HAL
open science

Optimal well-placement using probabilistic learning

Roger Ghanem, Christian Soize, Charanraj Thimmisetty

To cite this version:

Roger Ghanem, Christian Soize, Charanraj Thimmisetty. Optimal well-placement using probabilistic learning. Data-Enabled Discovery and Applications, 2018, 2 (1), 10.1007/s41688-017-0014-x . hal-01703255

HAL Id: hal-01703255

https://hal.science/hal-01703255v1

Submitted on 7 Feb 2018

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.

Noname manuscript No.
(will be inserted by the editor)

Optimal well-placement using probabilistic learning

Roger Ghanem · Christian Soize · Charanraj Thimmisetty

the date of receipt and acceptance should be inserted later

Abstract A new method based on manifold sampling is presented for formulating and solving the optimal well-placement problem in an uncertain reservoir. The method addresses the compounded computational challenge associated with statistical sampling at each iteration of the optimization process. An estimation of the joint probability density function between well locations and production levels is achieved using a small number of expensive function calls to a reservoir simulator. Additional realizations of production levels, conditioned on well locations and required for evaluating the probabilistic objective function, are then obtained by sampling this jpdf without recourse to the reservoir simulator.

Notations

A lower case letter x is a real variable.
A boldface lower case letter $\mathbf{x}$ is a real vector.
An upper case letter X is a real random variable.
A boldface upper case letter $\mathbf{X}$ is a real random vector.
A lower case letter between brackets $[x]$ is a real matrix.
A boldface upper case letter between brackets $[\mathbf{X}]$ is a real random matrix.
$\mathbb{N} = \{0, 1, 2, \ldots\}$: set of all the integers.
$\mathbb{R}$: set of all the real numbers.
$\mathbb{R}^n$: Euclidean vector space on $\mathbb{R}$ of dimension n.

R. Ghanem
210 KAP Hall
University of Southern California
Los Angeles, CA 90089, USA E-mail: ghanem@usc.edu

C. Soize
Université Paris-Est, Laboratoire Modélisation et Simulation Multi-Echelle
MSME UMR 8208 CNRS
5 bd Descartes
77454 Marne-la-Vallée, France
E-mail: christian.soize@univ-paris-est.fr

C. Thimmisetty
Center for Applied Scientific Computing
Lawrence Livermore National Laboratory
Livermore, CA 94551, USA E-mail: thimmisetty1@llnl.gov

$\|\mathbf{x}\|$: usual Euclidean norm in $\mathbb{R}^n$.
$\mathbb{M}_{n,N}$: set of all the $(n \times N)$ real matrices.
$\mathcal{C}_{\mathbf{w}}$: admissible set for optimization
$\mathcal{C}_{\mathbf{w}_0}$: set of N_0 initial design variables
$\mathcal{C}_{\mathbf{w}_g}$: set of N_g design variables on search grid
$\mathbb{1}_{\mathcal{A}}(a)$ is the indicator function of set $\mathcal{A}$: $\mathbb{1}_{\mathcal{A}}(a) = 1$ if $a \in \mathcal{A}$ and $= 0$ if $a \notin \mathcal{A}$.
E: Mathematical expectation.
α: production confidence level for optimal solution
$\mathbf{w}^{\text{opt}}$: optimal solution
$\mathbf{w}_r^{\text{opt}}$: reference optimal solution
$\mathbf{w}_d^{\text{opt}}$: optimal solution using statistical surrogate
$\mathbf{w}_{\text{ar}}, \mathbf{x}_{\text{ar}}$: additional realizations of $\mathbf{W}$ and $\mathbf{X}$
q_r^{opt}: production level at $\mathbf{w}_r^{\text{opt}}$
q_d^{opt}: production level at $\mathbf{w}_d^{\text{opt}}$
$\mathcal{Q}(\mathbf{w})$: random cumulative production after 2,000 days
$F_{\mathcal{Q}(\mathbf{w})}(q)$: probability distribution function for $h(q, \mathbf{w})$:$1 - F_{\mathcal{Q}(\mathbf{w})}(q)$
$p_{\mathcal{Q}(\mathbf{w})}(q; \mathbf{w})$: probability density function for $\mathcal{Q}$ at given $\mathbf{w}$
$\mathcal{Q}$ at a given $\mathbf{w}$
x_I, y_I: coordinates of injection well
x_P, y_P: coordinates of production well
$\mathbf{W}$: random well locations
$\mathbf{Q}$: random production levels for given $\mathbf{W}$
$\mathbf{X}$: $(\mathbf{W}, \mathbf{Q})$
$\mathbf{w}_g$: optimization variable on search grid
$\mathbf{w}_0$: optimization variable on initial set of points
N_0: number of initially available solutions
N_g: number of points for grid search by optimization algorithm
θ: element of sample space Θ
n_{rep}: number of repetitions for each design variable
n_{MC}: number of samples drawn from statistical surrogate

1 Introduction

The *well-placement problem* is paramount to efficient reservoir production and development. Its objective, of maximizing return on investment, is challenged by the fact that the subsurface is neither illuminated nor does it lend itself to simple constitutive laws of flow and deformation. Thus a rational treatment of the problem must tackle significant uncertainty about the medium and the governing equations as well as the resolution of relevant interacting physical processes. Conceptually, the uncertainty challenge embeds the problem into an ensemble of statistically similar problems, requiring a simultaneous consideration of the ensemble. Depending on the form of the objective function, either an average or extreme values of this ensemble may be relevant for evaluating its optimal value. The complexity of the physics, on the other hand, is typically approached from the perspective of computational modeling, leading to increasingly finer spatio-temporal resolutions in order to accurately capture the evolution of distinct interacting physical processes. Even if a statistical model of the subsurface is already available, the well-placement problem thus entails the repetitive numerical solution of potentially large computational models. The burden of this repetitive evaluation is manifested both in statistical sampling and optimization. A number of simplifications have been pursued to address these challenges, ranging from focusing the objective function on statistical averages instead of extremes, to reducing the spatial or physics resolution of the computational model, and developing adapted

optimization logic [1, 20]. Clearly these assumptions have consequences both on the selected well placement and on the credibility of the computed optimal solution.

For instance particle swarm optimization (PSO) [17] and the simultaneous perturbation stochastic approximation (SPSA) [24, 4] were introduced to accelerate the convergence of the iterative process towards an optimal solution. The SPSA was also applied to an optimal well-placement problem under geologic uncertainty [18], where the optimal solution was the maximum over an average of a statistical ensemble of size 100 reflecting uncertainty in permeability. It was observed that accounting for geologic uncertainty in this manner had a noticeable effect on the optimal well-placement. Similar observations were noted in other related studies [28]. Notwithstanding its importance, the computational challenge is clearly exacerbated when accounting for uncertainty in the optimization problem. This is due to the need to treat, simultaneously, several credible scenarios. The number of such scenarios depends both on the level of scatter in observed data and on the risk attitude adopted by the relevant decision maker. Thus, if tails of production are critical for decision making, a larger statistical ensemble is required, whereas assessments based on mean and variance would be treated sufficiently well with a smaller ensemble. In this context, both robust optimization [13, 28], involving a combination of mean and variance, risk-based optimization involving probabilities of exceedance [2], conditional value-at-risk (CVaRO) [8] dealing with risks associated with low returns, and utility functions [16] have been considered. Several recent studies dealing with optimal well-placement have stressed the value of joint location/control optimization whereby both the location of the wells and production strategy over some time horizon are simultaneously optimized [5, 6, 18, 28, 29]. Clearly, in such cases the computational aspects of the optimization problem are significantly more challenging, specially when accounting for uncertainties in geology, investment, and pricing.

Specialized optimization algorithms have been proposed to tackle challenges stemming from the confluence of optimization and uncertainty. Some of these include a retrospective framework [27] where the intrinsic structure of genetic algorithms is leveraged to maximize the information gleaned from each expensive numerical calculation, a polynomial chaos methodology [3, 26] which develops adapted surrogates to expedite the optimization task, specialized workflows to reuse expensive function evaluations in a sensible manner [19], and statistical proxies [2, 25]. The present paper extends these statistical proxies, adapting them to intrinsic structure hidden within the data, thus allowing us to maximize the value of information being inferred form this data. In particular, we delineate the intrinsic structure using the method of diffusion manifolds, construct a joint statistical model of parameters (i.e. permeability), reservoir quantities of interest (i.e. production at producing well), and design variables (i.e. well locations) from a handful of expensive model-based simulations. This joint probability density function (jpdf) is then used, with no further recourse to the physics simulator, to estimate the full probability of production for a given set of well locations. The concentration of the joint occurrences around the intrinsic manifold permits the construction of accurate jpdf from fewer samples than would otherwise be possible.

Once this jpdf is available, the optimization problem proceeds without the need for additional calls to any physics-based analysis code such as reservoir simulator. Indeed conditional on a specific choice of the design variables, the probability distribution of production can be readily evaluated by integrating a scalar function (jpdf). There are two challenges to this picture. First, constructing such a joint pdf requires itself a very large number of joint observations of well locations and production levels. The number of such observations being a monotonic function of the magnitude of the associated scatter. Second, except for simple convex sets in high-dimensional space, sampling from arbitrary sets is not straightforward.

In this paper we tackle the above two challenges by introducing to the well-placement problem a recent procedure [23] that permits the optimization iterator and the statistical sampler to share numerical samples, thus greatly reducing the computational burden required to attain an accurate optimal solution with severe probabilistic objectives or constraints. The essential idea is to recognize that while a collection of numerical samples may be observed to have some particular scatter, the scatter in these same samples would be smaller if they are construed to fluctuate around some implicit structure (when such a structure indeed exists). With smaller fluctuations, the same samples can be described and smoothed more accurately using standard techniques of probability density estimation.

The mathematical underpinnings of the above observations were developed in a recent paper by the authors [23], which proposes a new methodology for generating realizations of a random vector that are statistically con-

sistent with a dataset of observations of this vector. The probability distribution of this random vector, while a-priori not known, will be learned from the data and is presumed to be concentrated on an unknown manifold that will itself also be learned as part of the formalism. A random matrix is introduced whose columns are independent copies of the random vector and for which the number of columns is the number of data points in the dataset. The approach is based on the use of (i) the multidimensional kernel-density estimation method for estimating the probability distribution of the random matrix [7, 21], (ii) a Markov Chain Monte Carlo method for generating realizations for the random matrix [22], (iii) the diffusion-maps approach for discovering and characterizing the intrinsic geometry and the structure of the dataset [11, 10], and (iv) a reduced-order representation of the random matrix, which is constructed using the diffusion-maps basis associated with the dominant eigenvalues of an underlying transition matrix relative to the given dataset [23].

