles of size $q = 10, 20$ and 40 . This problem is more complex than the quadratic minimization of Section 5.2. We have the same level of uncertainties in both problems and the same size of the problem ($n = 100$). Still an ensemble of size 20 seems to produce convergent statistics. This can be explained with our geometric analysis. Fig. 5 also shows the outcome of the Gram–Schmidt algorithm detecting independent directions in the global sensitivity space $D_q = \text{Span}\{\nabla_{\mathbf{x}} j(\mathbf{x}_k^l, \mathbf{u}), l = 1, \dots, q\}$ for $q = 20$ and 40 as defined in Section 5.3 and eventually their dimensions which remain below 10. This gives an indication on how to choose the ensemble size in EnKF.

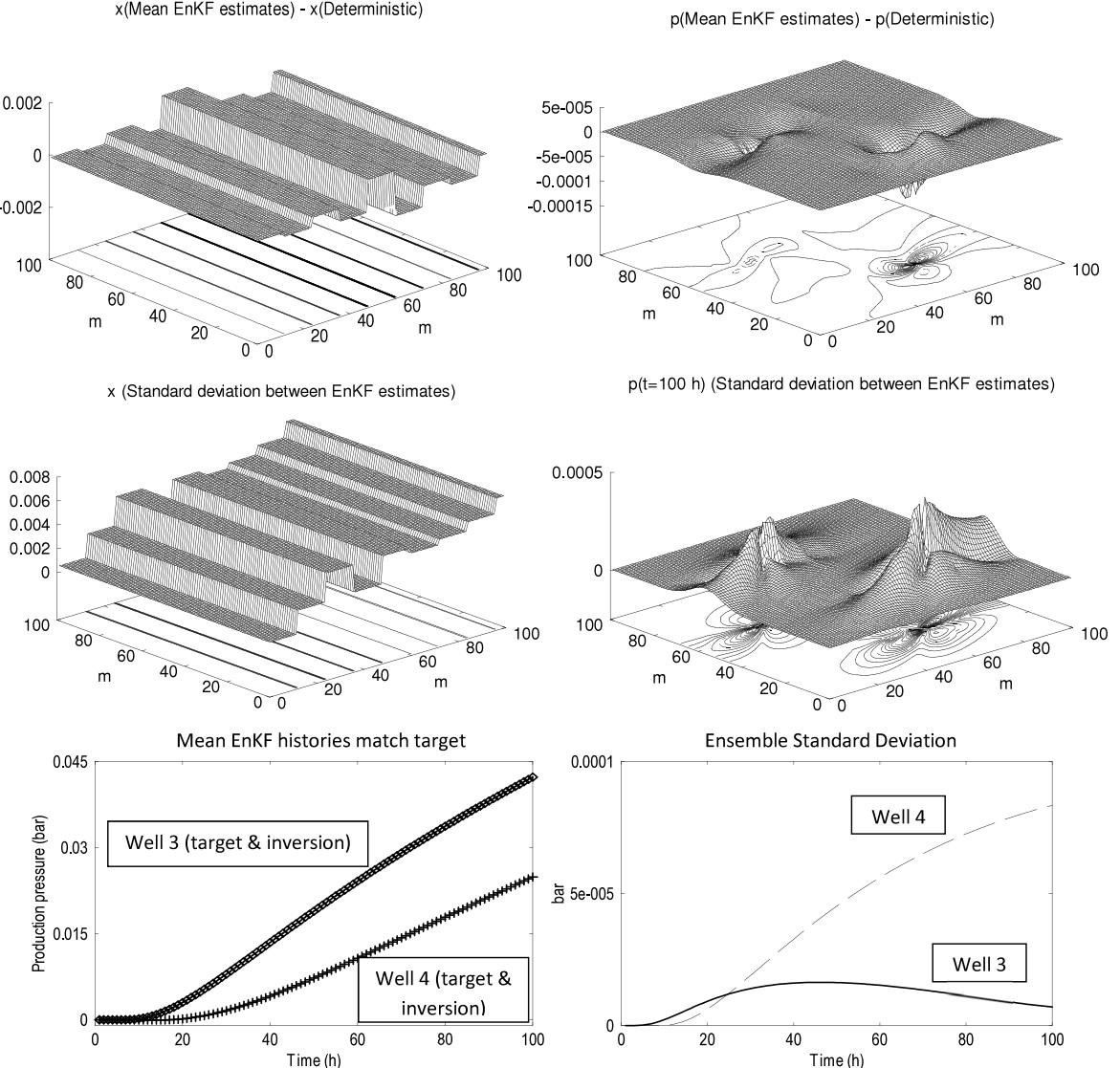


Fig. 4. Inversion with EnKF (ensemble of size 40) in a mono-point \mathbf{u} case. Differences between the ensemble mean solutions from the target \mathbf{x} and the field pressure $\mathbf{p}(t = 100h)$ and the deviation between ensemble solutions. Target pressure histories at production wells $s_{3,4}$ are well recovered by the ensemble mean but uncertainties on the production histories do not follow the same pattern.

The figure also illustrates the application of the concept of principal angles between subspaces to estimate the relative position of the global sensitivity spaces D_{20} and D_{40} . In addition to the dimension of the spaces, these angles provide additional information on the impacts on the search spaces due to an enrichment of the ensemble.

6.1.1. Var-based Directional Extreme Scenarios in the mono-point case

As described in Section 5.3, the EnKF inversion permits to define a sampling of the extreme scenarios region (15). We also saw in Section 4 how to define two VaR-based directional extreme scenarios during a gradient based minimization. We would like to compare the outcome of the two approaches with the EnKF analysis considered as the reference solution.

