Gradient-based constrained well placement optimization

Oleg Volkov*

Department of Energy Resources Engineering, Stanford University

Mathias C. Bellout†

Department of Geoscience and Petroleum, NTNU

Published: 16 August 2018, Journal of Petroleum Science and Engineering vol. 171 pp.1052–1066
DOI: 10.1016/j.petrol.2018.08.033

Abstract

A novel well placement gradient approximation methodology is developed based on performing finite difference approximations of augmented Lagrangian derivatives within the adjoint formulation. The methodology is efficient because it requires only a pair of (forward and backward) simulations to yield a cost function sensitivity with respect to well placement coordinates. The approximated derivative is used within a Sequential Quadratic Programming (SQP) solver ensuring fast convergence and efficient constraint-handling. An extensive error analysis is performed to identify the gradient approximation errors associated with different perturbation ranges. This analysis provides information regarding the appropriate perturbation step size range needed to maintain a consistent gradient approximation while reducing errors associated with the simulation and the discretized nature of the reservoir. We validate the efficiency of the approach by solving for optimal well placement and comparing the results against two major gradient-based well placement approaches from the literature. For these cases, the methodology developed in this work delivers higher or similar final objective values while providing better performance in terms of fewer cost function evaluations. Finally, the methodology is used to find the optimal configuration of multiple deviated producers both in a binary channelized case and in a case based on the Brugge reservoir. These applications show that the proposed methodology can handle cases with more complex grid and production scenarios that require derivative information for the location of deviated wellbores in continuous space.

1 Introduction

In this work, we present a novel and efficient way to approximate derivative information with respect to well location. The proposed approach relies on an augmented Lagrangian approximation of the cost function provided by an adjoint formulation. Within the adjoint framework, the approach evaluates the gradient of the objective using finite differences. This yields an inexpensive gradient assembly procedure that produces a well placement gradient after only two simulations, i.e., one forward- and one backward-in-time. Specifically, this approach extends the application of the adjoint gradient framework to well location variables, which typically lack a continuous representation within the reservoir model. Crucially,

*e-mail: ovolkov@stanford.edu
†e-mail: mathias.bellout@ntnu.no

the ability to estimate the sensitivity with respect to wellbore location in an efficient manner enables the use of gradient-based routines for optimization of complex wellbore designs and configurations.

Despite being local, gradient-based techniques are efficient, rely on well-established convergence theory, and can be implemented in a straightforward manner by making use of a variety of off-the-shelf solvers. These techniques are therefore attractive to use for the well placement problem. In this work, we couple the approximated gradient with a Sequential Quadratic Programming (SQP) method.

The gradients obtained by the proposed methodology depend on the accuracy of the finite difference approximation of the adjoint partial derivative terms. Clearly, to perform an efficient well placement local search, it is important that these gradients are consistent and sufficiently accurate. For this reason, we perform an extensive analysis to investigate the influence of different sources of error on the accuracy of the approximation.

This paper is structured as follows: Section 2, *Background*, describes the general problem of deriving well placement gradients and classifies the different gradient approximation approaches found in the literature based on their primary mode of approximation. Section 3, *Methodology*, explains in details, including pseudocode, how the adjoint solution approach is enhanced to approximate the well placement gradient. Furthermore, this section identifies and performs an extensive study of the different error sources associated with this type of gradient approximation. Following this, we select an appropriate perturbation size range based on maintaining the highest possible consistency in the finite difference expression. Section 4, *Application*, tests the methodology against two cases from the literature. These first two cases serve to validate and compare the performance of the proposed approach against existing methodologies. The methodology is then applied to two more advanced test cases: a case involving a binary channelized system and a case derived from the Brugge reservoir. These last two cases demonstrate the versatility of the approach in terms of dealing with horizontal or deviated wellbores within channelized geologies and in realistic grids, respectively. In Section 5, *Summary*, we summarize the methodology and results presented in this work and provide ideas for further studies.

2 Background

Various procedures relying on gradient-based or derivative-free algorithms have been proposed for the well placement problem. Since this work targets well placement gradient approximation, the following review addresses gradient-based procedures only. See, e.g., Forouzanfar and Reynolds [6] for references to multiple procedures concerning derivative-free algorithms.

Gradient-based procedures are difficult to implement for the well placement problem due to the general lack of derivatives. Reservoir fluid flow is typically modeled using mass balance equations commonly discretized by finite volumes. In this computational model, a well is defined as a collection of source/sink terms that correspond to those reservoir blocks where the well is completed. Well placement derivatives are difficult to derive analytically from this formulation since well position coordinates enter implicitly as discrete variables in the governing equations.

This section provides an overview of multiple approaches proposed in the literature to derive well placement derivatives. Different well placement gradient-approximation approaches are categorized based on whether authors (1) reinstate a continuous well model, (2) perform a direct estimation of the gradient by approximating the gradient through a difference scheme, or (3) perform an indirect estimation using adjoint-based gradients corresponding to (necessarily continuous) variables associated with production, i.e., well controls. These categories are broad and non-exclusive, in the sense that a particular approach may have attributes that extend beyond its designated category; however, the various approaches are categorized according to what is seen as their main mode of operation. The classification thus helps to organize and explain, in general terms, the different methodologies proposed in the literature. These outlined categories serve

as background for the adjoint-based well placement gradient approximation method presented in this work, and are further described below.

2.1 Gradient derivation through continuous well model approximation

In [23], a cost function gradient with respect to well coordinates is obtained by formulating a continuous approximation of the originally discrete well placement problem. The approximation is achieved by first defining the well source/sink terms in the mass balance partial differential equations as depending explicitly on continuous well location variables. These modified (still discontinuous) well source/sink terms are then approximated by a Gaussian type function (other functions are possible) that is a continuous expression with respect to the defined well coordinates. This results in a set of PDEs with continuous source/sink terms dependent on continuous well location coordinates. The discretization of this set of equations yields (additional) nonzero well terms beyond the original well blocks. The extent and total number of these additional terms are determined by the size of the base of the approximating Gaussian function. In practice, the set of additional well terms is truncated to constitute a neighborhood of pseudo-well blocks surrounding the original well block. Notice that the use of pseudo-wells created during the later stages of this method is an attribute shared with an indirect type of approach (discussed below). Regardless of that, the primary goal of this method is to enable the use of the adjoint framework for gradient computation by introducing a continuous representation of the discrete problem. This method, therefore, falls within the *continuous approximation* category.

Finally, the cost function within this approach is modified to include the additional well terms, and the system is solved within an adjoint formulation to yield a well placement gradient. Since all well terms are now established as continuous functions dependent on continuous well location variables, the adjoint gradient with respect to these variables can be solved for in a straightforward manner.

The approach proposed by Sarma and Chen [23] was developed for a two-dimensional (x, y) coordinate well placement problem on a regular grid. It is not obvious how this gradient approximation approach can be extended to the deviated (multi-block) well placement case. For such a case, one would need to specify at least two (x, y, z)-points (heel and toe) to define a horizontal well trajectory. For each well block traversed by this trajectory, this approach would specify a particular distribution of pseudo-wells, which would, at the very least, yield overlapping issues between pseudo-well distributions that would have to be resolved.

A related approach, in terms of establishing a distribution of completions surrounding a well trajectory, is described in [6]. Analogously to the latter part of the procedure given in Sarma and Chen, in this approach, the well rate associated with the original well trajectory is allocated among the surrounding perforations according to a logic that allocates a higher portion of the rate to perforations close to the original trajectory. In Forouzanfar and Reynolds, however, the main objective behind this distribution is not to approximate gradients (neither through adjoints or in any other way), but rather to smoothen the cost function to enable the applicability and further improve the efficiency of a model-based derivative-free algorithm. Taken together, both these approaches yield a considerable set of additional source/sink terms in the governing system of equations (even after removing completions with very low connection factors), thus increasing overall complexity and possibly making the simulation less stable and harder to converge.

2.2 Direct well placement gradient approximation

In this work, we categorize as a *direct approach* a type of methodology that relies fundamentally on a certain perturbation of the well coordinate vector to derive sensitivity information. For the most part, approaches within this category rely on a stochastic perturbation of the parameter vector, i.e., they perturb the vector only along a random direction, thus avoiding the high computational cost associated with proper finite difference schemes that require performing a function call for each vector component. More

sophisticated implementations, e.g., EnOpt, extend this technique to perform stochastic perturbation of the optimization vector over an ensemble of geological realizations, thus yielding an approximate gradient averaged over the uncertainty set [see 15].

Within the direct approach category, Bangerth et al. [1] were the first to apply an integer variant of the Simultaneous Perturbation Stochastic Approximation (SPSA) method to well placement optimization, while in [11] and later [16], SPSA was used not only to derive well placement sensitivities but also to jointly optimize for well placement and control. Finally, Jesmani et al. [12] optimized well trajectories (vertical, horizontal, and deviated) using a continuous variant of the SPSA algorithm.