We rely on this manifold characterization and sampling capability to construct, with very few samples, an accurate estimate of the joint density function of well locations and production levels. We then integrate this jpdf into a standard constrained optimization procedure to improve its convergence. This extension is applied to the well-placement problem with tail probabilities in the objective function. This problem has been addressed using kriging and regression interpolation [1] of the objective function, resulting in the introduction of significant error in the tail of the distributions. Our approach does not make such assumptions, permitting the accurate and efficient evaluation of optimal solutions associated with statistics of extremes. A novel contribution of the present paper, in addition to new approach for the well-placement problem, is in adapting the sampling on manifold procedure [23] to the efficient evaluation of tail probabilities.

The next section of the paper introduces the physical and mathematical setups. This is followed by a presentation of the optimization algorithm for our well-placement objective function. We conclude with an assessment of the method.

2 Problem set-up

2.1 Physical problem

A typical oil reservoir consists of one or more injection and production wells, respectively used to inject water into and extract oil from the reservoir. Uncertainty in both micro-scale and large-scale features of the subsurface, such as reflected in permeability, renders the oil production and water production rates random. The objective of the well-placement problem is to place the injection and production wells such that some probabilistic measure of the cumulative oil production is maximized.

We consider a reservoir with one injection and one production well, and seek to optimally emplace them so as to maximize some measure of production. In particular, we seek to maximize the cumulative production level, after 2000 days from start of production, that can be certified at some confidence level $1 - \alpha$. While we recognize that a more realistic setting would entail several injection and several production wells, the efficacy and implementation of our proposed methodology are demonstrated sufficiently well with our proposed layout. Furthermore, we only aim to optimize well location and do not include production controls in the optimization task. As indicated in the introduction section, such joint optimization has already been investigated in the literature and could readily be coupled with our proposed framework.

We describe the physics of flow in the porous medium using a black-oil model with intrinsic permeability characterized as a lognormal stochastic process. We use the SPE-10 permeability data set [9], discretized into a $60 \times 220 \times 85$ regular Cartesian grid with real range of $600 \times 2200 \times 170$ (ft)^3. The 85 layers along depth are treated as statistically independent realizations of a two-dimensional permeability field, thus serving to construct the sample mean and sample covariance matrix. This is clearly a gross assumption, but one that can be easily relaxed in the presence of more extensive field data. The implementation, efficiency, and conclusions associated with our proposed approach are not affected by this assumption. With the covariance of the permeability field thus estimated, a twenty-term Karhunen-Loève expansion is constructed as a discretization of this random field

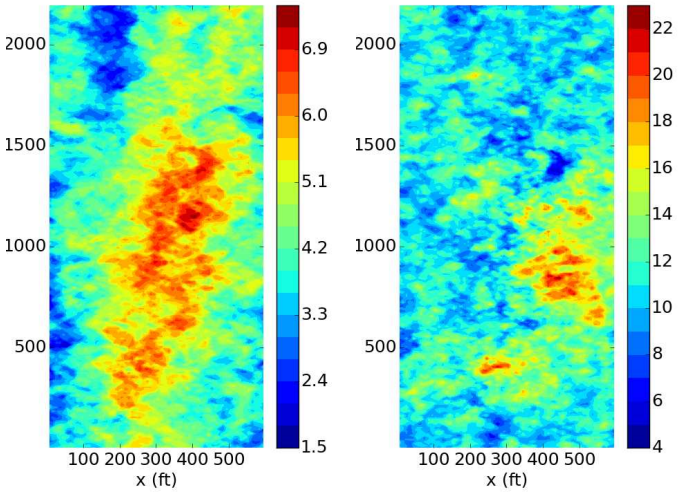

Fig. 1 Mean (left) and variance (right) of the log-permeability field.

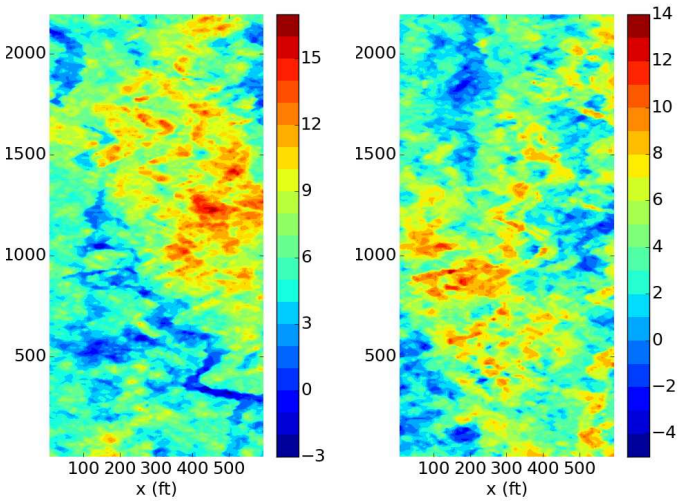

Fig. 2 Sample realizations of the log-permeability field.

using eigenvectors and eigenvalues of the sample covariance matrix. The mean and variance of the generated log-permeability field are shown in Fig. 1 and two sample realizations drawn from the Karhunen-Loève expansion are shown in Fig. 2. The permeability values shown in these figures have units of millidarcy.

Letting (x, y) be the coordinates in a cartesian system of the plane. Figure 3 shows the top view of the reservoir under consideration with arbitrarily placed injection and production wells. Here, (x_P, y_P) are the coordinates of a design production point P, and (x_I, y_I) are the coordinates of a design injection point I. In order to simplify the presentation, without loss of generality, while simultaneously demonstrating features of the proposed methodology,

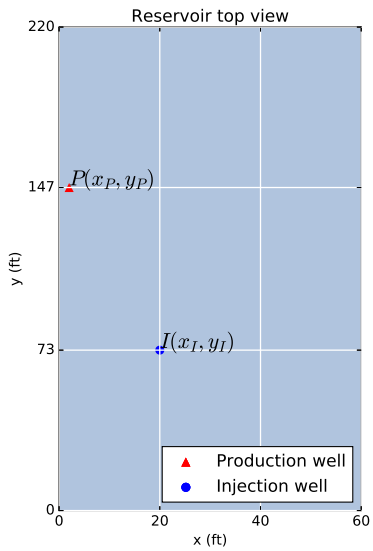

Fig. 3 Top view of the reservoir with arbitrarily located injection and production wells

we consider that the initial choice for each of the two wells is restricted to 16 spatial locations indicated by the grid in Fig. 3. This results in 256 (16×16) possible choices for the pair of (injection/production) well locations. Thus the reservoir simulator needs to be called for only 256 configurations, each with different location of the injection and the production well. As is subsequently explained, the constructed statistical proxy will then be able to identify optimal locations over continuous space, most likely not falling on the initial grid. We emphasize that for each of these 256 configurations, uncertainty in the subsurface is described through a Karhunen-Loève expansion of the intrinsic permeability field, as described above. Normally, sufficient samples must be evaluated (through reservoir simulations) for each configuration so as to accurately evaluate the averages of probabilities associated with the objective functions and constraints. As explained below, our proposed methodology requires far fewer samples, since these samples are shared between the optimizer and the statistical estimation and also since the statistical fluctuations are construed as variations around an intrinsic manifold in parameter space (and not within the whole hypercube defined by the uncertain parameters).

We use ECLIPSE [12] as our reservoir simulator. Each ECLIPSE run is completed by first generating 20 Gaussian variables, forming the Karhunen-Loève expansion, and taking the exponential of the sum to obtain one spatially varying realization of intrinsic permeability. A polynomial chaos (PCE) surrogate [14] is also constructed using several eclipse runs. A basis adaptation procedure is implemented to accelerate the convergence while addressing the curse of dimensionality [26]. The purpose of the PCE surrogate is to enable a thorough convergence analysis of the optimization procedure presented in the present paper. Convergence of this surrogate and its properties are presented elsewhere [26].

We note that the introduction of additional uncertainties beyond the permeability field does not present any challenge to the proposed methodology. The simultaneous observation of several parameters will serve to condition the quantity of interest, which should then exhibit smaller statistical fluctuations allowing it to be characterized with fewer samples. We also note that the introduction of additional parameters does not affect the computational effort in the proposed methodology, which is much more sensitive to the number of available data points. The added computational burden in these cases is typically dwarfed by the cost of a single simulation from the phsyics-based analysis code.

2.2 Optimization problem

Our objective will be to determine the locations, in a reservoir, of an injection well and a production well so as to maximize the production level, q, that is exceeded with some preset confidence $(1-\alpha)$. In this section, we define the mathematical constructs required to specify this problem as well as define required stochastic approximations.

Let $\mathbf{w} = (w_1, \ldots, w_{m_w})$ be a parameters vector that belongs to an admissible set $\mathcal{C}_{\mathbf{w}}$. In the present case, $m_w = 4$, and $\mathbf{w}$ denotes coordinates of the two wells (2 coordinates per well). We consider the general probabilistic non-convex optimization problem,

$$\mathbf{w}^{\text{opt}} = \arg \max_{\mathbf{w} \in \mathcal{C}_{\mathbf{w}}} f(\mathbf{w}) \,, \tag{1}$$

in which the scalar objective function $f(\mathbf{w})$, defined on $\mathcal{C}_{\mathbf{w}}$ with values in $\mathbb{R}$, is specified as

$$f(\mathbf{w}) = \arg \{q : \ \text{Proba}\{\mathcal{Q}(\mathbf{w}) > q\} = 1-\alpha\} \ , \tag{2}$$

where α is a given real number fixed in $]0\,,1[$. Equivalently, for each $\mathbf{w}$, we have $f(\mathbf{w}) = F^{-1}_{\mathcal{Q}(\mathbf{w})}(\alpha)$, with $F_{\mathcal{Q}(\mathbf{w})}(q)$ the cumulative distribution function (CDF) of random variables $\mathcal{Q}(\mathbf{w})$. Here, $\mathcal{Q}(\mathbf{w})$ is the cumulative production after $2{,}000$ days when the well locations have specified coordinates $\mathbf{w}$. Clearly, the evaluation of the objective $f(\mathbf{w})$ for every value of $\mathbf{w}$ is tantamount to a probability integral requiring several evaluations of $\mathcal{Q}$, the predicted cumulative production. The probability distribution on $\mathcal{Q}$ is assumed to be induced by uncertainty in model parameters of the underlying reservoir simulator, and thus for every $\mathbf{w}$, $\mathcal{Q}(\mathbf{w})$ is a random variable. For well-posed reservoir models (where production levels do not become infinite), this implies that $\{\mathcal{Q}(\mathbf{w}), \mathbf{w} \in \mathcal{C}_{\mathbf{w}}\}$ is a second-order scalar stochastic process indexed by $\mathcal{C}_{\mathbf{w}}$. For each $\mathbf{w}$ fixed in $\mathcal{C}_{\mathbf{w}}$, the random variable $\mathcal{Q}(\mathbf{w})$ is a mapping from sample space, Θ, into $\mathbb{R}$, which is assumed to have a probability density function $p_{\mathcal{Q}(\mathbf{w})}(r;\mathbf{w})$. The first argument of this jpdf reflects uncertainty in material permeability, while the second argument reflects uncertainty in well placement. Consequently one has,

$$E\{\mathcal{Q}(\mathbf{w})^2\} = \int_{-\infty}^{+\infty} r^2 \, p_{\mathcal{Q}(\mathbf{w})}(r;\mathbf{w}) \, dr \,. \tag{3}$$

It is also assumed that, for each $\mathbf{w}$ fixed in $\mathcal{C}_{\mathbf{w}}$, and each θ_ℓ in the sample space Θ the value $\mathcal{Q}(\mathbf{w};\theta_\ell)$ of production level is calculated by using a computational model (e.g. ECLIPSE[12]) with uncertainties, and that the probabilistic optimization problem defined by Eq. (1) has a unique solution $\mathbf{w}^{\text{opt}}$ in $\mathcal{C}_{\mathbf{w}}$. We will refer to the "stochastic computational model" when we mean that the computational model is being exercised a number of times as required by the probabilistic objective function.