Knowing the Gaussian probability density function for \mathbf{x}_k , we define two VaR-based scenarios along $\nabla_{\mathbf{x}} j(\mathbf{x}_k, \mathbf{u})$:

$$(\mathbf{x}_k)^{\pm} = \nabla_{\mathbf{x}} j(\mathbf{x}_k, \mathbf{u}) \cap \partial \mathbf{B}_{0.95}(\mathbf{x}_k),$$

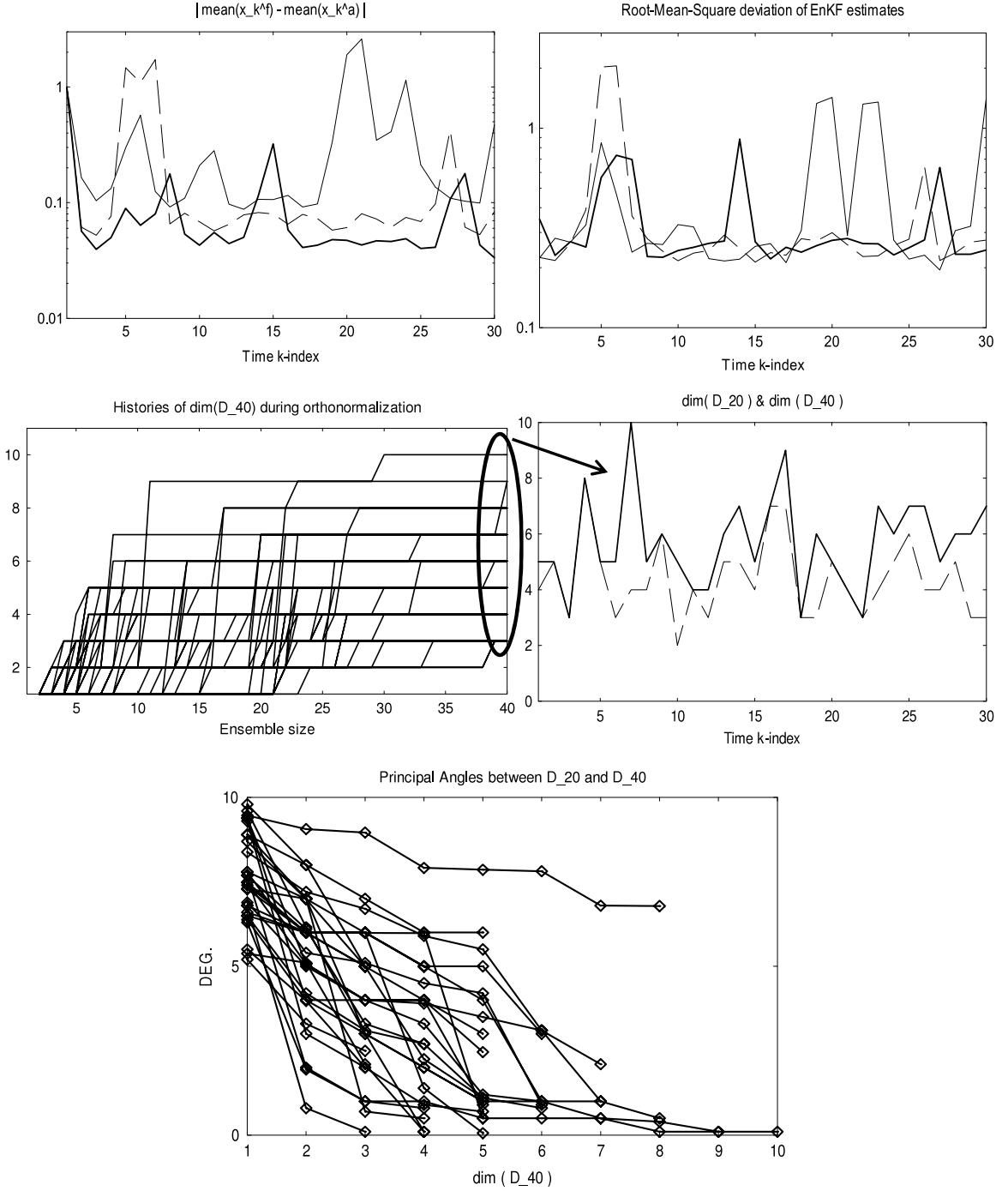


Fig. 5. Inversion with EnKF (ensemble of size 10, 20 and 40) in a mono-point \mathbf{u} case. Deviation between EnKF analysis and forecast estimates during EnKF-gradient iterations and the RMS of EnKF \mathbf{x}_k^f estimates. The histories of the Gram–Schmidt orthonormalization during ensemble iterations (middle-left) and the dimension of the free subspace generator for spaces D_{20} and D_{40} (middle-right) explains why an ensemble of size 10 seems to be efficient. Lower: principal angles between subspaces D_{20} and D_{40} defined in Section 5.3.

where $\mathbf{x}_k = \mathbf{x}_{k-1} - \rho \nabla_{\mathbf{x}} j(\mathbf{x}_{k-1}, \mathbf{u})$ with $\partial \mathbf{B}_{0.95}$ is the boundary of the VaR_{0.95}-based variability ball defined by (8) around \mathbf{x}_k . Knowing the correlation matrix Q of Gaussian variables \mathbf{x}_k , the extreme scenarios have their i th component