The stochastic nature and underlying sampling property of SPSA offers certain robustness against moderate cost function noise. In [17] and [13], this property is used to handle the increased noise in the expected objective caused by using only a random subset of realizations during well placement optimization under uncertainty.

The SPSA methodology effectively reduces the computational cost associated with approximating the gradient. However, the gain in computational cost might be offset by the stochastic procedure producing a sufficiently poor gradient estimate that in turn reduces the overall performance of the optimization routine. Issues regarding how to retain computational gain and assure reasonable conditional convergence when using this gradient are linked to proper selection of perturbation vector and to tuning parameters of the algorithm. These issues are further discussed by Li et al. [17] and Jesmani et al. [13].

2.3 Indirect gradient approximation

Finally, an *indirect approach* refers to the third type of approach that derives well placement sensitivity by using associated (most commonly adjoint) well control gradients. One main branch of this type of approach considers a well surrounded by an array of pseudo-wells, and then uses the well control gradient associated with each of the surrounding wells to determine a direction for well re-location that improves the objective. Commonly, this direction is determined using the magnitude of the pseudo-well control gradients and their relative position with respect to the original wellbore location. Another main branch within this category uses associated control gradients for well placement optimization while also (implicitly) being able to optimize on the number of wells. This type of approach performs control optimization starting with an artificially large number of wells (subject to a total rate constraint) and improves iteratively on the objective by removing low-performing wells. These two main branches and similar approaches within the *indirect approach* category are further discussed below.

The adjoint-based technique of placing an array of low-producing pseudo-wells surrounding a vertical wellbore was first developed by Zandvliet et al. [29]. Zandvliet et al. [29] placed pseudo-wells at each of the blocks neighboring the well whose vertical, discrete, (i, j)-position is to be optimized. Importantly, these pseudo-wells are set to produce at a very low rate compared to the rate of the main wellbore, so as not to significantly disturb the overall reservoir flow pattern. The existence of the pseudo-wells introduces additional source/sink terms within the reservoir simulation equations, which increases the complexity of the adjoint-gradient computation for the controls of these wells. The control rate derivatives for each of the pseudo-wells are summed, and the original well is moved to the location of the pseudo-well with the largest derivative. An iterative procedure for well location optimization is then applied, but only for a two-dimensional well configuration in a relative simple reservoir. The method was later extended by Castineira and Alpak [4] and Vlemmix et al. [26]. The former introduced a double ring of pseudo-wells to speed convergence and extend the search capacity to a two-grid-block neighborhood. The latter applied the method to optimize the trajectory of a deviated well subject to a curvature constraint within three-dimensional reservoir space. In that work, a number of trajectory points defining the well path are each surrounded by pseudo-sidetracks which extend to all the grid blocks adjacent each of these points. The sidetracks have a very little perforation so that their production accounts only for an insignificant fraction of the production through the main wellbore. An at-

tractor point is constructed for each trajectory point by assembling the sensitivities of the surrounding side tracks. Then, analogous to the vertical well application, each attractor point is used to reposition its corresponding trajectory point in a direction of higher NPV. Since this implementation modifies only the positioning of the trajectory points, it is limited to optimizing only the shape and not the overall location of the wellbore.

Another type of indirect method solving an alternate continuous control problem is described in [28] and Zhang et al. [30]. In those works, the problem is initialized as a well control problem with a well in every grid block, and the optimization subject to a total rate distributed over the number of wells existing at any point of the procedure. The procedure optimizes production and decreases drilling costs by shutting-in, and thus eliminating, wells (ultimately determining their number and position). Forouzanfar and Reynolds [7] proposed a two-stage variant of this indirect method that introduces an initialization step to determine an appropriate total rate for the given production time frame, thus decreasing the dependence of the solution on the a priori specification of the total rate constraint. A general benefit of this type of indirect method is that, in addition to well rates, well number and location are implicitly subject to optimization. The implementation of this type of indirect method for three-dimensional deviated trajectories, however, would be limited by the capacity of the simulator to handle a potentially large increase in the number of pseudo-well equations in the governing system.

2.4 Adjoint-based gradient approximation

Within this classifications scheme, the gradient approximation method presented in this work can be referred to as a direct approach. However, the main distinction of this methodology is that the direct approximation effort, based on finite differences, operates on key partial derivative terms within the adjoint solution procedure itself, and not at the cost function level. The main advantage then is that our method inherits the efficiency of the adjoint framework to compute the desired sensitivity. The next section explains the methodology in detail and provides extensive error analysis.

3 Methodology

This section presents the general well placement optimization problem, and main features of gradient-based iterative solution procedures. It then briefly describes the adjoint framework, and explains in detail the algorithm for developing the well placement gradients. Three sources of error affecting the gradient-approximation procedure are identified, and an in-depth analysis of their relative influence is performed.

3.1 Optimization problem and solution procedure

The general well placement optimization problem is defined as

$$\widehat{\mathbf{u}} = \underset{\mathbf{u}}{\operatorname{argmax}}\ \mathcal{J}(\mathbf{x}, \mathbf{u}), \tag{1a}$$

$$\text{s.t. } g(\mathbf{x}, \mathbf{u}) = 0, \tag{1b}$$

$$\mathbf{u} \in D. \tag{1c}$$

$\mathcal{J}$ represents the problem cost function which quantifies the objective for the optimization. $\mathcal{J}$ when representing the Net Present Value (NPV) is usually expressed in the general form

$$\mathcal{J}(\mathbf{x}, \mathbf{u}) = \sum_{i=1}^{N_w} C_l \operatorname{length}_i(\mathbf{u}) + \sum_{n=1}^{N} \left(\sum_{i=1}^{N_w} \sum_{p=1}^{N_p} C_{p,i}\, q_{p,i}(\mathbf{x}, \mathbf{u}) \right) \Bigg|_{t=t_n} \Delta t_n, \tag{2}$$

where N_w is the number of wells, N_p is the number of phases, t_n is the time discretized into N steps required by the simulation. $q_{p,i}$ and $C_{p,i}$ denote, respectively, the production/injection phase flow rate and the discounted price/cost of the phase p, in the i-th well. Finally, $C_l \operatorname{length}_i(\mathbf{u})$ defines the cost of extending the i-th well to its current length.

$\mathcal{J}$ depends on both state variables $\mathbf{x}$ and well control and configuration parameters $\mathbf{u}$. $\mathbf{x}$ is composed of pressure and saturation states at each of the reservoir grid blocks and at each time step, while $\mathbf{u}$ consist of time-dependent well controls, such as bottomhole pressure and liquid rates, in addition to variables that determine the general configuration of the wellbore within the reservoir grid. In this work, only wells consisting of a single segment are considered. The single-segment wellbore is parametrized in three-dimensional space by six values, i.e., the coordinates of the heel and toe for the i-th wellbore $[\, x_{\text{heel}} \;\; y_{\text{heel}} \;\; z_{\text{heel}} \;\; x_{\text{toe}} \;\; y_{\text{toe}} \;\; z_{\text{toe}} \,]_i \in \mathbb{R}^6$. In the case of vertical wells, this parametrization is modified by constraining some of the degrees of freedom, i.e., we set $x_{\text{toe}} = x_{\text{heel}}$, $y_{\text{toe}} = y_{\text{heel}}$, while the z coordinates of the heel and toe are kept constant. In general, one completion is defined at each grid cell traversed by the well trajectory. Thus, for vertical wells, completions are defined in those grid cells intersected by the line between z_{heel} and z_{toe}. For deviated wellbores, completions are defined only in those grid cells traversed by the specified well segment while any preceding section of the well is assumed non-producing. Future work will likely extend this overall parametrization, e.g., use wells consisting of multiple segments, to approximate more complex trajectories.

(1b) defines the mass balance equations discretized in space using finite volume formulation and in time using an implicit Euler scheme. In this formulation, the equations are thought of as a set of constraints that determine the state given by the controls [22], i.e. $\mathbf{x} = \mathbf{x}(\mathbf{u})$. The system of equations is solved numerically with respect to $\mathbf{x}$. This computation is performed by the reservoir simulator.

Time-dependent well controls enter explicitly into the reservoir simulation equations g through the well model defined at each source/sink reservoir grid block. On the other hand, variables that determine the configuration of the well, e.g., well placement coordinates, do not typically have an explicit representation within the reservoir simulation solution scheme. Rather, these variables exert their influence implicitly by determining the number and, through the well connection factor, the magnitude of the reservoir source/sink terms.