The objective of this paper is to propose a probabilistic approach that allows for solving this probabilistic non-convex optimization problem by using a small number N of evaluations of the computational model.

Remarks. (i) Given the form of the objective function as stated in equation (2), it will be useful to introduce, for any given production level q and well locations $\mathbf{w}$, the quantity $h(q,\mathbf{w})$ with values in $[0\,,1]$ defined as,

$$h(q,\mathbf{w}) = \text{Proba}\{\mathcal{Q}(\mathbf{w}) > q\} = \int_q^{+\infty} p_{\mathcal{Q}(\mathbf{w})}(r;\mathbf{w}) \, dr \,. \tag{4}$$

(ii) The integral in this last equation can be evaluated as an expectation by rewriting it in the form,

$$h(q,\mathbf{w}) = E\{\mathbb{1}_{[q\,,+\infty[}(\mathcal{Q}(\mathbf{w}))\} \,, \tag{5}$$

in which E denotes the mathematical expectation and where $\mathbb{1}_{[q\,,+\infty[}(r) = 1$ if $r \in [q\,,+\infty[$ and $= 0$ otherwise. We could use the classical estimate of the mathematical expectation by using samples as follows. For every $\mathbf{w}$ given in $\mathcal{C}_{\mathbf{w}}$, the stochastic computational model would allow for calculating ν_s independent samples (or realizations),

$\mathcal{Q}(\mathbf{w};\theta_{\ell'})$ of random variable $\mathcal{Q}(\mathbf{w})$, for $\theta_{\ell'}$ in Θ with $\ell' = 1,\ldots,\nu_s$. For ν_s sufficiently large, an accurate estimate of $h(q,\mathbf{w})$ defined by Eq. (5) could then be computed as

$$h(q,\mathbf{w}) \simeq \frac{1}{\nu_s}\sum_{\ell'=1}^{\nu_s} \mathbb{1}_{[q\,,+\infty[}(\mathcal{Q}(\mathbf{w};\theta_{\ell'}))\,. \tag{6}$$

Using standard formalisms, if the optimization algorithm requires N_0 evaluations of f, then the stochastic computational model must be called $\nu_s \times N_0$ times, which could be prohibitive. The probabilistic approach that is proposed allows for drastically limiting the number of calls of the stochastic computational model to a value $N \ll \nu_s \times N_0$. Let $\mathcal{C}_{\mathbf{w}_0}$ be the set consisting of the N_0 values of the design variable $\mathbf{w}$ for which the stochastic computational model is called, and refer to these points as $\mathbf{w}_0^i$ for $i = 1,\ldots,N_0$. Thus $\mathcal{C}_{\mathbf{w}_0}$ is the initial design grid which, in our case, consists of $N_0 = 256$ points.

2.3 Data-driven model for uncertain parameters and design variables

In this section, we first define the dataset of a fixed number N_0 of data points denoted by $\mathbf{x}^\ell = (\mathbf{w}^\ell, q^\ell)$ for $\ell = 1,\ldots,N_0$ that are generated by the optimization algorithm and which require N_0 calls of the computational model. We then construct the random variables $\mathbf{W}$, Q, and $\mathbf{X} = (\mathbf{W}, \mathrm{Q})$ that admit these N_0 data points as independent samples. The dependence of these random variables on N_0 is omitted for notational convenience. We note that Q denotes the random production level conditional on well location, whereas $\mathcal{Q}$ denotes the random production level including uncertainty about well location. A distinction between the two random variables is important since this formalism is being used in an optimization setting where well location is the design variable.

Definition of the dataset. For any $\mathbf{w}$ fixed in $\mathcal{C}_\mathbf{w}$ and for any θ_ℓ fixed in Θ, let $\mathcal{Q}(\mathbf{w};\theta_\ell)$ be a sample of the random variable $\mathcal{Q}(\mathbf{w})$ computed by using the stochastic computational model. Let us now consider a fixed number N_0 of values $\mathbf{w}^1,\ldots,\mathbf{w}^{N_0}$ in $\mathcal{C}_\mathbf{w}$ of vector $\mathbf{w}$. These values can correspond either to a training procedure applied to $\mathbf{w}$ or are some values of $\mathbf{w}$ generated by the optimization algorithm. Let $q^1,\ldots,q^{N_0}$ be the real numbers in $\mathbb{R}$ that are associated with $\mathbf{w}^1,\ldots,\mathbf{w}^{N_0}$ in $\mathcal{C}_\mathbf{w}$ and that are defined by

$$q^\ell = \mathcal{Q}(\mathbf{w}^\ell;\theta_\ell) \quad , \quad \ell = 1,\ldots,N_0\,. \tag{7}$$

We then introduce the N_0 data points $\mathbf{x}^1,\ldots,\mathbf{x}^{N_0}$ in $\mathbb{R}^n$ obtained by pairing $\mathbf{w}^\ell$ with q^ℓ,

$$\mathbf{x}^\ell = (\mathbf{w}^\ell, q^\ell), \quad \ell = 1,\ldots,N_0\,, \tag{8}$$

where $n = m_w + 1$.

Definition of the random variables associated with the dataset. We now define the random variables $\mathbf{W}$, Q, and $\mathbf{X}$. Let $\mathbf{W} = (W_1,\ldots,W_{m_w})$ and Q be the statistically dependent second-order random variables with values in $\mathbb{R}^{m_w}$ and $\mathbb{R}$, respectively, for which the joint probability distribution on $\mathcal{C}_\mathbf{w} \times \mathbb{R}$ is unknown but for which a set of N_0 independent samples is given by $\mathbf{w}^\ell$, q^ℓ, $\ell = 1,\ldots,N_0$, $\mathbf{w}^\ell = (w_1^\ell,\ldots,w_{m_w}^\ell)$ and q^ℓ are N_0. Let $\mathbf{X} = (X_1,\ldots,X_n)$ be the second-order random variable with values in $\mathbb{R}^n$ defined by pairing $\mathbf{W}$ and Q,

$$\mathbf{X} = (\mathbf{W}, \mathrm{Q})\,. \tag{9}$$

While the probability density function $p_\mathbf{X}(\mathbf{x})$ of $\mathbf{X}$ is unknown, N_0 independent samples $\mathbf{x}^1,\ldots,\mathbf{x}^{N_0}$ are specified by Eq. (8).

Increasing the size of the dataset for large fluctuations. In some instances, such as the present case involving large variations in subsurface permeability, a single permeability realization for each design point may not be sufficient to capture the variability in the objective function, even when an intrinsic manifold has been discovered. To systematically address this issue, we introduce n_{rep} repetitions of the construction introduced in the previous

paragraph, resulting in $N = N_0 \times n_{\text{rep}}$ total number of calls to the expensive physics-based simulator. In section 3 below, we will investigate the convergence of the optimization algorithm as a function of n_{rep}. We will then observe that the required value of n_{rep} is significantly smaller than the number of repetitions typically required for optimization under uncertainty. As explained previously, this can be attributed both to sharing samples between optimization and expectation, and to the detection of the underlying manifold. *Remarks*.

(i) As the stochastic computational model defines a random mapping between parameter $\mathbf{w}$ and the random variable $\mathcal{Q}(\mathbf{w})$, the unknown probability density function (pdf) $p_{\mathbf{X}}$ is concentrated in a neighborhood of an unknown subset $\mathcal{S}_n$ of $\mathbb{R}^n$ (this neighborhood of subset $\mathcal{S}_n$ will be discovered by using the method that is briefly summarized in Appendix B).

(iii) For every $\mathbf{w}$ fixed in $\mathcal{C}_{\mathbf{w}}$, the probability distribution of the real-valued random variable $\mathcal{Q}(\mathbf{w})$ is not explicitly known.

Rewriting $h(q, \mathbf{w})$ in terms of a conditional probability density function. Taking into account the construction of the random variables $\mathbf{W}$ and Q, for all q fixed in $\mathbb{R}$ and for any given $\mathbf{w}_g$ fixed in $\mathcal{C}_{\mathbf{w}}$, the quantity $h(q, \mathbf{w}_g)$ defined by Eq. (4) can be rewritten as

$$h(q, \mathbf{w}_g) = \int_q^{+\infty} p_{\text{Q}|\mathbf{W}}(r|\mathbf{w}_g)\, dr\,, \tag{10}$$

in which $p_{\text{Q}|\mathbf{W}}(r|\mathbf{w}_g)$ is the conditional probability density function of the real-valued random variable Q given $\mathbf{W} = \mathbf{w}_g$.

2.4 Algorithm used for solving the optimization problem

The algorithm for evaluating $h(q, \mathbf{w}_g)$ by using a dataset is detailed in Appendix A, and the algorithm for solving the probabilistic optimization problem with a fixed number of function evaluations is detailed in Appendix B. The flow chart of the algorithm is summarized in Figs. 4 to 6 and briefly described next. In a first step shown in Figure 4, the physics-based computational model is exercised N times in order to construct an initial data set from which a statistical surrogate, relating well-placement to production level, will be learned. This set is denoted by $\mathbf{x}^\ell$. A joint probability density function (jpdf) for production, well location, and permeability, is constructed from these N samples. Clearly, the value of N must be sufficiently high to account for both dependence of production level on uncertain permeability and on well location. In a second step, depicted in Figure 5, an intrinsic structure of the dataset, in the form of a diffusion manifold, is delineated using diffusion maps algorithms. The step produces a set of basis functions that characterizes the manifold. Fluctuations in the N data points will henceforth be viewed as fluctuations around this intrinsic manifold. Also in this second step, a particular stochastic differential equation is constructed. The solution of this equation, at each step, is guaranteed to be a multivariate sample drawn from the jpdf in step 1 and also to lie on the manifold. We then draw n_{MC} such N-sized samples, resulting in a total of $\nu_{\text{sim}} = N \times n_{\text{MC}}$ samples. The stochastic differential equation is low dimensional and can be solved very efficiently, requiring typically a few minutes, independently of how expensive the solution of the physics code is. In a third step, shown in Figure 6, the new realizations are used in a stochastic optimization code, where statistical averages are carried out on samples synthesized from the statistical surrogate constructed in step 2, without any additional recourse to the phsyics-based code. Details of the underlying theory are presented elsewhere [23, 15].