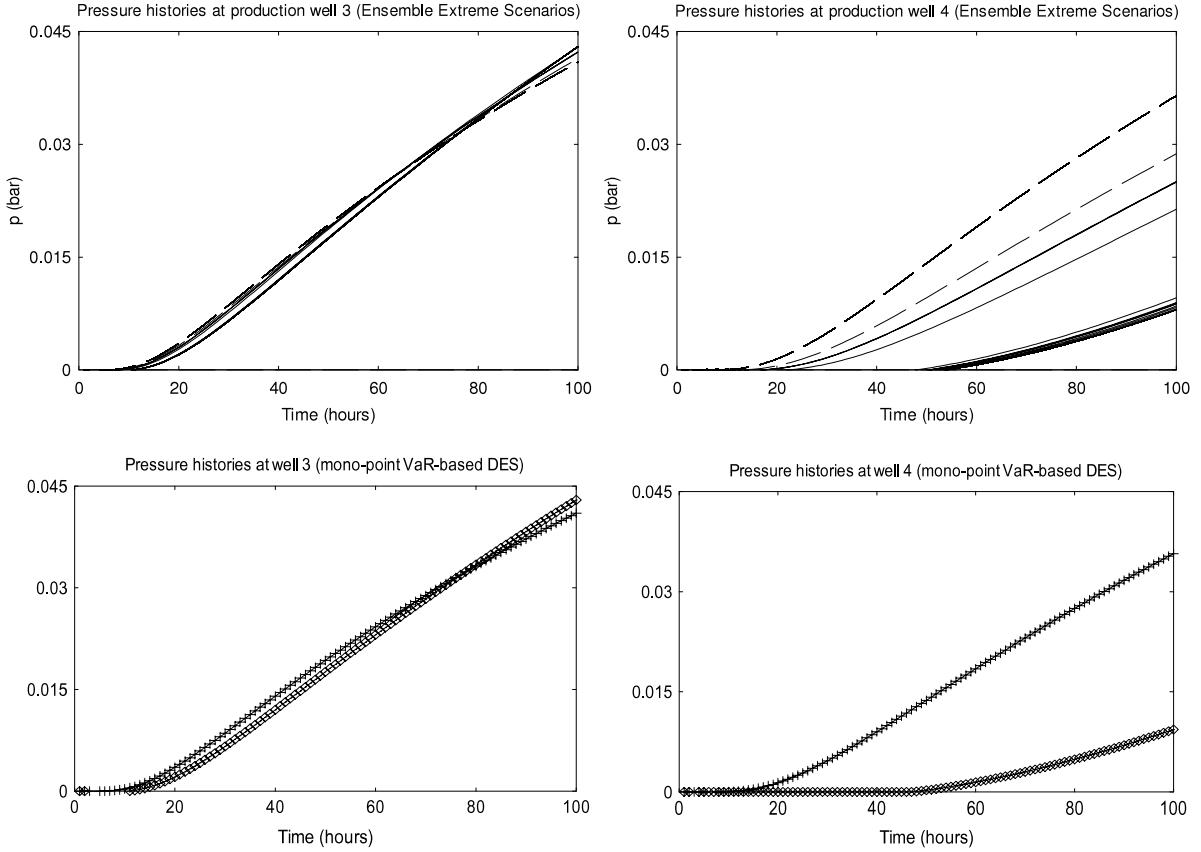


Fig. 6. Inversion with EnKF (ensemble of size 40) in a mono-point \mathbf{u} case. Upper: pressure histories for ensemble extreme scenarios $(\mathbf{x}_{30}^l)^{\pm}, l = 1, \dots, 40$ described in Section 6.1.1. Production at well 4 is more sensitive to the variability in \mathbf{x} . Lower: pressure histories for the two VaR-based scenarios given by (18) with a single-point gradient based minimization (no ensemble simulation involved).

defined by:

$$(\mathbf{x}_k)^{\pm}|_i = \mathbf{x}_k|_i \left(1 \pm Q \frac{\nabla_{\mathbf{x}} j(\mathbf{x}_k, \mathbf{u})}{\|\nabla_{\mathbf{x}} j\|}|_i \right). \quad (18)$$

In our case Q is diagonal ($Q = 0.02I_{100}$) and from Section 4 we have

$$\text{VaR}_{0.95}(N(0, 0.02)) = 0.02 \text{ VaR}_{0.95}(N(0, 1)) = 0.02 \times 1.65 = 0.032,$$

which defines the value of the i th component of \mathbf{x}_k^{\pm} following the sign of the i th component of $\nabla_{\mathbf{x}} j(\mathbf{x}_k, \mathbf{u})$: $(\mathbf{x}_k)^{\pm}|_i = \mathbf{x}_k|_i (1 \pm 0.032 \text{ sgn}(\nabla_{\mathbf{x}} j)_i)$. In the same way, $2q$ $(\mathbf{x}_k^l)^{\pm}$ VaR-based Ensemble Directional Extreme Scenarios can be defined applying (18) to each of the ensemble members $\mathbf{x}_k^l, l = 1, \dots, q$.

Fig. 6 shows pressure histories for these ensemble extreme scenarios after an inversion with an ensemble of size $q = 40$ giving 80 scenarios $\{(\mathbf{x}_{30}^l)^{\pm}, l = 1, \dots, 40\}$ after 30 iterations of EnKF. As reported in Fig. 4, the production at well 4 is more sensitive to the variability in \mathbf{x} than at well 3. The figure also compares these with the pressure histories for the two VaR-based scenarios $(\mathbf{x}_{30})^{\pm}$ by (18). This low-complexity estimation of the impact of the variability in \mathbf{x} gives satisfactory bounds in this single-point situation. The situation is less satisfactory in the multi-point situation justifying ensemble simulations.

6.2. Multi-point inversion under uncertainties

Now, we consider the two components of the control \mathbf{u} (pressure at injection wells) not anymore single valued but each defined through an interval with a uniform PDF: $u \in \mathbf{I} = [0.8, 1.2] \times [1.1, 1.5]$. We would like to have the

inversion to be robust also with respect to this variability. The multi-point minimization algorithm given in Section 2.1 is cast into the Kalman filters framework as described in Section 5.1 to control the first two moments μ and σ of j defined in Section 2 accounting for the variability on both \mathbf{x} and \mathbf{u} .

We consider again three ensembles of size $q = 10, 20$ and 40 . The multi-point optimization of 2.1 is defined for \mathbf{I}_m being a 10×10 sampling of \mathbf{I} .