Well variables $\mathbf{u}$ are constrained by simple upper and lower bounds. However, these variables can, and are usually, also constrained by more sophisticated types of restrictions. Restrictions on $\mathbf{u}$ range from various input and output nonlinear production constraints on the controls, to various forms of well placement constraints that limit the overall configuration of the wellbores, e.g., in terms of length, orientation, curvature and/or inter-well distance [14]. The various types of constraints that may apply to $\mathbf{u}$ form a feasible region designated by D.

(1) is commonly solved for $\hat{\mathbf{u}}$ by using iterative procedures. A nonlinear programming solution approach is to sequentially approximate the nonlinear $\mathcal{J}$ using linear or quadratic functions or using other functions whose optima are readily available. The local approximations are used to find current best points, and the iterative procedure yields a sequence of optima ultimately leading to the optimum of $\mathcal{J}$. Both linear and quadratic approximations rely on the first order sensitivities of $\mathcal{J}$ with respect to $\mathbf{u}$.

3.2 Adjoint framework

Considering $\mathcal{J}$ continuously differentiable at point $\mathbf{u}$, the sensitivities of $\mathcal{J}$ can be computed exactly following the mathematical definition of a Gâteaux differential. When $\mathbf{x}(\mathbf{u})$ is an explicit function, the Gâteaux differential of $\mathcal{J}$ is defined as

$$d\mathcal{J}(\mathbf{x}(\mathbf{u}), \mathbf{u}; \delta\mathbf{u}) = \lim_{\varepsilon \to 0} \frac{\mathcal{J}(\mathbf{x}(\mathbf{u} + \varepsilon\delta\mathbf{u}), \mathbf{u} + \varepsilon\delta\mathbf{u}) - \mathcal{J}(\mathbf{x}(\mathbf{u}), \mathbf{u})}{\varepsilon}. \quad (3)$$

The following gradient definition relies on the Riesz Representation theorem given below (here $\langle \cdot, \cdot \rangle$ designates the inner product).

Theorem (Berger 1977, p.30) *Let X be a Hilbert space. Then any bounded linear functional $h(x)$ defined on X can be uniquely written as $h(x) = \langle x, y \rangle$ for some $y \in X$.*

According to this theorem, if $d\mathcal{J}(\mathbf{x}(\mathbf{u}), \mathbf{u}; \delta\mathbf{u})$ is a bounded linear functional of $\delta\mathbf{u}$, then there exist a unique vector, $\nabla_{\mathbf{u}}\mathcal{J}$, called the gradient, such that for every $\delta\mathbf{u}$ we have

$$d\mathcal{J}(\mathbf{x}(\mathbf{u}), \mathbf{u}; \delta\mathbf{u}) = \left\langle \nabla_{\mathbf{u}}\mathcal{J}, \delta\mathbf{u} \right\rangle. \quad (4)$$

In contrast to $d\mathcal{J}(\mathbf{x}(\mathbf{u}), \mathbf{u}; \delta\mathbf{u})$, the term $\nabla_{\mathbf{u}}\mathcal{J}$ is independent of the perturbation direction $\delta\mathbf{u}$, and therefore the sought sensitivities of $\mathcal{J}$ are expressed entirely by this term. When the explicit form of $\mathbf{x}(\mathbf{u})$ is unknown, an efficient way to derive $\nabla_{\mathbf{u}}\mathcal{J}$ in (4) is to reformulate the problem into an augmented form, and apply the first order optimality conditions to this formulation.

In the following, we refer to the adjoint gradient derivation outlined in [21]. According to Plessix, the augmented Lagrangian is an extension of the original cost function definition that includes the equality constraint functions $g(\mathbf{x}, \mathbf{u})$ as additional terms. Note that in this derivation, $\mathbf{x}$ is kept independent of $\mathbf{u}$. Furthermore, each term added to the original cost function is weighted by a Lagrange multiplier. Such multipliers form a vector denoted here by $\boldsymbol{\lambda}$, which, as a part of the Lagrangian, is also kept independent of $\mathbf{u}$. The augmented Lagrangian formulation for the problem stated above is thus given as

$$\mathcal{L}(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}) = \mathcal{J}(\mathbf{x}, \mathbf{u}) + \big\langle \boldsymbol{\lambda}, g(\mathbf{x}, \mathbf{u}) \big\rangle, \tag{5}$$

with $\mathbf{x}$ and $\boldsymbol{\lambda}$ both being kept independent of $\mathbf{u}$ by definition.

Notice that $\mathcal{L}$ is a continuously differentiable function of $\boldsymbol{\lambda}$ and $\mathbf{x}$. Applying the first order optimality conditions with respect to those variables yields the following expressions:

$$\frac{\partial \mathcal{L}(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} = 0 \rightarrow g(\mathbf{x}, \mathbf{u}) = 0; \tag{6a}$$

$$\frac{\partial \mathcal{L}(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda})}{\partial \mathbf{x}} = 0 \rightarrow \left(\frac{\partial g(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}}\right)^T \boldsymbol{\lambda} = -\frac{\partial \mathcal{J}(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}}. \tag{6b}$$

(6a) is satisfied as $\mathbf{x}$ is determined to be the solution of (1b) by numerical simulation. (6b) is solved as a system of equations with $\boldsymbol{\lambda}$ being the unknown (in this computation, (6b) is referred to as the *adjoint variable system* while $\boldsymbol{\lambda}$ is called the *adjoint variable*).

Given a solution to (6), here denoted as $(\widetilde{\mathbf{x}}, \widetilde{\boldsymbol{\lambda}})$, [21] formulates the expression for the *adjoint-based gradient* of $\mathcal{J}$ as

$$\nabla_{\mathbf{u}}\mathcal{J} = \frac{\partial \mathcal{L}(\widetilde{\mathbf{x}}, \mathbf{u}, \widetilde{\boldsymbol{\lambda}})}{\partial \mathbf{u}}, \tag{7}$$

where $\widetilde{\mathbf{x}}$ and $\widetilde{\boldsymbol{\lambda}}$, are kept independent of $\mathbf{u}$ when the derivative of $\mathcal{L}$ with respect to $\mathbf{u}$ is computed. From (5) and (7) we obtain

$$\nabla_{\mathbf{u}}\mathcal{J} = \frac{\partial \mathcal{J}(\widetilde{\mathbf{x}}, \mathbf{u})}{\partial \mathbf{u}} + \left(\frac{\partial g(\widetilde{\mathbf{x}}, \mathbf{u})}{\partial \mathbf{u}}\right)^T \widetilde{\boldsymbol{\lambda}}, \tag{8}$$

where, as in (7), the partial derivatives on the right-hand side take into account only the explicit dependence of $\mathcal{J}$ and g with respect to $\mathbf{u}$.

An important conclusion based on (8) is that the gradient $\nabla_{\mathbf{u}}\mathcal{J}$ may be approximated by replacing the operator $\partial/\partial\mathbf{u}$ in (8) by a finite difference operation. This conclusion is central in the proposed derivation of well placement gradients described next.

3.3 Adjoint-based well placement gradient approximation procedure

In the previous section, the sensitivities of $\mathcal{J}$ with respect to $\mathbf{u}$ are computed under the assumption that $\mathcal{L}$ and $\mathcal{J}$ are continuously differentiable functions of $\mathbf{u}$. However, time/space-discretizations are often applied to fundamental parts of the overall problem, e.g., to solve $g(\mathbf{x}, \mathbf{u}) = 0$, and for well control variables in $\mathbf{u}$. Therefore, continuous differentiability of $\mathcal{L}$ and $\mathcal{J}$ cannot automatically be asserted for all cases. In the following, we inspect whether this assumption holds for the given discretizations of time and space. For the case of time discretization, continuous differentiability, both with respect to well configuration variables and with respect to piecewise linear constant-in-time well controls, can be confirmed. The former is valid because well configuration parameters, e.g., well placement coordinates, are independent of time, while the latter is valid because well controls are shown to be consistent with the standard adjoint formulation described in Section 3.2; for further explanation, see [27]. However, for the case of space discretization, the assumption of differentiability still holds for well control type of variables, but is no longer valid with respect to well configuration variables.