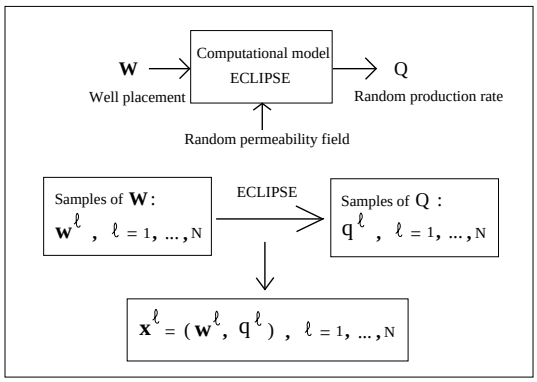

Fig. 4 Using the computational model, compute N samples $\{q^\ell, \ell = 1, \ldots, N\}$ of random variable Q from N realizations $\{\mathbf{w}^\ell, \ell = 1, \ldots, N\}$ of random variable **W**; construct the set $\{\mathbf{x}^\ell, \ell = 1, \ldots, N\}$ of N data points $\mathbf{x}^\ell = (\mathbf{w}^\ell, q^\ell)$

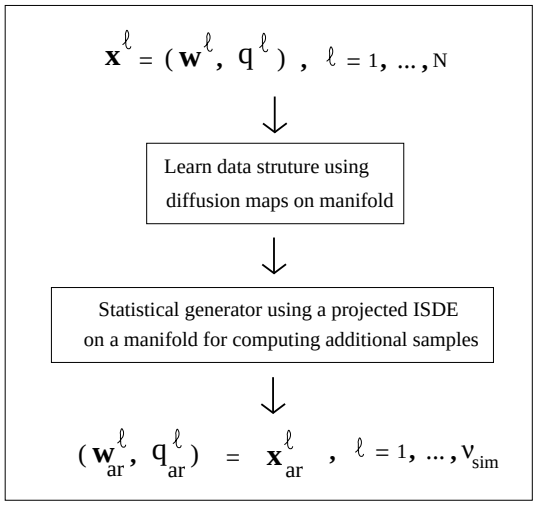

Fig. 5 From the set $\{\mathbf{x}^\ell, \ell = 1, \ldots, N\}$ of N data points $\mathbf{x}^\ell = (\mathbf{w}^\ell, q^\ell)$, using a probabilistic learning of the data structure and then using a statistical generator, compute $\nu_{\text{sim}} \gg N$ additional samples $\{\mathbf{x}^\ell_{\text{ar}}, \ell = 1, \ldots, \nu_{\text{sim}}\}$;construct the additional realizations $\mathbf{x}^\ell_{\text{ar}} = (\mathbf{w}^\ell_{\text{ar}}, q^\ell_{\text{ar}})$

3 Application

The well-placement problem is solved by relying on the machinery described in section 2. Implementation details and results are discussed in this section. To facilitate the interpretation of the results that follow, we reiterate some key concepts. **First**, we identify N_0 critical values of design variables. In our case these are 16 locations scattered over a coarse grid resulting in 256 possible permutations of well locations. **Second,** for each of these N_0 design variables, we draw one independent sample of permeability and solve one "expensive" ECLIPSE problem. **Third,** in order to enrich these N_0 function evaluations, we may have to draw a total of n_{rep} samples for each of the N_0 design points. With these $N = N_0 \times n_{\text{rep}}$ "expensive" function evaluations, we construct a statistical surrogate model which we then exercise (almost in real time and very inexpensively) to draw additional samples, as required by the optimization algorithm. We thus draw n_{MC} additional sets for a total of $N_0 \times n_{\text{rep}} \times n_{\text{MC}}$ samples available to the optimization code. Only N of these require calls to the "expensive" evaluator, with savings of the order of n_{MC}.

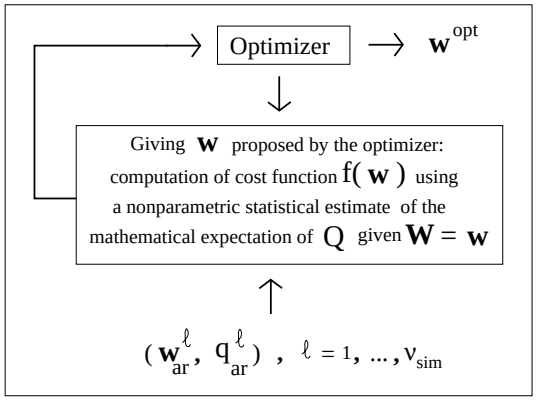

Fig. 6 Optimal solution sought using a random search algorithm, such as a grid search or a genetic algorithm, without new calls to the computational model, using only the additional samples $(\mathbf{w}_{\text{ar}}^{\ell}, q_{\text{ar}}^{\ell})$ for $\ell = 1, \ldots, \nu_{\text{sim}}$.

3.1 Description of the database and construction of the reference solution

We then have $m_w = 4$ and the vector $\mathbf{w} = (w_1, w_2, w_3, w_4)$ is such that $w_1 = x_P$, $w_2 = y_P$, $w_3 = x_I$, and $w_4 = y_I$. We also consider $N_0 = 256$ pairs $\{P, I\}$ of design points on the grid shown in Figure 3. Therefore, vector $\mathbf{w}$ with units in feet, belongs to the discrete subset $\mathcal{C}_{\mathbf{w}_0} = \{10, 200, 400, 600\} \times \{10, 730, 1470, 2200\} \times \{10, 200, 400, 600\} \times \{10, 730, 1470, 2200\}$ of $\mathbb{R}^4$ for which the N_0 elements are denoted by $\{\mathbf{w}_0^{\ell}, \ell = 1, \ldots, N_0\}$. For every given design point $\mathbf{w}_0^{\ell}$, $\nu_s = 100{,}000$ independent samples $\mathcal{Q}(\mathbf{w}_0^{\ell}; \theta_{\ell'}) \geq 0$ of the random production rate $\mathcal{Q}(\mathbf{w}_0^{\ell})$ are generated through the polynomial chaos surrogate and are denoted by $q_0^{\ell,\ell'} = \mathcal{Q}(\mathbf{w}_0^{\ell}; \theta_{\ell'})$ for $\ell' = 1, \ldots, \nu_s$. This represents the reference database which will be used to obtain the reference optimal solution. The objective of the proposed method is to achieve comparable optimal performance with only a fraction of these samples.

Generation of the database. Carrying out standard UQ analysis in 20-dimensional parameter space requires thousands of statistical samples to accurately characterize the distributions for each proposed pair of injection-production wells. In the present paper, it is assumed that such points are available at will, and are synthesized from the polynomial chaos decomposition that has been constructed. We will carry out a convergence study to explore the significance of the initial number of samples on the accuracy of optimization. Specifically, we will assume an upper bound of $\nu_s = 100{,}000$ samples $\mathcal{Q}(\mathbf{w}_0^{\ell}; \theta_{\ell'})$ for every design point represented by $\mathbf{w}_0^{\ell}$. Therefore, the database is made up of $256 \times 100{,}000$ values for $q_0^{\ell,\ell'}$ ($\ell = 1, \ldots, 256$ and $\ell' = 1, \ldots, 100{,}000$) and 256×4 values for $\mathbf{w}_0^{\ell}$ ($\ell = 1, \ldots, 256$).

Quantification of levels of statistical fluctuations. In order to clarify the level of the statistical fluctuations of $\{\mathcal{Q}(\mathbf{w}_0^{\ell}), \ell = 1, \ldots, N_0\}$, for each of the 256 design options, we compute the cumulative production level for 100,000 samples. We then compute the mean and standard deviations by averaging over these 100,000 samples. The coefficient of variation for each of these 256 designs is then estimated by dividing the sample standard deviation by the sample mean. Figure 7 displays the empirical estimate of the coefficient of variation (COV) for each of these 256 designs. The observed scatter in the COV is indicative of significant statistical fluctuations of production level across the design space, in addition to fluctuations induced by uncertainty in permeability. Such fluctuations would present significant computational challenges to standard optimization algorithms.

Construction of the reference solution and convergence analysis. In order to limit the length of the paper, the results are presented for only one value of the probability level α equal to 0.1. It should be noted that the objective of this paper is not to present a stochastic analysis of the physical problem under consideration (such an analysis can be

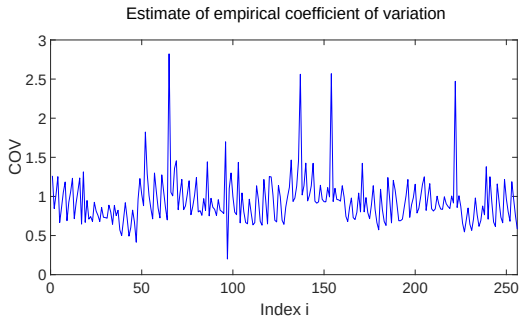

Fig. 7 Graph of the empirical estimate of the coefficient of variation of discrete stochastic process $\{\mathcal{Q}(\mathbf{w}_0^\ell), \ell = 1, \ldots, N_0\}$.