Fig. 7 shows the histories over time k -index of the deviation between the ensemble analysis and forecast EnKF estimates and the root-mean-square deviation of the forecast estimates for three ensembles of size $10, 20$ and 40 . **Fig. 7** also shows the outcome of the Gram–Schmidt algorithm detecting independent directions in the global sensitivity space $D_q = \text{Span}\{d(\mu(\mathbf{x}_k^l), \sigma(\mathbf{x}_k^l)), l = 1, \dots, q\}$ defined in Section 5.3 and eventually its dimension which remains below 15. This is often larger than what we had in the mono-point situation in **Fig. 5**, but sometimes smaller (at given iteration) and would then mean that in these situations the multi-point direction constructed to account for the variability on \mathbf{u} actually reduces the effect of the variability on \mathbf{x} . In all cases, this explains why an ensemble of size 10 is clearly not enough rich. The figure also illustrates the application of the concept of principal angles between subspaces to estimate the relative orientation of the global sensitivity spaces D_{20} and D_{40} for the ensemble sizes of 20 and 40. These spaces are generated by the 20 (resp. 40) directions $d(\mu(\mathbf{x}_k^l), \sigma(\mathbf{x}_k^l))$ as described in Section 5.3. In addition to the dimension of the spaces, these angles provide additional information on possible impacts on the search spaces due to an enrichment of the ensemble. In the present case, the angles increase with the EnKF iterations which mean that one needs more accuracy with convergence and that the additional information by the ensemble of 40 is not present in what available with an ensemble simulation of size 20.

Fig. 8 shows the difference between the target \mathbf{x} and \mathbf{p} and the mean ensemble estimates for the ensemble of size 40, to be compared with **Fig. 4** in the mono-point situation for \mathbf{u} . The inversion gives a very different estimation for \mathbf{x} when the control \mathbf{u} is not anymore single valued. Also field deviations between the EnKF estimates and between these EnKF estimates and the solution by a deterministic inversion are different from what obtained in the mono-point situation showing the utility of the multi-point formulation. The pressure histories at production wells $s_{3,4}$ also deviate more from the targets which were recovered by both the inversions in the deterministic and uncertain mono-point cases. The deviation is not the same for the two wells. One should have in mind that there is no guarantee an inversion finding a distribution of \mathbf{x} maximizing the likelihood with respect to the uncertain parameters and observation data would produce a design giving the deterministic target pressure output. Indeed and surprisingly, despite one has lower deviation between ensemble pressure simulations for well 3 (**Fig. 8**, lower-right), the inversion under uncertainty produces a solution farther from the deterministic target than for well 4. This is a kind of information which help discriminating between different production points in term of reliability and necessary monitoring.

As in the mono-point situation, these can be compared to the two scenarios obtained by the Directional Extreme Scenarios (DES) concept presented in Section 5.3 which permits a low-complexity quantification of the uncertainties due to the variability on \mathbf{x} .

6.2.1. Var-based Directional Extreme Scenarios in the multi-point case

We proceed with the same analysis than in Section 6.1.1 in the single-point situation but using the directions $d(\mu(\mathbf{x}_k), \sigma(\mathbf{x}_k))$ instead of $\nabla_{\mathbf{x}} j(\mathbf{x}_k, \mathbf{u})$ and the multi-point control algorithm of Section 2.1 instead of a simple descent for j as presented earlier.

Fig. 9 (to be compared to **Fig. 6** in the mono-point case) shows pressure histories for the ensemble extreme scenarios after an inversion with the multi-point optimization algorithm of 2.1 and an ensemble of size $q = 40$ after 30 iterations of EnKF. The figure also compares these with the pressure histories for the two directional VaR-based scenarios $(\mathbf{x}_{30})^{\pm}$ defined by:

$$(\mathbf{x}_{30})^{\pm} = \mathbf{x}_{30} \left(1 \pm Q \frac{d(\mu(\mathbf{x}_{30}), \sigma(\mathbf{x}_{30}))}{\|d(\mu, \sigma)\|} \right). \quad (19)$$

This low-complexity estimation does not give satisfactory bounds for the extreme scenarios in this multi-point situation. This is unfortunate as otherwise there would be no need for an ensemble analysis to access the extreme scenarios. This situation shows that we need to be cautious, especially when uncertainties are present in both \mathbf{x} and \mathbf{u} . In these cases, the variability in the search direction $d(\mu(\mathbf{x}^l), \sigma(\mathbf{x}^l))$ is larger than what we have in the mono-point situation with $\nabla_{\mathbf{x}} j(\mathbf{x}^l, \mathbf{u})$ for $l = 1, \dots, q$. This can also be seen from increasing principal angles comparing