Following this, (7) can only be applied in a straightforward manner to well control variables. Crucially, however, (7) may still be applied to well configuration

variables if the partial derivatives in (8) are approximated, e.g., by finite differences, as indicated at the end of Section 3.2. Following this reasoning, we first consider a particular point $\mathbf{u}$ and obtain $\widetilde{\mathbf{x}}$ and $\widetilde{\boldsymbol{\lambda}}$ as the solution of (6) to be used in the approximation of (7). Because the derivatives with respect to $\mathbf{u}$ in (8) are partial derivatives, we take into account only the explicit variation of $\mathcal{J}(\widetilde{\mathbf{x}}, \mathbf{u})$ and $g(\widetilde{\mathbf{x}}, \mathbf{u})$ as functions of $\mathbf{u}$ with respect to some small perturbations $\mathbf{u} \pm \varepsilon\delta\mathbf{u}$. Here ε denotes a perturbation size and $\delta\mathbf{u}$ is a perturbation direction. Next, we apply a central difference scheme to evaluate the terms of (8) for the direction $\delta\mathbf{u}$. The approximation of the adjoint gradient along this direction thus becomes

$$\begin{aligned}\widetilde{\nabla}_{\delta\mathbf{u}}\mathcal{J} =& \frac{\mathcal{L}(\widetilde{\mathbf{x}}, \mathbf{u} + \varepsilon\delta\mathbf{u}) - \mathcal{L}(\widetilde{\mathbf{x}}, \mathbf{u} - \varepsilon\delta\mathbf{u})}{2\,\varepsilon} = \\ & \frac{\mathcal{J}(\widetilde{\mathbf{x}}, \mathbf{u} + \varepsilon\delta\mathbf{u}) - \mathcal{J}(\widetilde{\mathbf{x}}, \mathbf{u} - \varepsilon\delta\mathbf{u})}{2\,\varepsilon} + \\ & \left(\frac{g(\widetilde{\mathbf{x}}, \mathbf{u} + \varepsilon\delta\mathbf{u}) - g(\widetilde{\mathbf{x}}, \mathbf{u} - \varepsilon\delta\mathbf{u})}{2\,\varepsilon}\right)^T \widetilde{\boldsymbol{\lambda}}.\end{aligned} \tag{9}$$

To approximate the full gradient, (9) is repeated using perturbations $\delta\mathbf{u}$ for all $6N_w$ well location components of $\mathbf{u}$, i.e., $\delta\mathbf{u}$ is defined as a unit vector $\mathbf{e}_j$ in the jth component of $\mathbf{u}$. Finally, we stress that, whenever applying (9), both $\widetilde{\mathbf{x}}$ and $\widetilde{\boldsymbol{\lambda}}$ at all times conform with the conditions stated for (8). Therefore, as in (8), the approximated adjoint gradient requires only one solution of (6). Approximation (9) is used in the context of an adjoint framework summarized in Algorithm 1. The steps of this algorithm are described in detail below.

All components of the gradient vector $\nabla_{\mathbf{u}}\mathcal{J}$ are assembled with the use of two runs only. In the first (forward-in-time) run, we perform a reservoir simulation and save the state variables into a binary backup (lines 2 through 5 in Algorithm 1). Recall N is the number of steps required by the simulation. In the second (backward-in-time) run, the state variables are read from the backup for each time step starting from the last time step. The state data is used to assemble the adjoint system (lines 7 through 30). At each step of the backward-in-time procedure, the adjoint system is solved with respect to the corresponding adjoint variable $\widetilde{\boldsymbol{\lambda}}$ (lines 9 and 10). At this point, the vector of optimization variables $\mathbf{u}$ is separated into two groups:

1. well control variables denoted v_j;
2. well configuration variables denoted w_j.

The gradient components corresponding to v_j are evaluated directly according to (7), (lines 11 through 13). The gradient components corresponding to the well configuration variables are obtained by perturbing the variables w_j individually by $\pm\varepsilon$ (lines 15 through 21). The residual and cost function terms corresponding to $w_j \pm \varepsilon$ are computed and used in (9) to approximate the gradient component at that time step (lines 22 through 27).

Perturbing slightly the variables defining a well configuration will yield blocks with slightly different well connection factors.

Recall that well configuration variables w_j have a set of reservoir grid blocks associated with it. Depending both on the perturbation size and on the geometry of the well traversing the grid, the perturbed variables may also yield a slight change in the cardinality of the set of well blocks. Lines 15 through 21 in Algorithm 1 compute the perturbed well trajectories, i.e., this part of the algorithm finds the (altered) sets of traversed grid blocks and calculates the appropriate connection factors for these well blocks (line 19 in Algorithm 1).

For perturbation sizes significantly smaller than the mean grid block size, applying the standard Peaceman model [19] does not provide an adequate measure of production change due to the underlying model assumption that the wellbore runs through the axis of the block. Since production changes from this type of perturbations are linked to wellbore length and/or geometry changes to the well section traversing the grid block, we introduce a flexible (approximate) calculation of the well connection factor. This calculation, presented by Shu [25] and Magnusson [18], relies on first projecting the deviated well trajectory onto the cell faces before using the Peaceman model independently on each component of the trajectory. Finally, the l^2-norm of the components is taken to obtain an appropriate connection factor rep-

Algorithm 1 Procedure to obtain well placement sensitivities from adjoint framework.

$\mathbf{u} = \{v_j, w_j\} \leftarrow$ input from SQP (request of cost function and gradient)
for $n = 1, \ldots, N$ (forward-in-time run) **do**
    solve (6a) for state variables $\widetilde{\mathbf{x}}|_{t=t_n}$
    store $\widetilde{\mathbf{x}}|_{t=t_n}$
end for
evaluate cost function value to be returned to SQP
for $n = N, \ldots, 1$ (backward-in-time run) **do**
    load stored $\widetilde{\mathbf{x}}|_{t=t_n}$
    assemble adjoint variable system (6b)
    solve for adjoint variables $\widetilde{\boldsymbol{\lambda}}|_{t=t_n}$
    for each well control variable v_j **do**
        compute $\nabla_{v_j}\mathcal{J}|_{t=t_n}$ according to (7)
    end for
    for each well configuration variable w_j **do**
        if $n = N$ (first backward-in-time step) **then**
            perturb w_j by $\pm\varepsilon$ and define well as a line segment
            find grid cells intersected by the segment
            gather intersected cells (well blocks) $I_{w_j \pm \varepsilon} = \{i_1, \ldots, i_{N_g}\}$
            compute well connection factors
            store $I_{w_j \pm \varepsilon}$ and connection factors
        end if
        for each well block i_k in stored $I_{w_j \pm \varepsilon}$ **do**
            assemble source/sink terms of g_{i_k}
            form $g_{i_k}(\widetilde{\mathbf{x}}, w_j + \varepsilon) - g_{i_k}(\widetilde{\mathbf{x}}, w_j - \varepsilon)$
        end for
        form $q_{p,i}(\widetilde{\mathbf{x}}, w_j + \varepsilon) - q_{p,i}(\widetilde{\mathbf{x}}, w_j - \varepsilon)$
        approximate $\nabla_{w_j}\mathcal{J}|_{t=t_n}$ using (9)
    end for
    add $\nabla_{\mathbf{u}}\mathcal{J}|_{t=t_n}$ to gradient
end for

resenting the deviated well trajectory. The decomposition performed by this calculation enhances the ability of the well connection factor to consistently capture production differences caused by small-scale geometric changes to the trajectory. Moreover, by taking into account the three-dimensional geometry of slanted wellbores within grid blocks, this calculation contributes to smoothening the cost function objective with respect to well coordinates, and thus ultimately helps increase the accuracy of the gradient

approximation.

We reiterate that the adjoint gradient computation described in Algorithm 1 encompasses only two runs: a *forward-in-time* run performed by a standard reservoir simulator with the capacity to store the state variables, and a *backward-in-time* run performed by a customized adjoint framework with the capability of assembling the flow rates of the residual source/sink terms. This yields an efficient gradient approximation approach compared to a straightforward function-level application of a finite difference scheme that would otherwise require a standalone reservoir simulation for each of its stencil points. Furthermore, the process of assembling the adjoint gradient components (lines 11 through 28 in Algorithm 1) is independent for each variable v_j and w_j. This feature enables potential parallelization, making the adjoint approach efficient vis-a-vis a parallelized finite difference gradient computation.

However, as for any approximation, the gradient estimate obtained through this procedure is subject to various types of error, e.g., truncation errors linked to the central difference scheme in (9). Therefore, to corroborate the overall gradient approximation procedure, it is important to identify and quantify the influence of pertinent error sources. The main goal is to determine, if possible, a perturbation size range within which sufficient accuracy can be obtained to efficiently drive the well placement search. Such error analysis is performed in the next section.

3.4 Consistency of gradient approximation

Finite differences are commonly used within engineering disciplines to approximate the sensitivity of functions. Unlike analytical derivatives, though, finite differences do not take into account the properties of the function in the interval between the current and perturbed point. This can often be considered an advantage, e.g., in the case of function noise or discontinuities, the finite difference gradient may be able to capture the coarse-scale slope of the function. However, a clear shortcoming is that the approximated gradient may not be able to come close to, or may even miss, a local optimum between the two points. In this work, we regard the approximate gradient as containing sufficiently accurate information about the slope of the function to assure convergence since the SQP method used for well placement optimization requires building no more than a coarse approximation of the objective function to find a line search direction. For context, see work by [3] contending an SQP method permits relative gradient errors of 50%, or more. Further details regarding the solver used in this work are given in Section 3.5.