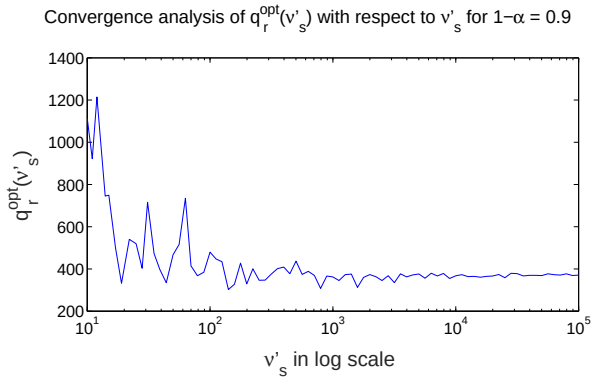

Fig. 8 Graph of the optimal solution $q_r^{\text{opt}}(\nu_s')$ for $10 \leq \nu_s' \leq 100{,}000$ in log scale.

found in [26]), but to use the generated database for demonstrating that the novel method presented is efficient for solving the probabilistic optimization problem by significantly reducing the number of calls to the computational physical model. The reference solution, $\mathbf{w}_r^{\text{opt}}$ is obtained by solving the optimization problem under uncertainty

$$\mathbf{w}_r^{\text{opt}} = \arg \max_{\mathbf{w}_0 \in \mathcal{C}_{\mathbf{w}_0}} f(\mathbf{w}_0) \quad , \quad q_r^{\text{opt}} = f(\mathbf{w}_0^{\text{opt}}) \, , \tag{11}$$

where the objective function $f(\mathbf{w}_0)$ is defined by Eq. (2) with $h(q, \mathbf{w}_0)$ (defined by Eq. (4)) is estimated by using Eq. (6) with $\nu_s = 100{,}000$. The optimal solution, which is simply calculated by taking the maximum on the discrete subset $\mathcal{C}_{\mathbf{w}_0}$ defined in the first paragraph of Section 3.1, is

$$\mathbf{w}_r^{\text{opt}} = (10, 1470, 600, 1470)(ft) \quad , \quad q_r^{\text{opt}} = 369.4 \text{ bbl} \, . \tag{12}$$

The convergence analysis of q_r^{opt} with respect to the number of samples is given by plotting in Figure 8, $q_r^{\text{opt}}(\nu_s')$ for $10 \leq \nu_s' \leq 100{,}000$. This figure shows that about $1{,}000$ samples per design point would suffice to converge the optimization algorithm to its optimal solution. We point out again that with our proposed statistical surrogate, the optimal solution is not restricted to $\mathcal{C}_{\mathbf{w}_0}$ and could be achieved anywhere within the spatial domain, as allowed by the optimization algorithm. Thus, if for instance a grid search algorithm is utilized, then the optimal solution is constrained to the underlying grid being searched. We thus further underscore a distinction between the grid associated with a grid search (which we label below as $\mathcal{C}_{\mathbf{w}_g}$ and the grid we define above ($\mathcal{C}_{\mathbf{w}_0}$), and which is associated with the construction of the probabilistic model (and which is associated with the initial data set obtained from the expensive computational model).

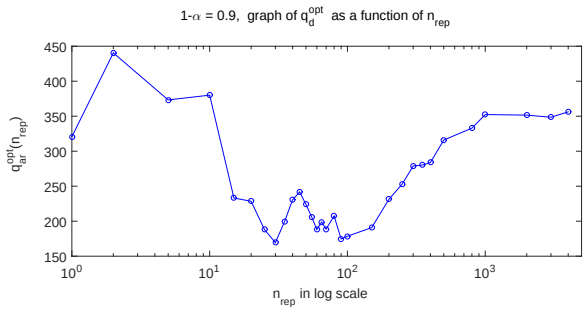

Fig. 9 Optimal production, q^{opt}, as function of number of repetitions n_{rep}; $\alpha = 0.1$; $1 \leq n_{\text{rep}} \leq 4\,000$ in log scale; units for q^{opt} in bbl.

3.2 Construction of the dataset from the database and validation of the algorithm for evaluating $h(q, \mathbf{w}_g)$

The objective of this section is to use the algorithm proposed in Appendix A for validating the nonparametric statistical estimation of $h(q, \mathbf{w}_g)$ given by Eqs. (27) to (29) and to analyze the convergence towards the reference solution with respect to N. As we have explained in Section 3.1, the level of statistical fluctuations are too big and $N_0 = 256$ may not be sufficient for constructing the dataset. Consequently, $n_{\text{rep}} >= 1$ repetitions are introduced such that $N = n_{\text{rep}} \times N_0$, and the convergence analysis with respect to n_{rep} is carried out in order to validate the formulation. Thus at each of the N_0 design points, n_{rep} random values of production levels are calculated using ECLIPSE. We recall that we have at our disposal (ν_s=100,000) samples already drawn from the polynomial chaos surrogate, which we are saving for such a convergence study.

Construction of a family of datasets from the database. Let us assume that n_{rep} is specified and fixed. Let $\ell'_1, \ldots, \ell'_{n_{\text{rep}}}$ be $n_{\text{rep}} \leq \nu_s$ pseudorandom integers drawn from the discrete uniform distribution on the interval $[1,\, \nu_s]$ (without repetition). The dataset $\{(\mathbf{w}^\ell, q^\ell), \ell = 1, \ldots, N\}$ (that depends on n_{rep}) is generated, for $i = 1, \ldots n_{\text{rep}}$ and for $j = 1, \ldots, N_0$, by

$$\mathbf{w}^\ell = \mathbf{w}_0^j \ , \ q^\ell = \mathcal{Q}(\mathbf{w}_0^j; \theta_{\ell'_i}) \ , \ \ell = j + (i-1) \times N_0 \, . \tag{13}$$

Solution of the optimization problem on a finite subset $\mathcal{C}_{\mathbf{w}_g}$ of $\mathcal{C}_{\mathbf{w}}$. In this paragraph, we construct the solution of the following probabilistic non-convex optimization problem on the finite subset $\mathcal{C}_{\mathbf{w}_g}$ of $\mathcal{C}_{\mathbf{w}}$ (introduced in Appendix A),

$$\mathbf{w}_d^{\text{opt}} = \arg \max_{\mathbf{w}_g \in \mathcal{C}_{\mathbf{w}_g}} f(\mathbf{w}_g) \quad , \quad q_d^{\text{opt}} = f(\mathbf{w}_g^{\text{opt}}) \, , \tag{14}$$

where the objective function $f(\mathbf{w}_g)$ is defined by Eq. (2), in which $h(q, \mathbf{w}_g)$ is estimated using the algorithm detailed in Appendix A. The calculated optimal solution $\mathbf{w}_d^{\text{opt}}$ and q_d^{opt} depends on n_{rep}. The finite set $\mathcal{C}_{\mathbf{w}_g}$ is made up of N_g elements that correspond to the nodes $\mathbf{w}_g^i = (w_{j_1}, w_{j_2}, w_{j_3}, w_{j_4}) \in \mathbb{R}^4$ of a grid in $\mathbb{R}^4$ such that $(j_1, j_2, j_3, j_4) \in \{1, \ldots, 12\}^4$ and where, for $k = 1, 2, 3, 4$, $w_{j_k} = 1 + (j_k - 1) \times \Delta_k$ with $\Delta_1 = \Delta_3 = 5.36$ and $\Delta_2 = \Delta_4 = 19.91$. Consequently, we have $i = 1, \ldots, N_g$ with $N_g = 12^4 = 20{,}736$. Note that this grid of $\mathbb{R}^4$ is such that there is no node $\mathbf{w}_g^i$ that exactly coincides with the reference solution $\mathbf{w}_r^{\text{opt}}$.

Validation of the algorithm for evaluating $h(q, \mathbf{w}_g)$. A validation of the proposed algorithm is obtained by solving Eq. (14) with the algorithm detailed in Appendix A and by comparing the obtained optimal solution $(\mathbf{w}_d^{\text{opt}}, q_d^{\text{opt}})$ with the reference solution $(\mathbf{w}_r^{\text{opt}}, q_r^{\text{opt}})$ defined by Eq. (12). The convergence analysis of $(\mathbf{w}_d^{\text{opt}}, q_d^{\text{opt}})$ with respect to the number n_{rep} of repetitions is shown in Figure 9, and the four components of function $\mathbf{w}_d^{\text{opt}}(n_{\text{rep}})$, are displayed in Figure 10. It can be seen that the convergence of the four coordinates (the four components of $\mathbf{w}_d^{\text{opt}}$ shown in Figure 10) is reached for $n_{\text{rep}} = 250$,

$$\mathbf{w}_d^{\text{opt}} = (10, 1404, 600, 1404) \ (ft) \quad , \quad q_r^{\text{opt}} = 253.2 \text{ bbl} \, , \tag{15}$$

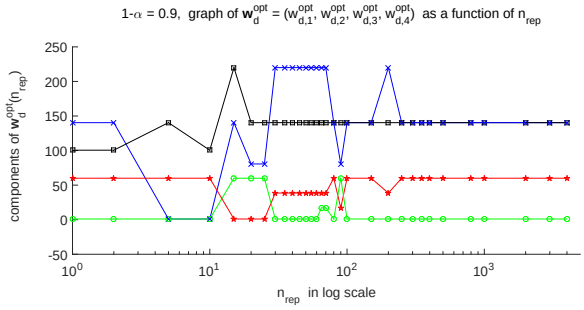

Fig. 10 Optimal well coordinates, $w_{d,1}^{\text{opt}}(n_{\text{rep}})$, as function of number of repetitions n_{rep}; $\alpha = 0.1$: $w_{d,1}^{\text{opt}}(n_{\text{rep}})$ (green circle), $w_{d,2}^{\text{opt}}(n_{\text{rep}})$ (black square), $w_{d,3}^{\text{opt}}(n_{\text{rep}})$ (red pentagram), $w_{d,4}^{\text{opt}}(n_{\text{rep}})$ (black cross); $1 \leq n_{\text{rep}} \leq 4\,000$ in log scale; vertical axis $\mathbf{w}_d^{\text{opt}}/10$ in ft.

which has to be compared to Eq. (12). As we have explained, the reference solution is obtained at the node $\mathbf{w}_r^{\text{opt}}$ belonging to the coarse grid $\{\mathbf{w}_0^\ell, \ell = 1, \ldots, N_0\}$ that does not coincide with a node of the finest grid $\{\mathbf{w}_g^i, i = 1, \ldots, N_g\}$. For $n_{\text{rep}} = 1\,000$ and $4\,000$, we have $q_r^{\text{opt}} = 352.4$ and 356.2 respectively.