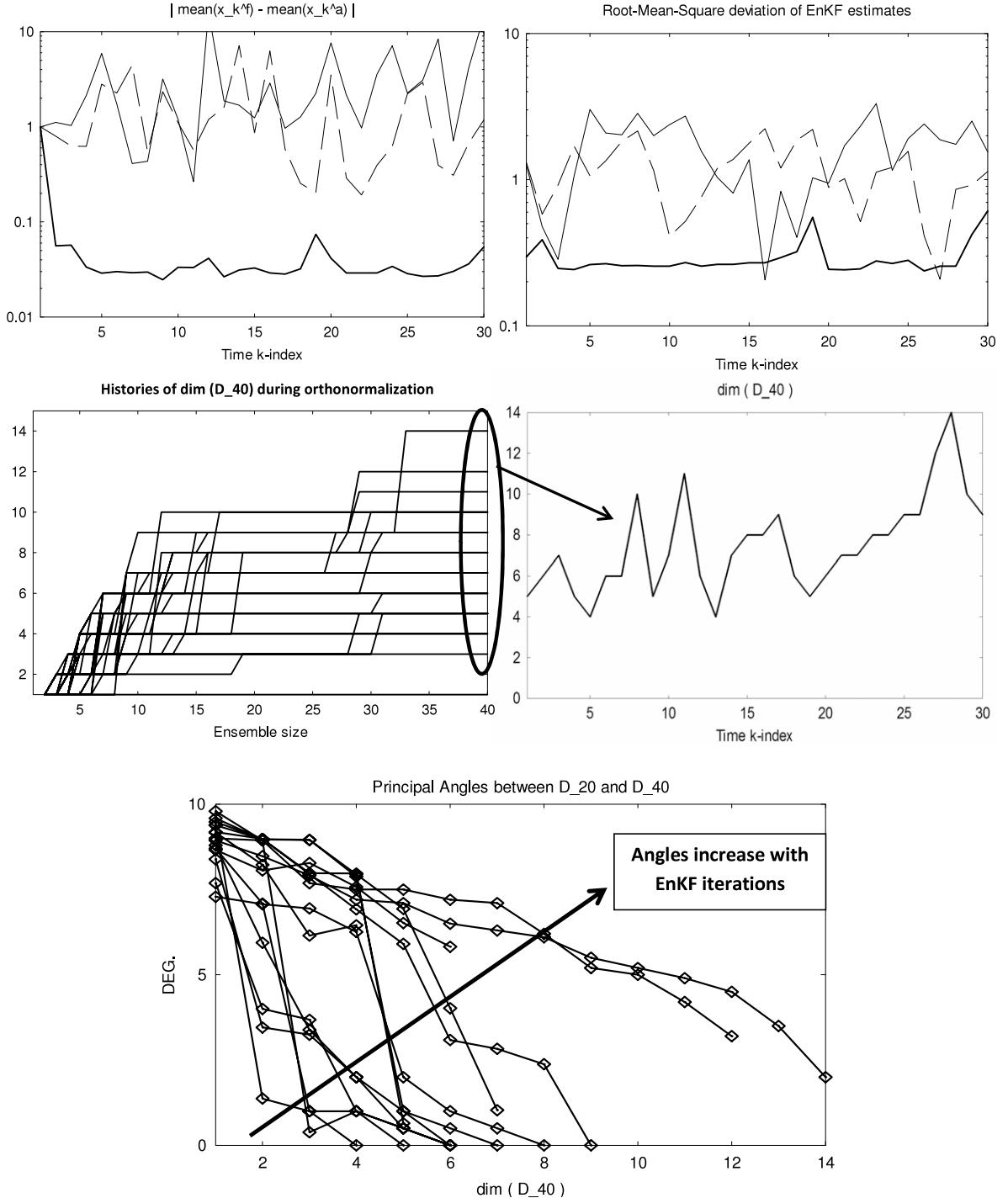


Fig. 7. Inversion with EnKF (ensemble of size 10, 20 and 40) in a multi-point \mathbf{u} configuration. Deviation between EnKF analysis and forecast estimates during EnKF-gradient iterations and the RMS of EnKF \mathbf{x}_k^f estimates. The histories of the Gram–Schmidt orthonormalization during ensemble iterations for the ensemble of size 40 (middle-left) and the dimension of the free subspace generator for space D_q (middle-right) to be compared to Fig. 5. Lower: principal angles between subspaces D_{20} and D_{40} defined in Section 5.3.

Figs. 5 and 7. Indeed, when accounting for the variability in \mathbf{u} (i.e. going from mono to multi-point) the angles between subspaces D_{20} and D_{40} increase as an indication of larger variability. As a consequence, when accounting