Notwithstanding the above argumentation, a measure of gradient accuracy is important for the performance of the nonlinear programming method used in this work. This section therefore seeks to provide a measure of the quality of the approximated gradient by determining a perturbation range within which the gradient estimate is valid, i.e., consistent. Since exogenous errors may influence the gradient estimate, determining the range of perturbation sizes, within which the error explicitly associated with the approximation is dominant, is a practical way of establishing when the approximation is consistent. In our case, this means finding the perturbation range within which the truncation error is dominant. To find this range, one has to identify the major sources of error that influence the overall estimate and evaluate how each error source varies with respect to the perturbation size. Finding this range can be done, for example, by equating the analytical estimates of the errors and then solving for perturbation size. Such an analysis, based on balancing the simulation and truncation error, is performed by Iott et al. [10] for a data-fitting problem and a wing design simulation. Due to the complexity of the reservoir system, however, a theoretical analysis of this kind cannot be easily done in our case. Instead, to find the range of perturbation sizes within which the truncation error is dominant, we first establish the main error sources, then compare numerically the relevant error orders determined using an extensive sampling of the well placement solution space.

3.4.1 Analysis of error sources

The following three error sources are considered the main components of the total error associated with

the finite difference gradient approximation presented in (9)

$$\nabla_{\mathbf{u}}\mathcal{J} = \widetilde{\nabla}_{\delta\mathbf{u}}\mathcal{J} + E_{\text{cancel}} + E_{\text{trunc}} + E_{\text{discr}}. \quad (10)$$

The first error source (E_{cancel}) is related to the cancellation error caused by subtracting two very similar computed values: As ε goes to zero, the computed values $\mathcal{L}(\widetilde{\mathbf{x}}, \mathbf{u}+\varepsilon\delta\mathbf{u})$ and $\mathcal{L}(\widetilde{\mathbf{x}}, \mathbf{u}-\varepsilon\delta\mathbf{u})$ approach each other; and for small enough ε, round-off errors begin dominating the numerator of (9). The second error source (E_{trunc}) refers to the truncation error associated with the neglected terms from the Taylor series expansion used to approximate the gradient. The third error source, termed discretization error (E_{discr}), is the error associated with the discontinuities originating from the discrete nature of the grid and its components, e.g., the different rock properties for each grid block. These errors are mutually independent, however, their relative contribution to the total error vary according to the perturbation size ε. To determine their relative error contribution, below we establish how the individual errors E_{cancel}, E_{trunc} and E_{discr} depend on ε.

Cancellation error As defined, the cancellation error originates from round-off error and is caused by the subtractions in the numerator of (9). This error is proportional to $\epsilon_{mach}\,\varepsilon^{-1}$, where ϵ_{mach} is the machine epsilon.

Truncation error The truncation error originates from the Taylor expansion underlying the central finite difference approximation (9) and is therefore expected to be of order $O(\varepsilon^2)$.

Discretization error The discretization error is zero for a well and its associated ε-neighborhood enclosed entirely by a zone of homogeneous rock properties (e.g., a grid cell). However, this error materializes as soon as the well trajectory leaves the homogeneous zone and continues to grow as additional cell interfaces are traversed in a heterogeneous field. Note that any point on the well trajectory may contribute to the discretization error. Specifically, perturbing the heel or toe within their respective endpoint grid cells only, does not guarantee (in fact, is often not the case) that other well trajectory points will stay within their respective grid cells and thus not contribute to the error. The magnitude of the discretization error, considering a particular point $\mathbf{u} \pm \varepsilon\delta\mathbf{u}$, does not depend explicitly on ε itself. Rather, it depends on ε indirectly through the number of grid blocks traversed by the well and on the heterogeneity of the grid block parameters. For this reason, a unique power law can not be expressed for this error in the same manner as for the other two error types. This renders a study of its dependence on ε beyond the scope of this work.

Having established how the individual error types are expected to vary with respect to ε, we now consider the properties of their linear combination. In such a combination, depending on the magnitude of ε, the following three scenarios are possible: either one of the three error types prevails predominantly over the others; or, the total error is controlled by some combination of both the truncation and cancellation errors; or, the total error is controlled by some combination of both the truncation and discretization errors. A goal of the error order experiment next is to quantify the relative contribution of each of these error types.

3.4.2 Error order experiment

The following experiment is devised to infer the expected order of the total error, $E_{\text{cancel}} + E_{\text{trunc}} + E_{\text{discr}}$, as a function of ε. The experiment is performed by perturbing the x, y, z coordinates of the heel of a wellbore (the toe remains fixed) and computing the gradient approximations (9) for three perturbation sizes $\{\varepsilon, 2\varepsilon, 4\varepsilon\}$. The following quantities (later referred to as *order estimator*) are then computed numerically

$$R = \frac{\widetilde{\nabla}_{4\delta\mathbf{u}}J - \widetilde{\nabla}_{2\delta\mathbf{u}}J}{\widetilde{\nabla}_{2\delta\mathbf{u}}J - \widetilde{\nabla}_{\delta\mathbf{u}}J}. \quad (11)$$

Following definition (10), a positive R is interpreted using the formula

$$R \approx \frac{O\big(4^n\varepsilon^n\big) - O\big(2^n\varepsilon^n\big)}{O\big(2^n\varepsilon^n\big) - O\big(\varepsilon^n\big)} \approx 2^n. \quad (12)$$

R provides the order of the total error $n \approx \log_2(R)$. For a given $\mathbf{u}$, (11) yields values R_x, R_y, R_z corresponding to the three perturbation directions x, y, z. To sample different configurations of the well, this procedure is repeated for several random locations of the heel. Following this, the sampled values of R_x, R_y are combined, and the quartiles (median, P25 and P75) of R for perturbations in the horizontal direction are computed from the distribution. Similarly, the sampled values of R_z are used to deduce the same quartiles of R for perturbations in the vertical direction. Finally, the above procedure is repeated for varying ε within $(0, \varepsilon_{\max}]$ to obtain the sensitivity with respect to the perturbation size. Figure 1 shows the statistical distributions of R as functions of ε. The order estimator statistics for perturbations performed in the horizontal and vertical directions are shown in figures 1(a) and 1(b), respectively. We treat the horizontal and vertical dimensions separately because of their different scaling in the reservoir grid. The statistical metrics shown in Figure 1 are obtained by evaluating Case 4, described in Section 4, using 150 uniformly-distributed random well endpoint locations.

At this point, we focus this experiment to study the role of the truncation error. This means we concentrate on the scenarios where either the truncation error is dominant, or both the truncation and the discretization error contribute in some measure to the total error. The order of the total error is computed using different perturbation sizes ε ranging from 0.18 to 29 meters for the horizontal direction and from 0.0075 to 1.5 meters for the vertical direction. To avoid a contribution from the cancellation error, the lower bounds of ε above are selected to be appropriately large. Furthermore, the upper bounds of ε are chosen significantly smaller than the characteristic grid cell size of $\Delta x \approx \Delta y \approx 125$ meters and $\Delta z \approx 5$ meters. Recall that even though the selected perturbation range is limited, the discretization error still contributes since points along the modified trajectories cross multiple cell boundaries for the various ε.

Figure 1 allows us to statistically study the influence of the truncation error on the total approximation error. Importantly, in each plot, we are able to identify a plateau of medians corresponding to $O(\varepsilon^2)$ which is the order associated with the truncation error. For very small ε the box plot is almost invisible, i.e., only a tiny distance between lower and upper quartiles, which means the well samples at this perturbation range do not generate discretization error. In the following ε range, the box plots demonstrate a skewed distribution, i.e., only a tiny distance is seen from median to the upper quartile. This indicates that the well samples still cluster largely around the gradient approximation free of discretization error. As ε increases in size, the number of well samples experiencing discretization errors increases and the distribution becomes less skewed. As the median approaches the lower quartile, the truncation error has an inferior relative role within the total error, which is reasonable and implies a greater relative influence by the discretization error. In Figure 1(a) we observe that, for horizontal perturbations, this occurs for ε larger than 3 meters. In Figure 1(b), on the other hand, we see that for vertical perturbations, the plateau of the median ends at $\varepsilon = 0.12$ meters.

It should be stressed that a total error dominated by the truncation error is preferred because under this condition the gradient $\nabla_{\mathbf{u}}\mathcal{J}$ is guaranteed to be consistent with the established approximation $\widetilde{\nabla}_{\delta\mathbf{u}}\mathcal{J}$. Furthermore, recall that the main objective is to determine the bounds of perturbation size ε for which the gradient approximation is still consistent. Given this objective, maintaining the above condition is important because it specifies the range of ε within which we have the greatest control over the gradient error, i.e., within this range we are assured that, by directly varying ε, the accuracy of the gradient can be consistently increased/decreased in an exact, consequent manner as given by (9).