3.3 Solving the probabilistic optimization problem with a fixed number of function evaluations

Solution of the optimization problem on the finite set $\mathcal{C}_{\mathbf{w}_g}$. For the reason given in Section 3.1, the results are presented for $\alpha = 0.1$. The algorithm detailed in Appendix B is used for solving the probabilistic non-convex optimization problem on the finite set $\mathcal{C}_{\mathbf{w}_g}$ that is defined in Section 3.2,

$$\mathbf{w}_{\text{ar}}^{\text{opt}} = \arg \max_{\mathbf{w}_g \in \mathcal{C}_{\mathbf{w}_g}} f(\mathbf{w}_g) \quad , \quad q_{\text{ar}}^{\text{opt}} = f(\mathbf{w}_0^{\text{opt}}) \, , \tag{16}$$

where the objective function $f(\mathbf{w}_g)$ is defined as before and where $h(q, \mathbf{w}_g)$ is estimated using Eq. (27) in which N is replaced by $\nu_{\text{sim}} = n_{\text{MC}} \times N$ with $N = n_{\text{rep}} \times N_0$ and $N_0 = 256$, and where n_{rep} belongs to the interval $[1\,, 100]$. As we have explained in Sections 3.1 and 3.2, the level of statistical fluctuations is too large and requires more samples than provided by the dataset $\{(\mathbf{w}_0^\ell, q_0^\ell), \ell = 1, \ldots, N_0\}$ with $N_0 = 256$. It has indeed been verified that the reference solution $\mathbf{w}_r^{\text{opt}}$ cannot be reached by $\mathbf{w}_{\text{ar}}^{\text{opt}}$ with $n_{\text{rep}} = 1$ and for any value of n_{MC}. Consequently, $n_{\text{rep}} > 1$ repetitions have been introduced (as in Section 3.2) and a convergence analysis with respect to n_{rep} has been carried out in the range $1 \leq n_{\text{rep}} \leq 100$ in order to validate the algorithm proposed in Appendix B. In such an analysis, for every fixed value of n_{rep} belonging to $[1\,, 100]$, the convergence analysis with respect to $n_{\text{MC}} > 1$ could be presented. Such analysis was performed, but in order to limit the presentation, the results presented herein are limited to those computed with $n_{\text{MC}} = 10$. For every value of n_{rep}, it was observed that $\mathbf{w}_{\text{ar}}^{\text{opt}}$ was improved for values of n_{MC} greater than 10 (justifying the choice $n_{\text{MC}} = 10$). Consequently, for n_{MC} fixed to 10, the optimal solution $(\mathbf{w}_{\text{ar}}^{\text{opt}}, q_{\text{ar}}^{\text{opt}})$ obtained by solving Eq. (16) depends only on n_{rep} and is thus rewritten as $(\mathbf{w}_{\text{ar}}^{\text{opt}}(n_{\text{rep}}), q_{\text{ar}}^{\text{opt}}(n_{\text{rep}}))$. The convergence analysis then consists in studying the dependence of $q_{\text{ar}}^{\text{opt}}(n_{\text{rep}}))$ and $\mathbf{w}_{\text{ar}}^{\text{opt}}(n_{\text{rep}}) = (w_{\text{ar},1}^{\text{opt}}(n_{\text{rep}}), w_{\text{ar},2}^{\text{opt}}(n_{\text{rep}}), w_{\text{ar},3}^{\text{opt}}(n_{\text{rep}}), w_{\text{ar},4}^{\text{opt}}(n_{\text{rep}})$ on n_{rep}.

Generation of the n_{MC} additional samples. In this paragraph, we give the values of the parameters that are used for generating the n_{MC} additional samples using the algorithm summarized in Appendix B and for which the notations are those used in [23]. We have: $m_w = 4$, $n = m_w + 1 = 5$, $N_0 = 256$, and $n_{\text{MC}} = 10$.

For every n_{rep} fixed in $[1\,, 100]$:

1. $N = N_0 \times n_{\text{rep}}$.

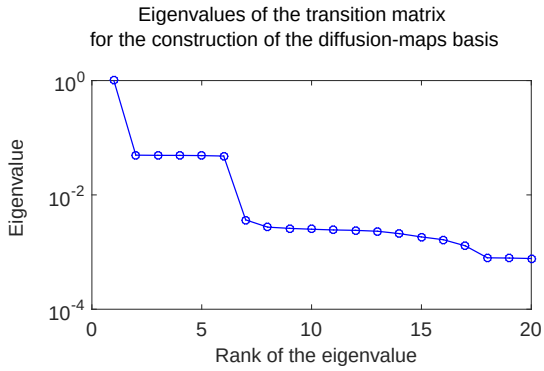

Fig. 11 Eigenvalues of the transition matrix.

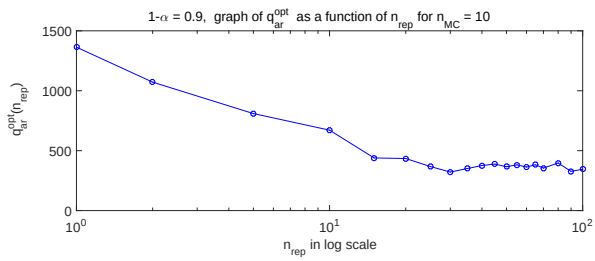

Fig. 12 Optimal production, $q_{\text{ar}}^{\text{opt}}(n_{\text{rep}})$ for $1 \leq n_{\text{rep}} \leq 100$ in log scale; units for q^{opt} in bbl; $\alpha = 0.1$.

2. The matrix $[x_d]$ in $\mathbb{M}_{n,N}$ for which the columns are the N data points $\mathbf{x}^1 \ldots \mathbf{x}^N$ with $\mathbf{x}^\ell = (\mathbf{w}^\ell, q^\ell) \in \mathbb{R}^n$ is generated by using Eq. (13).
3. A normalization of matrix $[x_d]$ is performed by using the principal component analysis for which all the eigenvectors are kept (no statistical reduction is done).
4. Concerning the transition matrix for the construction of the diffusion-maps basis, the smoothing parameter is $\varepsilon = 10$ and $\kappa = 1$. The eigenvalues of the transition matrix are displayed in Figure 11 with a clear delineation of a six dimensional eigen-space which we adopt for defining the diffusion manifold.
5. For the numerical integration of the reduced Itô stochastic differential equation for generating the additional samples, the parameters are $f_0 = 1.5$, $M_0 = 100$, and $\Delta r = 0.142$ (for $n_{\text{rep}} = 1$) to 0.0914 (for $n_{\text{rep}} = 100$).

Results and convergence analysis. The convergence analysis of $(\mathbf{w}_{\text{ar}}^{\text{opt}}, q_{\text{ar}}^{\text{opt}})$ with respect to the number n_{rep} of repetition is demonstrated in Figure 12 showing $q_{\text{ar}}^{\text{opt}}(n_{\text{rep}})$ as function of n_{rep} and Figure 13 showing the four components of $\mathbf{w}_{\text{ar}}^{\text{opt}}(n_{\text{rep}})$. It can be seen that the convergence of the four coordinates (the four components of $\mathbf{w}_{\text{ar}}^{\text{opt}}$ shown in Figure 13) is reached for $n_{\text{rep}} = 55$, yielding the following optimal solution,

$$\mathbf{w}_{\text{ar}}^{\text{opt}} = (10, 1404, 600, 1404)(ft) \quad , \quad q_r^{\text{opt}} = 381.5 \text{ bbl}\,, \tag{17}$$

which has to be compared to Eq. (12). As we have explained the reference solution is attained at node $\mathbf{w}_r^{\text{opt}}$ belonging to the coarse grid $\{\mathbf{w}_0^\ell, \ell = 1, \ldots, N_0\}$ that does not coincide with a node of the finest grid $\{\mathbf{w}_g^i, i = 1, \ldots, N_g\}$.

Gain obtained by using the algorithm proposed in Appendix B. As described in the introduction to this section, the gain is of the order of n_{MC}, which in the present case represents one order of magnitude.

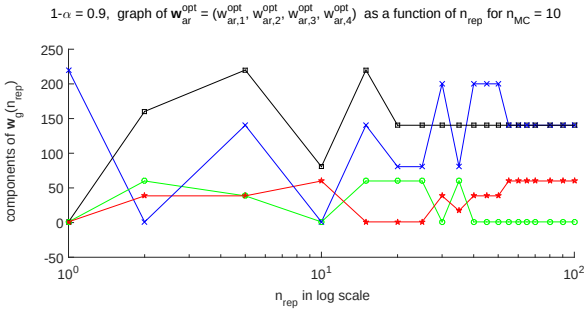

Fig. 13 Optimal coordinates for $\alpha = 0.1$, $w^{\text{opt}}_{\text{ar},1}(n_{\text{rep}})$ (green circle), $w^{\text{opt}}_{\text{ar},2}(n_{\text{rep}})$ (black square), $w^{\text{opt}}_{\text{ar},3}(n_{\text{rep}})$ (red pentagram), $w^{\text{opt}}_{\text{ar},4}(n_{\text{rep}})$ (black cross) for $1 \leq n_{\text{rep}} \leq 100$ in log scale; vertical axis $\mathbf{w}^{\text{opt}}_{\text{ar}}/10$ in ft.

4 Conclusions

A new method is presented for the solution of the optimal well-placement problem. The method tackles successfully the computational challenge where the number of iterations in an optimization setting is compounded by the number of samples required for evaluating the objective function. This is achieved through a very efficient construction of the joint density function between objective function and design variables, together with an associated sampling algorithm. The density function is constructed through a significantly smaller ensemble than is required using other methods. A convergence study is also performed to assess the influence of the number of initial points on the statistical accuracy of the method. The present work builds on recent developments by the authors characterizing the construction and stability of these manifolds. With these new perspectives, optimization problems under uncertainty can be accurately evaluated without recourse to probabilistic simplifications. The performance of the method relies on the existence of an underlying manifold that acts like an attractor to the statistical fluctuations. The shape of the manifold delineates the statistical dependence between inputs and outputs. Thus such a manifold is expected to underlie many physical phenomena where first principles induce such a dependence. It is also expected that the stronger the statistical fluctuations in the observed data (the N_0 original samples), the more valuable the proposed method will be since traditional methods would require many more samples to capture these fluctuations, whereas the proposed method tackles by securing additional realizations of production level from the statistical surrogate and not from the numerical simulator.

A Algorithm for evaluating $h(q, \mathbf{w}_g)$ by using a dataset

The available information is only constituted of the fixed number N of data points $\mathbf{x}^\ell = (\mathbf{w}^\ell, q^\ell)$ for $\ell = 1, \ldots, N$, which are N independent samples of the constructed random variable $\mathbf{X} = (\mathbf{W}, \mathrm{Q})$ defined by Eq. (9) (see Section 2.3). The problem consists in calculating, for any q fixed in $\mathbb{R}$ and for any point $\mathbf{w}_g$ fixed in $\mathcal{C}_{\mathbf{w}}$, an estimate of $h(q, \mathbf{w}_g)$ defined by Eq. (10). Such an estimate is computed by using the nonparametric statistics. As explained in Section 2.3, Eq. (6) cannot be used because, for given $\mathbf{w}^\ell$, only one sample q^ℓ is assumed to be known, and for using the classical statistical approach defined by Eq. (6), a large number of samples $q^{\ell,1}, \ldots, q^{\ell,\nu_s}$ of Q for $\mathbf{W} = \mathbf{w}^\ell$ would be necessary. In order to overcome this difficulty, a data smoothing technique based on the use of the Gaussian kernel-density estimation method is used for estimating $h(q, \mathbf{w}_g)$ defined by Eq. (10).