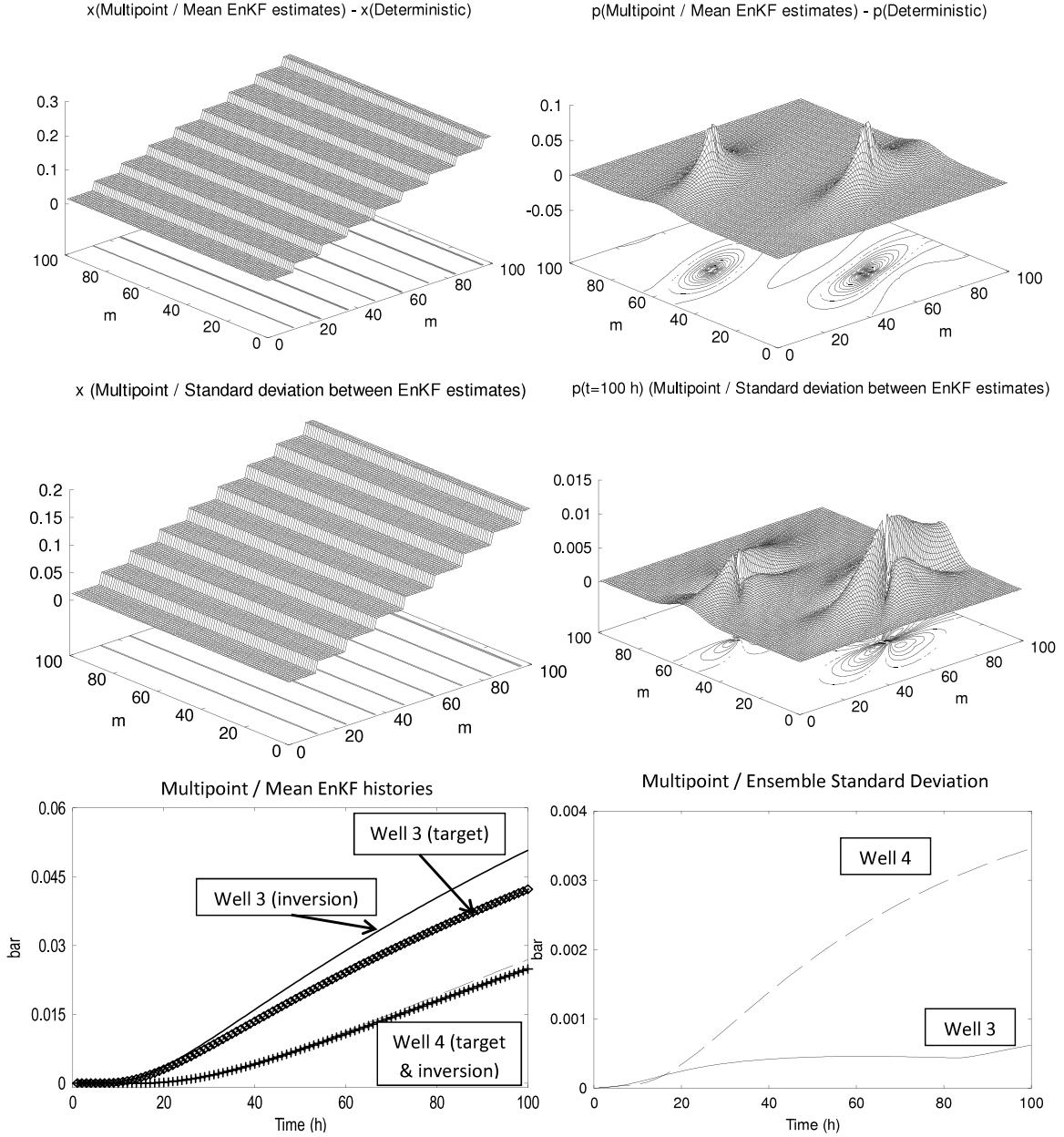


Fig. 8. Inversion with EnKF (ensemble of size 40) in a multi-point \mathbf{u} configuration. Differences between the ensemble mean solutions from the target \mathbf{x} and the field pressure $\mathbf{p}(t = 100h)$ and the deviation between ensemble solutions. Target pressure histories at production wells $s_{3,4}$ are not well recovered by the ensemble mean and the uncertainties on the production histories do not follow the same pattern than in the mono-point case. The results are quite different from those with a mono-point inversion in Fig. 4.

for the uncertainty on \mathbf{u} we need to a posteriori define the worst case scenarios from a more exhaustive ensemble analysis.

7. Concluding remarks

In order to be easily integrated in engineering environments and quantify our confidence on the optimal solution without intensive sampling of large dimensional parameter spaces a cascade of geometric uncertainty quantification concepts, with growing computational complexity, has been introduced.

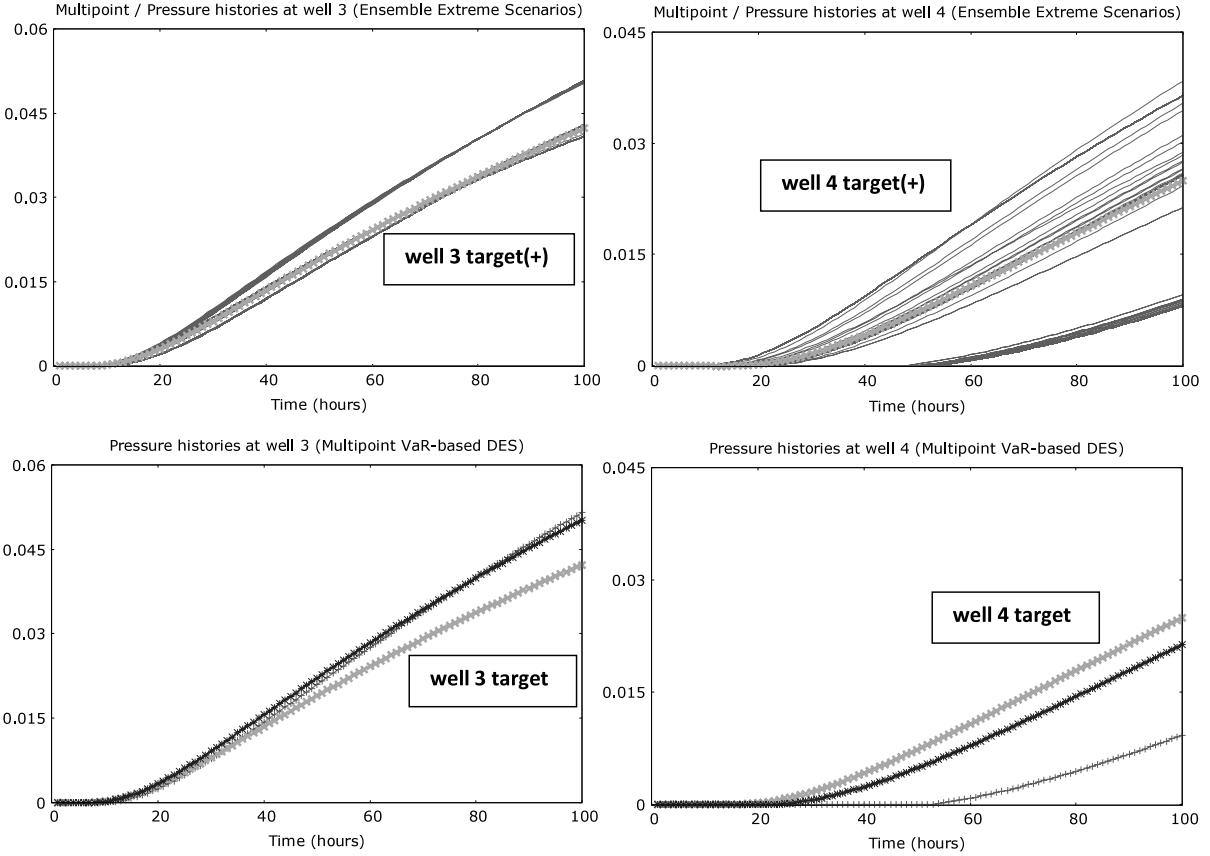


Fig. 9. Inversion with EnKF (ensemble of size 40) in a multi-point \mathbf{u} case. Upper: pressure histories for ensemble extreme scenarios $(\mathbf{x}_{30}^l)^{\pm}, l = 1, \dots, 40$ described in Section 6.1.1. Lower: pressure histories for the two directional VaR-based scenarios given by (19) with the multi-point minimization Algorithm 2.1 (no ensemble simulation involved) inefficient in this multi-point case.