As outlined at the start of this section, the main goal of this analysis is to determine a perturbation size range within which sufficient accuracy can be obtained for the well placement gradient approximation. This analysis is a helpful preprocessing step to obtain useful tuning parameters for the proposed gradient methodology. However, the acquired information is case-dependent, which means the analysis has to performed for each different reservoir model considered. Moreover, the analysis entitles a signifi-

cant computational load since producing a figure such as Figure 1 requires a large sampling of gradient approximations, e.g., the analysis in this section takes 150 pairs of forward- and backward-in-time simulations. In this regard, future applications are expected to decrease computational load significantly through parallelization of Algorithm 1. Finally, a main assumption behind this analysis is that the obtained ranges remain valid for all feasible well configurations explored by the optimization procedure. Future work will further explore this assumption, and, through further analysis involving more case study data, also attempt to develop general rules that can apply across different reservoir models.

3.5 SQP and constraint handling

In our work, the well placement optimization problem (1) is solved using SQP code implemented in the sparse nonlinear optimizer (SNOPT) library. A detailed description of this library can be found in [9]. This code, through a sequence of $\nabla_{\mathbf{u}}\mathcal{J}$, constructs and updates a quadratic model of $\mathcal{J}$. This quadratic model is solved as a QP-subproblem and contributes to the update of $\mathbf{u}$ through a line search procedure. Additionally, the quadratic model of $\mathcal{J}$ may be supplemented with linear models of the nonlinear constraint functions.

An important feature of the SQP solver is that it deals with the various constraints in a proactive manner, i.e., solution candidates are maintained largely feasible throughout the optimization. In our work, this strategy is instrumental in satisfying the well placement constraints such as restrictions on the well length, angle of inclination, direction, inter-well distance, and signed well-to-boundary distance. Crucially, these constraint functions are independent of the state variables $\mathbf{x}$ and therefore do not rely on the adjoint framework. The linearization of these functions is done either using analytical derivatives or approximations with a central difference scheme.

Several of the constraints mentioned earlier have been implemented, as necessary, for the various cases tested in this work. As described in Section 3.1, both vertical and deviated wellbores are parametrized as line segments in three-dimensional space. Given this parametrization, practically all well placement constraint-handling computations entitle solving a series of, often standard, geometric problems dealing with the relative positioning of lines and points in three dimensional space, e.g., finding the minimum distance between a line segment and a point in space. Because of this, most constraint enforcement in this work is implemented using readily available, efficient geometric algorithms [5].

Still, some specific algorithmic developments, have been necessary to deal with the inter-well distance constraint (implemented for all study cases in Section 4) and the signed well-to-boundary distance constraint featured in cases 2 and 4. Next we explain some of the main features of this constraint-handling methodology.

Inter-well distance In the $\mathbb{R}^6$-parameterization of wellbores, the inter-well distance metric used in this work between two wells is the shortest distance between two line segments. In this study, we implement an efficient algorithm adapted from [5] to compute this metric. The distance to a group of wells is then defined as the distance to the nearest well in the group.

Signed well-to-boundary distance In this work, a boundary constraint that maintains a wellbore within the reservoir grid is implemented by computing a signed well-to-boundary distance. Since the well-to-boundary distance has to be computed frequently during optimization, in this work, we use a discrete analog of the boundary consisting of a set of three-dimensional points (this analog is later referred to as discretized boundary). For example, the reservoir boundary may be represented by the outermost grid cell vertices of the boundary cells. At each constraint function evaluation, the Euclidean distance to the boundary point closest to the wellbore segment is computed; see Algorithm 2 for pseudo-code of this computation. The computation is based on a point-to-segment algorithm (lines 6 through 19) found in [5] and a sign-guessing procedure (lines 21 through 26) introduced ad-hoc by the authors. The latter classifies the distance as positive if the wellbore is inside

the boundary and negative if outside. Future work will involve region constraints that can enclose wells within promising reservoir areas and/or sands delimited by geological features such as faults.

Algorithm 2 Procedure for estimating signed well-to-boundary distance for a single well

$\vec{O} \leftarrow$ centroid of discretized boundary
$\vec{r} \leftarrow \vec{P}_{\text{toe}} - \vec{P}_{\text{heel}}$
$R = \langle \vec{r}, \vec{r} \rangle + \epsilon_{mach}$
$d \leftarrow \infty$
for each point $\vec{Q}_i$ of discretized boundary **do**
    $s \leftarrow \langle \vec{r}, \vec{Q}_i - \vec{P}_{\text{heel}} \rangle$
    if $s \leqslant 0$ (near heel) **then**
        $\vec{P} \leftarrow \vec{P}_{\text{heel}}$
    else if $s \geqslant R$ (near toe) **then**
        $\vec{P} \leftarrow \vec{P}_{\text{toe}}$
    else
        $\vec{P} \leftarrow \vec{P}_{\text{heel}} + \vec{r}\, s\,/R$
    end if
    $d_{\text{old}} \leftarrow d$
    $d \leftarrow \min(d, \langle \vec{P}_i - \vec{P}, \vec{P}_i - \vec{P} \rangle)$
    if $d < d_{\text{old}}$ **then**
        $\vec{Q}_{\min} \leftarrow \vec{Q}_i$
        $\vec{P}_{\min} \leftarrow \vec{P}$
    end if
end for
for $i \in x, y, z$ **do**
    $\alpha_i \leftarrow \min(P_{i,\min}, Q_{i,\min}, O_{i,\min})$
    $A_i \leftarrow \max(P_{i,\min}, Q_{i,\min}, O_{i,\min})$
    scale $P_{i,\min}, Q_{i,\min}, O_{i,\min}$ by $(A_i - \alpha_i)^{-1}$
end for
distance $\leftarrow \text{sign}\langle \vec{P}_{\min} - \vec{Q}_{\min}, \vec{O} - \vec{Q}_{\min} \rangle \sqrt{d}$

The sensitivities of both the inter-well and well-to-boundary distance functions are calculated using finite differences. These sensitivities, in addition to constraint violation function values, are passed to the SQP solver which uses linearized forms of the constraints within its internal constraint-handling logic. This logic restricts iterates to lie within a feasible region D up to a given tolerance. It should be noted that feasibility in the SQP solver is influenced by the accuracy of the linear approximation of the constraints. Thus, original constraints that are highly nonlinear may generate infeasible iterates. Details regarding how the SQP deals with constraints can be found in [9].

3.6 Gradient-based termination criteria

The following termination criteria (in order of verification order) are enforced on the gradient-based procedure in this work:

1. maximal number of solver iterations;
2. first-order optimality condition;
3. stationary point.

The first criterion is straightforward. The practical implications of the second and third criteria are described next.

The first-order optimality criterion is used by the SQP solver to determine convergence to a local optimum at which point the optimization is terminated. In practice, this criterion is satisfied when the magnitude of the gradient is less than a given tolerance. As discussed in Section 3.4, the accuracy of the well placement gradient is dependent on the perturbation size ε. If ε has a lower bound, then the truncation error corresponding to the smallest nontrivial ε may be larger than the specified optimality tolerance. This problem arises, e.g., in a discrete well placement problem where a cost function change is obtained only for a perturbation $\varepsilon\delta\mathbf{u}$ large enough to displace the wellbore to a different grid block. The truncation error for such an ε would render the first-order optimality criterion inoperable which is likely to reduce algorithmic performance by wasting SQP iterations trying in vain to satisfy the condition. For this reason, the implementation in this work has, for each case, matched gradient accuracy (by modifying ε) to the scale of the tolerance, or vice-versa, adjusted

the solver tolerance according to the possible lower-bound gradient accuracy attainable by the particular problem.

The third criterion considers the stationary point of the problem. The stationary point is defined as a subsequent repetition of the gradient norms over multiple iterations without convergence. In practice, the condition verifies, within a specified tolerance, whether the sequence of gradient norm values consisting of the last three solver iterations has already occurred. If this condition is true, the optimization process is terminated, and the current solution is accepted as an optimal solution.

Next, we apply the methodology to four cases. The first two cases test the validity and performance of the methodology against other approaches in the literature. The last two cases study the performance of the methodology given strict heterogeneity and a more complex grid, respectively.