A.1 Nonparametric statistical estimation of $h(q, \mathbf{w}_g)$

In order to apply the nonparametric statistics for estimating $h(q, \mathbf{w}_g)$ on the basis of the dataset that is made up of the N independent realizations $\{(\mathbf{w}^\ell, q^\ell), \ell = 1, \ldots, N\}$, it is necessary to normalize the dataset in order to obtain well conditioned numerical calculations.

Normalizing the dataset. For $j = 1, \ldots, m_w$, let $\underline{w}_j$ and σ_j be the empirical estimates of the mean value and of the standard deviation of random component W_j of random vector $\mathbf{W}$, computed by using the N independent samples $\{w_j^\ell, \ell = 1, \ldots, N\}$ of W_j. Similarly, let $\underline{q}$ and σ be the empirical estimates of the mean value and of the standard deviation of random variable Q constructed with the N independent samples $\{q^\ell, \ell = 1, \ldots, N\}$ of Q. We then introduce the normalized random variable $\widehat{\mathrm{Q}}$ and, for $j = 1, \ldots, m_w$, the normalized random variable $\widehat{W}_j$, such that

$$\widehat{\mathrm{Q}} = (\mathrm{Q} - \underline{q})/\sigma \quad , \quad \widehat{W}_j = (W_j - \underline{w}_j)/\sigma_j \, , \tag{18}$$

for which the N independent samples are given, for $\ell = 1, \ldots, N$, by

$$\widehat{q}^\ell = (q^\ell - \underline{q})/\sigma \quad , \quad \widehat{w}_j^\ell = (w_j^\ell - \underline{w}_j)/\sigma_j \, . \tag{19}$$

The corresponding normalized value $\widehat{q}$ of q and the corresponding normalized value $\widehat{\mathbf{w}}_g = (\widehat{w}_{g,1}, \ldots, \widehat{w}_{g,m_w})$ of any $\mathbf{w}_g = (w_{g,1}, \ldots, w_{g,m_w})$ fixed in $\mathcal{C}_{\mathbf{w}}$, are written as

$$\widehat{q} = (q - \underline{q})/\sigma \quad , \quad \widehat{w}_{g,j} = (w_{g,j} - \underline{w}_j)/\sigma_j \, . \tag{20}$$

It can easily be verified that

$$h(q, \mathbf{w}_g) \simeq h(\widehat{q}, \widehat{\mathbf{w}}_g) \, , \tag{21}$$

$$h(\widehat{q}, \widehat{\mathbf{w}}_g) = \int_{\widehat{q}}^{+\infty} p_{\widehat{\mathrm{Q}}|\widehat{\mathbf{W}}}(\widehat{r}|\widehat{\mathbf{w}}_g) \, d\widehat{r} \, , \tag{22}$$

in which $p_{\widehat{Q}|\widehat{\mathbf{W}}}(\widehat{r}|\widehat{\mathbf{w}}_g)$ is the probability density function of the real-valued random variable $\widehat{\mathrm{Q}}$ given $\widehat{\mathbf{W}} = \widehat{\mathbf{w}}_g$. Introducing the joint pdf $p_{\widehat{\mathbf{W}},\widehat{\mathrm{Q}}}(\widehat{\mathbf{w}}, \widehat{r})$ of random variables $\widehat{\mathbf{W}}$ and $\widehat{\mathrm{Q}}$ and the pdf $p_{\widehat{\mathbf{W}}}(\widehat{\mathbf{w}}) = \int_{\mathbb{R}} p_{\widehat{\mathbf{W}},\widehat{\mathrm{Q}}}(\widehat{\mathbf{w}}, \widehat{r}) \, d\widehat{r}$ of random variable $\widehat{\mathbf{W}}$, Eq. (22) can be written as

$$h(\widehat{q}, \widehat{\mathbf{w}}_g) = \frac{1}{p_{\widehat{\mathbf{W}}}(\widehat{\mathbf{w}}_g)} \int_{\widehat{q}}^{+\infty} p_{\widehat{\mathbf{W}},\widehat{\mathrm{Q}}}(\widehat{\mathbf{w}}_g, \widehat{r}) \, d\widehat{r} \, . \tag{23}$$

Nonparametric statistical estimate of $h(\widehat{q}, \widehat{\mathbf{w}}_g)$. Each one of the dependent random variables $\widehat{W}_1, \ldots, \widehat{W}_{m_w}$ and $\widehat{\mathrm{Q}}$, has a zero empirical mean value and a unit empirical standard deviation (calculated by using the N samples $(\widehat{\mathbf{w}}^\ell, \widehat{q}^\ell)$). The nonparametric statistical estimation of the joint pdf $p_{\widehat{\mathbf{W}},\widehat{\mathrm{Q}}}(\widehat{\mathbf{w}}_g, \widehat{q})$ at point $(\widehat{\mathbf{w}}_g, \widehat{q})$, constructed by using the Gaussian kernel-density estimation method [7,21] and by using the N independent realizations $\{(\widehat{\mathbf{w}}^\ell, \widehat{q}^\ell), \ell = 1, \ldots, N\}$, is written, for N sufficiently large, as

$$p_{\widehat{\mathbf{W}},\widehat{\mathrm{Q}}}(\widehat{\mathbf{w}}_g, \widehat{q}) \simeq \frac{1}{N} \sum_{\ell=1}^{N} \frac{1}{(\sqrt{2\pi}\, s)^{m_w+1}} \exp\left\{ -\frac{1}{2s^2} \{ \|\widehat{\mathbf{w}}^\ell - \widehat{\mathbf{w}}_g\|^2 + (\widehat{q}^\ell - \widehat{q})^2 \} \right\} \, , \tag{24}$$

in which s is the bandwidth parameter that can be chosen as the usual multidimensional optimal Silverman bandwidth [22] (in taking into account that the empirical estimation of the standard deviation of each component is unity),

$$s = \left\{ \frac{4}{N(2 + m_w + 1)} \right\}^{1/(4+m_w+1)} \, . \tag{25}$$

From equation (24), it can be deduced that $p_{\widehat{\mathbf{W}}}(\widehat{\mathbf{w}}_g) = \int_{\mathbb{R}} p_{\widehat{\mathbf{W}},\widehat{\mathrm{Q}}}(\widehat{\mathbf{w}}_g, \widehat{r}) \, d\widehat{r}$ can be estimated, for N sufficiently large, by

$$p_{\widehat{\mathbf{W}}}(\widehat{\mathbf{w}}_g) \simeq \frac{1}{N} \sum_{\ell=1}^{N} \frac{1}{(\sqrt{2\pi}\, s)^{m_w}} \exp\left\{ -\frac{1}{2s^2} \|\widehat{\mathbf{w}}^\ell - \widehat{\mathbf{w}}_g\|^2 \right\} \, . \tag{26}$$

By using equations (23), (24), and (26), it can be deduced that, for N sufficiently large, an estimate of $h(\widehat{q}, \widehat{\mathbf{w}}_g)$ is given by

$$h(\widehat{q}, \widehat{\mathbf{w}}_g) \simeq \frac{\sum_{\ell=1}^{N} \widehat{r}^\ell(\widehat{q}) \, \widehat{v}^\ell(\widehat{\mathbf{w}}_g)}{\sum_{\ell=1}^{N} \widehat{v}^\ell(\widehat{\mathbf{w}}_g)} \, , \tag{27}$$

in which, for $\ell = 1, \ldots, N$, the real numbers $\widehat{v}^\ell(\widehat{\mathbf{w}}_g)$ and $\widehat{r}^\ell(\widehat{q})$ are written as

$$\widehat{v}^\ell(\widehat{\mathbf{w}}_g) = \exp\left\{ -\frac{1}{2s^2} \|\widehat{\mathbf{w}}^\ell - \widehat{\mathbf{w}}_g\|^2 \right\} \tag{28}$$

$$\widehat{r}^\ell(\widehat{q}) = \frac{1}{2} \left(1 - \mathrm{erf}\{(\widehat{q} - \widehat{q}^\ell)/(s\sqrt{2})\} \right) \, , \tag{29}$$

where $\mathrm{erf}\{y\} = \frac{2}{\sqrt{\pi}} \int_0^y e^{-t^2} \, dt$ is the erf function.

A.2 For given q and $\mathbf{w}$, algorithm for estimating $h(\widehat{q}, \widehat{\mathbf{w}}_g)$ by using only the given N data points

The proposed algorithm can be used for estimating $h(\widehat{q}, \widehat{\mathbf{w}}_g^i)$ in N_g given points $\mathbf{w}_g^i$ in $\mathcal{C}_{\mathbf{w}}$ belonging to the subset $\mathcal{C}_{\mathbf{w}_g} = \{\mathbf{w}_g^i = (w_{g,1}^i, \ldots, w_{g,m_w}^i), i = 1, \ldots, N_g\} \subset \mathcal{C}_{\mathbf{w}}$ by using only the N data points $\mathbf{x}^1, \ldots, \mathbf{x}^N$ in $\mathbb{R}^n$.

Data for the algorithm .

1. Given m_w and N.
2. set $n = m_w + 1$.
3. Given $\mathbf{x}^1, \ldots, \mathbf{x}^N$ in $\mathbb{R}^n$ such that $\mathbf{x}^\ell = (\mathbf{w}^\ell, q^\ell)$ is in $\mathbb{R}^n$ with $\mathbf{w}^\ell$ in $\mathbb{R}^{m_w}$ and q^ℓ in $\mathbb{R}$.
4. Given scalar q and $\mathbf{w}_g$ in $\mathcal{C}_{\mathbf{w}_g}$.

Steps of the algorithm for estimating $h(\widehat{q}, \widehat{\mathbf{w}}_g^i)$ by using only the given N data points.

1. Compute $\underline{q} = \frac{1}{N}\sum_{\ell=1}^N q^\ell$ and $\sigma_q^2 = \frac{1}{N}\sum_{\ell=1}^N (q^\ell - \underline{q})^2$.
2. For $j = 1, \ldots, m_w$, compute $\underline{w}_j = \frac{1}{N}\sum_{\ell=1}^N w_j^\ell$ and $\sigma_j^2 = \frac{1}{N}\sum_{\ell=1}^N (w_j^\ell - \underline{w}_j)^2$.
3. For $\ell = 1, \ldots, N$, compute $(\widehat{\mathbf{w}}^\ell, \widehat{q}^\ell) \in \mathbb{R}^{m_w} \times \mathbb{R}$ by using Eq. (19).
4. Compute $\widehat{\mathbf{w}}_g$ and $\widehat{q}$ by using Eq. (20).
5. Compute s by using Eq. (25).
6. For $\ell = 1, \ldots, N$,
   (a) compute $\widehat{v}^\ell(\widehat{\mathbf{w}}_g)$ by using Eq. (28).
   (b) compute $\widehat{r}^\ell(\widehat{q})$ by using Eq. (29).
7. Compute $h(\widehat{q}, \widehat{\mathbf{w}}_g)$ by using Eq. (27).