The analysis starts with the geometric characterization of global sensitivity spaces through their dimensions and relative positions by the principal angles between global search subspaces. Then, joining a multi-point descent direction and extreme values information from the probability density functions of design variables the concept of Directional Extreme Scenarios (DES) has been introduced.

The construction goes beyond DES with Ensemble Kalman Filters (EnKF) after the multi-point optimization algorithm is cast into an ensemble simulation environment. This permits to account for the variability on the functioning parameters through the multi-point formulation and for the variability on the optimization parameters and observation data through the ensemble Kalman filter formulation.

The UQ cascade ends with the joint application of the EnKF and DES leading to the concept of Ensemble Directional Extreme Scenarios (EDES) which provides exhaustive possible extreme scenarios knowing the Probability Density Function of the optimization parameters and this without a sampling of the admissible space.

These ingredients have been illustrated on an history matching application. The interest of casting a multi-point minimization algorithm into an EnKF framework has been demonstrated when simple directional extreme scenarios do not provide satisfactory bounds for the impact of the uncertainties on the optimization parameters on the results.

The method (both the DES and EDES constructions) can theoretically be applied to any level of uncertainty on \mathbf{x} . But this will increase the regularity constraints on the functional over $B_a(\mathbf{x})$ as discussed in Sections 3 and 5.3. Also, the Kalman ensemble size will need to be much larger as the dimension of the ensemble sensitivity space D_q will certainly increase due to this variability.

The Gaussian framework is necessary for the Kalman filter framework. But, the other ingredients of the paper can be applied to non-Gaussian situations: dimensional analysis, principal angles, quantile-based extreme scenarios as far as one can (even by numerical integration of tabulated laws) find the a -quantiles quantifying the tail of the PDFs.

The framework presented features several sources of natural parallelism making the time to solution comparable to a mono-point deterministic inversion.

Looking at the bigger picture, this analysis shows the importance of considering all sources of variability in the problem to make sure the global search space is neither too large nor too small. The latter should obviously be avoided for the design to be robust. But, the former tells us that a more efficient design is maybe possible better monitoring the design space.

Acknowledgments

The adjoint code for the history matching problem has been obtained using Tapenade AD tool developed at INRIA-Sophia Antipolis by L. Hascoet and his team Tropics.

References

- [1] AIAA Guide for the Verification and Validation of Computational Fluid Dynamics Simulations, American Institute of Aeronautics and Astronautics, AIAA-G-077, 1998.
- [2] J. Caers, *Modeling Uncertainty in the Earth Sciences*, Wiley-Blackwell, 2011.
- [3] K.C. Kapur, L.R. Lamberson, *Reliability in Engineering Design*, John Wiley & Sons, New York, 1977.
- [4] B. Mohammadi, O. Pironneau, Shape optimization in fluid mechanics, *Ann. Rev. Fluid Mech.* 36 (1) (2004) 255–279.
- [5] J.R.R.A. Martins, A.B. Lambe, Multidisciplinary design optimization: A Survey of architectures, *AIAA J.* 51 (9) (2013) 2049–2075.
- [6] B. Mohammadi, Value at Risk for confidence level quantifications in robust engineering optimization, *Optimal Control Appl. Methods* (2013) <http://dx.doi.org/10.1002/oca.2061>.
- [7] C. Schillings, S. Schmidt, V. Schulz, Efficient shape optimization for certain and uncertain aerodynamic design, *Comput. & Fluids* 46 (1) (2011) 78–87.
- [8] P.G. Hoel, *Introduction to Mathematical Statistics*, John Wiley, London, 1971.
- [9] H.R. Lindman, *Analysis of Variance in Complex Experimental Designs*, Freeman, New York, 1974.
- [10] J.T. Spooner, M. Maggiore, R.O. Onez, K.M. Passino, *Stable Adaptive Control and Estimation for Nonlinear Systems: Neural and Fuzzy Approximator Techniques*, John Wiley, New York, 2002.
- [11] R. Ghanem, A. Doostan, On the construction and analysis of stochastic models: characterization and propagation of the errors associated with limited data, *J. Comput. Phys.* 217 (2006) 63–81.
- [12] G. Iaccarino, Quantification of Uncertainty in Flow Simulations Using Probabilistic Methods, VKI Lecture Series, 2008.
- [13] D. Xiu, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*, Princeton University Press, 2010.
- [14] R. Ghanem, P. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Springer Verlag, New York, 1991.
- [15] X. Wan, G.E. Karniadakis, Multi-element generalized polynomial chaos for arbitrary probability measures, *SIAM J. Sci. Comput.* 28 (3) (2006) 901–928.
- [16] P.M. Congedo, C. Correa, J.-M. Martinez, Shape optimization of an airfoil in a bzt flow with multiple-source uncertainties, *CMAME* 200 (2011) 16–32.
- [17] P. Cinnella, S.J. Hercus, Robust optimization of dense gas flows under uncertain operating conditions, *Comput. & Fluids* 39 (2010) 1893–1908.
- [18] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM, NY, 1987.
- [19] K. Ide, P. Courtier, M. Ghil, A. Lorenc, Unified notation for data assimilation: operational, sequential and variational, *J. Meteorol. Soc. Japan 75 (1B)* (1997) 181–189.
- [20] G. Casella, R. Berger, *Statistical Inference*, second ed., Duxbury Press, London, 2001.
- [21] R.E. Kalman, A new approach to linear filtering and prediction problems, *Trans. ASME. J. Basic Eng.* 82 (1960) 35–45.
- [22] S.A. Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of functions, *Dokl. Akad. Nauk SSSR* 148 (1963) 1042–1043; Russian Engl. Transl.: Soviet Math. Dokl. 4 (1963) 240–243.
- [23] H.-J. Bungartz, M. Griebel, Sparse grids, *Acta Numer.* 13 (2004) 147–269.
- [24] W.H. Schilders, H.A. Van der Vorst, J. Rommes, *Model Order Reduction: Theory, Research Aspects and Applications*, in: Springer Math in Industry Series, vol. 13, Berlin, 2008.
- [25] G. Obinata, B. Anderson, *Model Reduction for Control System Design*, Springer, Berlin, 2000.
- [26] Z. Qu, *Model Order Reduction Techniques with Applications in Finite Element Analysis*, Springer, Berlin, 2004.
- [27] K. Veroy, A. Patera, Certified real-time solution of the parametrized steady incompressible Navier–Stokes equations: Rigorous reduced-basis *a posteriori* error bounds, *Int. J. Numer. Meth. Fluids* 47 (8) (2005) 773–788.
- [28] B. Mohammadi, Reduced sampling and incomplete sensitivity for low-complexity robust parametric optimization, *Internat. J. Numer. Methods Fluids* (2013) <http://dx.doi.org/10.1002/fld.3798>.
- [29] F. Gallard, B. Mohammadi, M. Montagnac, M. Meaux, An adaptive multipoint formulation for robust parametric optimization, *J. Opt. Theory & Appl.* 165 (1) (2014) <http://dx.doi.org/10.1007/s10957-014-0595-6>.
- [30] C. Jordan, *Essay on geometry in n dimensions*, *Bull. Soc. Math. France* 3 (1875) 103–174.
- [31] H. Gluck, F. Warner, Great circle fibrations of the three-sphere, *Duke Math. J.* 50 (1983) 107–132.
- [32] S. Jiang, Angles between Euclidean subspaces, *Geom. Dedicata* 63 (2) (1996) 113–121.
- [33] C. Shonkwiler, *Poincaré duality angles for riemannian manifolds with boundary* (Ph.D. thesis), Univ. Pennsylvania, 2009.