4 Application

Four cases are presented to both compare the adjoint-based well placement gradient method against current approaches in the literature, and to test the methodology on more complex cases involving deviated wellbores. In Case 1 and 2, we test the methodology against results from a direct and an indirect approach, presented by Li and Jafarpour [16] and Zandvliet et al. [29], respectively. Notice that these proof-of-concept cases contain only vertical wells. The main goal behind these tests is to show that the coupling of the gradient approximation approach with the SQP solver presented in this work yields an efficient method that competes against state-of-the-art gradient-based methods on their terms. The other two cases optimize the placement of horizontal and deviated wellbores and are based, respectively, on a binary channelized system (Case 3) and the Brugge reservoir (Case 4).

Since joint optimization of well placement and controls is not the main focus of this work, control optimization is only considered in one of the two benchmark cases (i.e., Case 1 includes the allocation of well rates) primarily to compare results with Li and Jafarpour [16]. For all other cases, well control parameters are kept fixed.

None of the following case studies implement economic limits in the underlying simulations. A recent study by [27] demonstrates that economic limits can significantly influence the consistency of the adjoint gradients with respect to well controls. Future work will extend this study to the development of adjoint-based well placement gradients.

Reservoir simulations in this work are performed using Stanford's Automatic-Differentiation-based General Purpose Research Simulator (AD-GPRS) [see 8]. One of the distinctive features of AD-GPRS is the capacity to store state variables and reassemble, on demand, the constitutive parts of the residual g and cost functional $\mathcal{J}$. This feature is illustrated in lines 4, 8 and 9 of Algorithm 1. AD-GPRS therefore serves as a well-suited and highly-flexible platform for the development of adjoint-based gradient methodology. The different comparisons against literature cases, in terms of final results and performance, are described next.

4.1 Case 1: SPE10 layer.

In Case 1, we implement the optimization case study corresponding to "Example 4: SPE10 top layer" in [16]. This means that we use the exact same reservoir model, production and well settings to reproduce the results presented in that paper. In the SQP, all optimization variables are continuous including w_j, which here are vertical well (x, y)-plane coordinates relative to a corner of the grid. However, to be consistent with [16], well connection factors for intersected grid cells are evaluated using Peaceman's formula for vertical wells and the horizontal perturbation size ε is set to Δx.

Li and Jafarpour [16] solved two problems: (1) they optimized single well placement and (2) they jointly optimized well placement and rate allocation, using gradients derived from an SPSA procedure. In our study, the same two problems are solved using the adjoint-based gradient approximation methodology presented in Section 3.

Case 1 comprises a two-dimensional isothermal incompressible oil-water reservoir model with a $60 \times$

220×1 grid (13200 blocks) with cell dimensions $\Delta x = \Delta y = 10$m, $\Delta z = 1$m. The geological distribution, shown in Figure 2(a) is taken from the first layer of the SPE10 model. A constant porosity of 0.2 is used. Optimization variables are the allocation of well fluid rates and location of fifteen vertical wells: five injectors and ten producers. As in [16], we inject and produce the fluid mass equivalent of one reservoir pore volume.

Figures 2(b) and 2(c) show the oil saturations at the end of reservoir life observed for the optimal well location and rates obtained using the adjoint-based well-placement-gradient procedure. The two results are analogous to the ones presented by Li and Jafarpour [16]. Figure 2(b) shows the oil saturation after optimizing only for well placement, while Figure 2(c) shows the saturation after performing optimization on both well location and rates. In both cases, minimum inter-well distance for all wells is set to 150m. In [16], the SPSA-based routine applied to well placement optimization yields an NPV increase of 43.7% after more than 500 iterations. The solution to the same problem using the adjoint-based routine presented in this paper yields an NPV increase of 57.1% after 141 function and 46 gradient evaluations, i.e., a total computational cost equivalent to 187 reservoir simulations. Qualitatively, Figure 2(b), compared to its analog figure in [16] (Figure 7b1), shows two producers moving away from the bounds, and a marked decrease in oil saturation in the western and northern regions of the layer.

Similarly, in [16], the SPSA-based routine optimizing well placement and controls jointly (as mentioned, using an alternate solution procedure) yields a 78.5% NPV increase after approximately 2000 iterations. The solution to the same problem using the adjoint-based procedure from this work yields an NPV increase of 79.5% after a total of 332 equivalent function evaluations (252 forward simulations and 80 gradient calculations). In combination, these results validate the well placement gradient approximation method proposed in this work, and they demonstrate the efficiency of the overall procedure, in terms of final result and performance, by arriving at improved or close-to equivalent solutions at less computational cost. This efficiency is seen as due to a combination of crucial improvements by the proposed procedure at both the gradient-approximation and algorithm levels. The former due to the effectiveness of the adjoint framework in deriving the gradient, while the latter due to the more accurate search direction provided by this gradient.

4.2 Case 2: Egg model.

Case 2 compares results obtained using the methodology proposed in this work against results from [29] where an indirect gradient-approximation approach based on pseudo-wells is used to optimize the location of a set of vertical wells. Here we have reimplemented the case corresponding to "Example 3" in [29] which consists of optimizing the location of eight injectors and four producers. The optimization is run from two initial well configurations, referred to as "standard" and "mini" configurations; these configurations are illustrated in Figure 3(a) and Figure 4(a), respectively.

Figure 3(a) and Figure 4(a) also show the log-permeability distribution of the reservoir model. The reservoir consists of a $60 \times 60 \times 7$ grid (with 18553 active cells) with cell dimensions $\Delta x = \Delta y = \Delta z = 20$m. Further details about the model can be found in [29].

Notice that in this model the vertical wells are completed through all 7 layers. Here, the standard Peaceman model is used for the well connection factor computation of the vertical wells, while optimization variables are the well location coordinates in (x, y)-plane determined relative to one of the corners of the grid. For consistency with the Peaceman model, the horizontal perturbation size ε is set to Δx. In our implementation, we supply the following nonlinear well placement constraints to the SQP solver to regularize the solution. These constraints, inter-well and signed well-to-boundary distance constraints, respectively, consist of keeping the wells two grid blocks apart from each other and within the region of the active (x, y)-plane of grid blocks. Optimizations runs are started using the two initial configurations and the same control and economic constraint (reactive shut-in based on water-cut) settings as the original work have been implemented.

Figure 3(b) and Figure 4(b) show the final oil saturation and well configurations obtained using the methodology presented in this work. In Figure 5, we show the NPV evolution curves from the original publication as black lines, i.e., the values of these curves have been copied from Figure 11 in [29]. In Figure 5, red lines correspond to NPV curves obtained using the methodology proposed in this work.

Comparing the final well configurations against those obtained by Zandvliet et al. [29], we observe that, for both optimization runs, i.e., starting using either the standard or the mini initial points, the resulting well locations are qualitatively a close match. For the standard initial point, the optimal well configuration ends in a local optimum with essentially the same features as the original five-spot pattern. For the mini initial point, the well locations evolve into a linear well configuration pattern with the producers aligned in the middle and injectors sitting on the west and east flanks.

The curves shown in Figure 5 demonstrate that, in terms of the NPV and equivalent number of function evaluations, the methodology presented in this work yields close-to equal results to the ones presented by Zandvliet et al. [29]. Compared to these results, the methodology proposed here yields NPV differences of 0.4% and -0.8%, when starting from the standard and mini initial positions, respectively.

In Figure 5, the function evolution curves for the standard initial configuration follow similar paths, i.e., these curves start from a high function value and obtain only a modest increase. This progression is indicative of an initial point close to a local optimum. On the other hand, the function evolutions curves corresponding to the mini initial configuration start from a clearly suboptimal point. Here the curves for each of the methods take distinct paths, in particular in the beginning, e.g., the methodology presented in this work shows a steeper increase at the start, and eventually converges to a slightly lower optimum. Overall, this comparison demonstrates that the gradient-approximation method presented in this work yields a sufficiently accurate direction of cost function increase and that this direction can be used to efficiently drive a gradient-based algorithm.

Next, the versatility and performance of the methodology are tested on two more advanced production cases in terms of well configuration: the third case in this work includes four horizontal wellbores and a vertical injector within a heavily channelized geology, while the fourth case consists of two deviated wellbores within a realistic three-dimensional grid.