End of the algorithm.

B Proposed algorithm for solving the probabilistic optimization problem with a fixed number of function evaluations

The proposed algorithm used for solving the probabilistic optimization problem defined by equations (1) to (2), is based on a procedure that evaluates $h(q, \mathbf{w}_g)$ as described above while necessitating only a small number (N) of calls to the stochastic computational model. The proposed method consists of enriching the N samples that are represented by the columns of matrix $[x_d] = [\mathbf{x}^1 \ldots \mathbf{x}^N]$ in $\mathbb{M}_{n,N}$, with n_{MC} additional realizations $[x_{\text{ar}}^1], \ldots, [x_{\text{ar}}^{n_{\text{MC}}}]$ in $\mathbb{M}_{n,N}$, which are computed without calling the stochastic computational model but that are computed by using the generator of samples that is summarized in Section 2.4. By using these $\nu_{\text{sim}} = n_{\text{MC}} \times N$ additional samples, a good estimate of $h(q, \mathbf{w}_g)$ can be obtained by using Eqs. (27) to (29). The algorithm is the following.

Data allowing the algorithm to be initialized.

1. Given m_w, N, and n_{MC}.
2. Set $n = m_w + 1$.
3. Given matrix $[x_d]$ in $\mathbb{M}_{n,N}$ for which the columns are the N data points $\mathbf{x}^1 \ldots \mathbf{x}^N$ and where $\mathbf{x}^\ell = (\mathbf{w}^\ell, q^\ell) \in \mathbb{R}^n$ with $\mathbf{w}^\ell \in \mathbb{R}^{m_w}$ and $q^\ell \in \mathbb{R}$.
4. If necessary, scale the data set represented by matrix $[x_d]$ in $\mathbb{M}_{n,N}$.
5. Introduce the random matrix $[\mathbf{X}]$ whose columns are N independent copies of random vector $\mathbf{X}$ and for which one given sample of random matrix $[\mathbf{X}]$ is matrix $[x_d]$.
6. Given q in $\mathbb{R}$ and $\mathbf{w}_g$ in $\mathcal{C}_{\mathbf{w}_g}$.

Steps of the algorithm.

1. Generate n_{MC} additional samples $[x_{\text{ar}}^1], \ldots, [x_{\text{ar}}^{n_{\text{MC}}}]$ of random matrix $[\mathbf{X}]$ by using the algorithm detailed in [23].
2. From the $\nu_{\text{sim}} = n_{\text{MC}} \times N$ columns of the n_{MC} matrices $[x_{\text{ar}}^1], \ldots, [x_{\text{ar}}^{n_{\text{MC}}}]$, extract ν_{sim} samples $(\mathbf{w}_{\text{ar}}^\ell, q_{\text{ar}}^\ell) \in \mathbb{R}^{m_w} \times \mathbb{R}$ for $\ell = 1, \ldots, \nu_{\text{sim}}$.
3. Compute $\underline{q}_{\text{ar}} = \frac{1}{\nu_{\text{sim}}}\sum_{\ell=1}^{\nu_{\text{sim}}} q_{\text{ar}}^\ell$ and $\sigma_{q_{\text{ar}}}^2 = \frac{1}{\nu_{\text{sim}}}\sum_{\ell=1}^{\nu_{\text{sim}}} (q_{\text{ar}}^\ell - \underline{q}_{\text{ar}})^2$.
4. For $j = 1, \ldots, m_w$, compute $\underline{w}_{\text{ar},j} = \frac{1}{\nu_{\text{sim}}}\sum_{\ell=1}^{\nu_{\text{sim}}} w_{\text{ar},j}^\ell$ and $\sigma_{\text{ar},j}^2 = \frac{1}{\nu_{\text{sim}}}\sum_{\ell=1}^{\nu_{\text{sim}}} (w_{\text{ar},j}^\ell - \underline{w}_{\text{ar},j})^2$.

5. For $\ell = 1, \ldots, \nu_{\text{sim}}$, compute $(\widehat{\mathbf{w}}_{\text{ar}}^{\ell}, \widehat{q}_{\text{ar}}^{\ell}) \in \mathbb{R}^{m_w} \times \mathbb{R}$ by using Eq. (19) in which N is replaced by ν_{sim}.
6. Compute $\widehat{\mathbf{w}}_g$ and $\widehat{q}$ by using Eq. (20).
7. Compute s by using Eq. (25) in which N is replaced by ν_{sim}.
8. For $\ell = 1, \ldots, \nu_{\text{sim}}$,
   (a) compute $\widehat{v}^{\ell}(\widehat{\mathbf{w}}_g)$ by using Eq. (28) in which N is replaced by ν_{sim}.
   (b) compute $\widehat{r}^{\ell}(\widehat{q})$ by using Eq. (29) in which N is replaced by ν_{sim}.
9. Compute $h(\widehat{q}, \widehat{\mathbf{w}}_g)$ by using Eq. (27) in which N is replaced by ν_{sim}.

End of the algorithm.

Acknowledgments

This research was supported by the US department of energy under the Scidac Institute for Uncertainty Quantification under Extremes (Quest).

References

1. Aanonsen, S., Eide, A., Holden, L., Aasen, J.: Optimizing reservoir performance under uncertainty with application to well location. In: In SPE annual technical conference and exhibition. Society of Petroleum Engineers, pp. 67–76 (1995)
2. Artus, V., Durlofsky, L., Onwunala, J., Aziz, K.: Optimization of nonconventional wells under uncertainty using statistical proxies. Computational Geosciences **10**, 389–404 (2006)
3. Babaei, M., Alkhatib, A., Pan, I.: Robust optimization of subsurface flow using polynomial chaos and response surface surrogates. Computational Geosciences **19**, 979–998 (2015)
4. Bangerth, W., Klie, H., Wheeler, M., Stoffa, P., Sen, M.: On optimization algorithms for the reservoir oil well placement problem. Computational Geosciences **10**, 303–319 (2006)
5. Beckner, B., Song, X.: Field development planning using simulated annealing-optimal economic well scheduling and placement. In: In SPE annual technical conference and exhibition. Society of Petroleum Engineers (1995)
6. Bellout, M., Echeverria-Ciaurri, D., Durlofsky, L., Foss, B., Kleppe, J.: Joint optimization of oil well placement and controls. Computational Geosciences **16**, 1061–1079 (2012)
7. Bowman, A., Azzalini, A.: Applied Smoothing Techniques for Data Analysis. Oxford University Press, Oxford, UK (1997)
8. Christiansen, L., Capolei, A., Jrgensen, J.: Time-explicit methods for joint economical and geological risk mitigation in production optimization. Journal of Petroleum Science and Engineering **146**, 158–169 (2016)
9. Christie, M., Blunt, M.: Tenth spe comparative solution project: a comparison of upscaling techniques. SPE Reservoir Evaluation & Engineering **4** (2001)
10. Coifman, R., Lafon, S.: Diffusion maps, applied and computational harmonic analysis. Applied and Computational Harmonic Analysis **21**(1), 5–30 (2006)
11. Coifman, R., Lafon, S., Lee, A., Maggioni, M., Nadler, B., Warner, F., Zucker, S.: Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. PNAS **102**(21), 7426–7431 (2005)
12. ECLIPSE: Reference Manual. Schlumberger, Houston, Texas (2009)
13. van Essen, G., Zandvilet, M., den Hof, P.V., Bosgra, O., Jansen, J.: Ribust waterflooding optimization of multiple geological scenarios. SPE Journal **14**(1), 202–210 (2009)
14. Ghanem, R.: Scales of fluctuation and the propagation of uncertainty in random porous media. Water Resources Research **34**(9), 2123–2136 (1998)
15. Ghanem, R., Soize, C.: Probabilistic non-convex constrained optimization with fixed number of function evaluations. International Journal for Numerical Methods in Engineering **to appear** (2017)
16. Guyaguler, B., Horne, R.: Uncertainty assessment of well placement optimization. SPE Reservoir Evaluation & Engineering **7**(1), 23–32 (2004)
17. Jesmani, M., Bellout, M., Hanea, R., Foss, B.: Well placement optimization subject to realistic field development constraints. Computational Geosciences **20**, 1185–1209 (2016)
18. Li, L., Jafarpour, B., Mohammad-Khaninezhad, M.: A simultaneous perturbation stochastic approximation algorithm for coupled well placement and control optimization under geologic uncertainty. Computational Geosciences **17**, 167–188 (2013)
19. Rashid, K., Bailey, W., Couet, B., Wilkinson, D.: An efficient procedure for expensive reservoir-simulation optimization under uncertainty. SPE Economics & Management **5**(4), 21–33 (2013)
20. Rian, D., Hage, A.: Automatic optimization of well locations in a north sea fractured chalk reservoir using a front tracking reservoir simulator. In: In International Petroleum Conference and Exhibition of Mexico. Society of Petroleum Engineers (1994)
21. Scott, D.: Multivariate Density Estimation: Theory, Practice, and Visualization, 2nd edn. John Wiley and Sons, New York (2015)

22. Soize, C.: Polynomial chaos expansion of a multimodal random vector. SIAM/ASA Journal on Uncertainty Quantification **3**(1), 34–60 (2015). DOI 10.1137/140968495
23. Soize, C., Ghanem, R.: Data-driven probability concentration and sampling on manifold. Journal of Computational Physics **321**, 242–258 (2016). DOI 10.1016/j.jcp.2016.05.044
24. Spall, J.: Introduction to Stochastic Searh and Optimization. Wiley-Interscience (2003)
25. Sun, W., Durlofsky, L.: A new data-space inversion procedure for efficient uncertainty quantification in subsurface flow problems. Mathematical Geosciences **49**, 679–715 (2017)
26. Thimmisetty, C., Tsilifis, P., Ghanem, R.: Paper petroleum. Artificial Intelligence for Engineering Design, Analysis and Manufacturing **31**(3), 265–276 (2017). DOI Homogeneous chaos basis adaptation for design optimization under uncertainty: Application to the oil well placement problem
27. Wang, H., Echeverria-Ciaurri, D., Durlofsky, L., Cominelli, A.: Optimal well placement under uncertainty using a retrospective optimization framework. SPE Journal **17**(1), 112–121 (2012)
28. Yeten, B., Durlofsky, L., Aziz, K.: Optimization of nonconventional well type, location and trajectory. In: In SPE annual technical conference and exhibition. Society of Petroleum Engineers (2002)
29. Zhang, Y., Lu, R., Forouzanfar, F., Reynolds, A.: Well placement and control optimization for wag/sag processes using ensemble-based method. Computers and Chemical Engineering **101**, 193–209 (2017)