- [34] Ph. Jorion, *Value at Risk: The New Benchmark for Managing Financial Risk*, McGraw-Hill, 2006.
- [35] B. Anderson, J. Moore, *Optimal Filtering*, Prentice-Hall, 1979.
- [36] A. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, NY, 1970.
- [37] A. Gelb, *Applied Optimal Estimation*, M.I.T Press, 1974.
- [38] G. Evensen, Advanced data assimilation for strongly nonlinear dynamics, *Mon. Weather Rev.* 125 (1997) 1342–1354.
- [39] G. Evensen, *Sequential Data Assimilation for Nonlinear Dynamics: The Ensemble Kalman Filter In Ocean Forecasting: Conceptual basis and applications*, Springer-Verlag, Heidelberg, 2002.
- [40] B. Mohammadi, Principal angles between subspaces and reduced order modeling accuracy in optimization, *Struct. Multidiscip. Optim.* (2014) <http://dx.doi.org/10.1007/s00158-013-1043-1>.
- [41] Z. Tang, J. Periaux, Uncertainty based robust optimization method for drag minimization problems in aerodynamics, *CMAME* 217 (220) (2012) 12–24.
- [42] R.E. Melchers, *Structural Reliability Analysis and Prediction*, John Wiley and Sons, Chichester, 1999.
- [43] A. Conn, K. Scheinberg, L. Vicente, *Introduction to Derivative-Free Optimization*, SIAM, NY, 2002.
- [44] J. Jahn, *Vector Optimization: Theory, Applications and Extensions*, Springer, 2004.
- [45] J. Nocedal, S. Wright, *Numerical Optimization*, Springer, NY, 2006.
- [46] B. Mohammadi, Uncertainty Quantification by geometric characterization of sensitivity spaces, *Comput. Methods Appl. Mech. Engrg.* (2014) <http://dx.doi.org/10.1016/j.cma.2014.07.021>.
- [47] S.J. Julier, J.K. Uhlmann, A new extension of the Kalman filter to nonlinear systems, in: Proc. of AeroSense: The 11th Int. Symp. on Aerospace/Defence Sensing, Simulation and Controls, 1997.
- [48] A. Seiler, G. Evensen, J.A. Skjervheim1, J. Hove1, J.G. Vabo, Using the EnKF for history matching and uncertainty quantification of complex reservoir models, in: L. Biegler, G. Biros, O. Ghattas, M. Heinkenschloss, D. Keyes, B. Mallick, L. Tenorio, B. van Bloemen Waanders, K. Willcox (Eds.), *Computational Methods for Large-Scale Inverse Problems and Quantification of Uncertainty*, John Wiley & Sons, Ltd., 2001.
- [49] N. Liu, O. Dean, Ensemble Kalman filter for automatic history matching of geologic facies, *J. Petrol. Sci. Eng.* 47 (2005) 147–161.
- [50] L. Heidaria, V. Gervaisa, M. Le Ravaleca, H. Hans Wackernagelb, History matching of petroleum reservoir models by the Ensemble Kalman Filter and parameterization methods, *Comput. Geosci.* 55 (2013) 84–95.
- [51] H. Darcy, *Les Fontaines Publiques de la Ville de Dijon*, Victor Dalmont, Paris, 1856, Translated in english by P. Bobeck: *The Public Fountains of the City of Dijon*, Kendall Hunt Publishing, Dubuque, 2004.
- [52] G. Chavent, G. Cohen, J. Jaffre, Discontinuous upwinding and mixed finite element for two-phase flows in reservoir simulation, *Comput. Methods Appl. Mech. Engrg.* 47 (1984) 93–118.
- [53] L. Hascoet, V. Pascual, Tapenade 2.1 user's guide, INRIA Tech. Rep. RT-300, 2004.
- [54] A. Griewank, *Computational Derivatives*, Springer, New York, 2001.