4.3 Case 3: Two-dimensional case in channelized permeability field.

Case 3 involves a binary channelized permeability field mapped onto a two-dimensional uniform 60×60 grid with $\Delta x = \Delta y = \Delta z = 24$m. The permeability is taken from a set of realizations generated in [24]. The sand permeability is 500mD, while the shale permeability is 10mD. The field, along with the initial well locations, is shown in Figure 6. Notice the four producers have short initial well lengths that range over two grid blocks only, while the injector in the middle is forced to remain vertical throughout the optimization. Initially, the reservoir is at equilibrium with an average pressure of 146bar. The well controls are fixed and set to ensure uninterrupted production during the production time frame of 500 days. The bottom-hole pressures of the producers and the injector are respectively 20 and 200bar. The horizontal length of the producers is constrained to a maximum of 350m, whereas the minimum inter-well distance for all wells is set to be 350m. Despite the relatively simple channelized structure, this case poses a challenging well placement problem because the abrupt changes in the permeability field yield a nonsmooth solution space. The objective for this case is NPV with only oil production with oil price set to 75\$/bbl and cost of horizontal producer trajectory 500\$/feet. The discount factor is set to 10%. The statistical distribution of total error orders with respect to the horizontal perturbation size is given in Figure 7. Based on this figure, two sizes of ε are selected for testing, i.e., 15m and 24m. Figure 7 shows that the former ε marks the end of the range where the truncation error ($R = 4$) is present within the upper quartile, while the latter ε (corresponding to grid dimension Δx for this case) is associated with predominantly the discretization error, (see the analysis in Section 3.4.2).

Results for this case are presented in figures 8 and 9. Figure 8 shows the initial and final well configurations, and the resulting saturation maps for each configuration, while Figure 9 shows the corresponding function evolution graph. For $\varepsilon = 24$m, the increase of cost function after 280 evaluations is from $4.061 \cdot 10^8$ to $6.285 \cdot 10^8$, i.e., an increase of 54.8%. For $\varepsilon = 15$m, however, the NPV increased to $6.319 \cdot 10^8$ after only 97 evaluations, i.e., an increase of 55.6%. These results are a clear demonstration of how a more accurate gradient approximation, i.e., using $\varepsilon = 15$m, increases the effectiveness of the gradient-based procedure.

Comparing the initial position against the final solution we observe a significant change in well configuration, both in terms of length and orientation, while both well length and inter-well distance constraints are satisfied. In Figure 8 we see the injector moving from a location at the border of a high-permeability channel to the middle of the high permeability area at the intersection of two main channels. This repositioning seems reasonable as it ensures the oil in this area is effectively pushed towards the surrounding wells. As expected, the producers extended their lengths to increase their drainage area, and two producers have orientation closely aligned with the direction of the channels. We also observe that the heels of three of the wells cross areas of low permeability (due to shale facies) to reach into a neighboring channel.

4.4 Case 4: Three-dimensional production case with deviated wellbores and realistic constraints.

Case 4 has been built using model data from the Brugge benchmark case [20]. Its purpose is to provide a more realistic grid and production scenario to test the performance of the proposed methodology. The same grid geometry and static cell data as the realization #73 of Brugge benchmark have been used, while modifications have been made such as a slight rotation of the grid, shallower oil-water and gas-oil contacts (now at 1610m and 1518m, respectively). In this case, the original fluid formulation (two phase dead oil) has been replaced with a three phase black oil formulation thus allowing free gas to form at the top of reservoir. The intention is to create a more challenging three-way trade-off between oil, gas and water production in the cost function which comprises a non-discounted NPV with oil price, water production cost and gas production cost equal to 47.7\$/bbl, 0.79\$/bbl and 5.3\$/MMscf, respectively. In this setup, e.g., wellbores have to avoid both high gas and/or water saturation areas. The new relative permeability data is shown in Figure 10. Well length penalization is not included in the cost.

Case 4 includes two deviated production wells and two vertical injectors. The well placement problem consists of optimizing the spatial configuration of the two producers. The injectors are placed at the outskirt of the reservoir for pressure support, and their location is fixed, i.e., these well are not part of the optimization. All wells are controlled using fixed BHP pressure, injectors are set at 170bar, while producers are set at 70bar. Production time frame for the case is 6 years. Figure 10 shows the permeability of the sector of the reservoir model that includes the producers (this sector is large enough to study the relevant saturation changes and provides an appropriate close-up of the resulting well configurations, and is therefore used in all remaining figures).

The production wellbores are represented as straight segments within the reservoir volume. As stated in Section 3.1, each horizontal well is parametrized using the spatial heel-toe endpoint coordinates. Thus, Case 4 has in total twelve optimization variables. Initially, the reservoir is produced by two very short wells traversing the upper part of the oil rim, see Figure 12(a). These short wells do not supply sufficient drainage capacity for this production scenario. Instead, this starting point is chosen to provide the optimization with ample improvement potential. The main goal of this setup is to study the efficiency of the gradient-based, constrained, local-search algorithm.

In Case 4, the following three types of nonlinear constraints are imposed on possible well configurations:

1. *well length constraint*: wellbores cannot be longer than 1500m;

2. *inter-well distance constraint*: any point along a wellbore cannot be closer than 400m to any point along any other wellbore;

3. *reservoir boundary constraint*: wellbores are kept within the reservoir grid by measuring the distance between any given well endpoint and the outer vertices of the reservoir grid.

The implementation of these three constraints is described in Section 3.5. In addition to these constraints, the upper and lower bounds on the well variables serve as a secondary boundary constraint delimiting a (cubic) volume of the reservoir. The x, y-sides of this box constraint are shown as white borders in figures 10, 12(c) and 12(d).

The results from this case are discussed in the following. Figures 12(a) and 12(b) show the initial and final well configurations, respectively. Comparing the optimal positioning of the wells in Figure 12(b) to the initial configuration in Figure 12(a), we observe that the final wellbores have remained at the top of the oil rim and that their lengths have expectedly increased from initial values of about 130m to well lengths of 1343m and 1395m. Note, however, that these wells did not reach their maximum allowed value of 1500m. Instead, this local maximum is the result of the trade-off between oil, gas, and water production. Overall, the final well configuration covers a large part of the allotted feasible area, which is a reasonable solution to maximize oil production, with the heel of one of the wells touching the boundary of the box constraint. Well coordinate changes (in meters) from initial to final configuration are

$$\begin{aligned}
&\text{producer \#1}\\
&\quad \Delta\vec{P}_{\text{heel}} = \{-376.9,\ -147.8,\ -9.5\},\\
&\quad \Delta\vec{P}_{\text{toe}} = \{817.3,\ 13.7,\ 37.7\};\\
&\text{producer \#2}\\
&\quad \Delta\vec{P}_{\text{heel}} = \{-1014.3,\ 79.1,\ 21.8\},\\
&\quad \Delta\vec{P}_{\text{toe}} = \{236.8,\ -98.7,\ 13.9\}.
\end{aligned}$$

For reference, the size of the horizontal bounding box (white) (shown Figure 12) is (4400m, 1350m). This result signifies a substantial development of the initial solution by the search procedure.

The distribution of oil, gas and water saturations at end of production time corresponding to the initial well configuration is shown in Figure 12(c), while the distribution corresponding to the optimal well configuration is shown in Figure 12(d). Comparing Figure 12(d) with 12(c) we see a clear decrease in oil saturation for the whole segment corresponding to the longer wellbores. Figure 13 shows the progression of the cost function where the NPV increases from $2.591 \cdot 10^9$ to $1.106 \cdot 10^{10}$ using 333 forward reservoir simulations and 333 backward gradient evaluations. From this progression we see that the customized convergence criteria described in Section 3.6 are efficient in the sense that they allow the algorithm to proceed without stopping it too early, allowing step-increases throughout.

5 Summary

This paper presents an efficient approach for approximating the well placement gradient through the adjoint framework. Three main categories: continuous, direct and indirect, have been defined to describe the multiple well placement gradient-approximation approaches proposed in the literature, and to provide context to the adjoint-based approach presented in this work. Common for these approaches is the goal of deriving well placement sensitivities to enable the use of local-search optimization procedures. This work implements an SQP-based procedure coupled with efficient well placement constraint enforcement. A core advantage of this methodology compared to the other approaches is the efficiency of the sensitivity estimation inherited by performing the approximation within the adjoint solution framework. Additionally, the flexibility of the adjoint-based approach enables the straightforward computation of gradients of associated functionals which facilitates the implementation of efficient constraint-handling.

Furthermore, this paper elaborates on three important implementation aspects: (1) the accuracy of the approximate gradient, (2) termination criteria and (3) boundary constraint-handling. In terms of gradient accuracy, this work performs an extensive error analysis to determine the limits within which the

approximation is consistent. Moreover, concerning termination criteria, this work presents a hierarchy of reasonable conditions suitable for the SQP implementation and discusses their practical implications. Finally, a low-cost algorithm for evaluating signed well-to-boundary distance is introduced to deal with realistic reservoir boundaries.

Finally, the adjoint-based gradient approximation approach has been tested against indirect and direct gradient computation approaches from the literature. Compared to a direct approach based on stochastic approximation the adjoint-based approach is shown to be substantially more efficient in terms of the total number of required function evaluations. Compared to an indirect approach that relies on pseudo-wells, the adjoint-based approach demonstrates similar performance in terms of final result and the total number of function evaluations. However, the adjoint-based approach has a clear advantage in terms of general applicability, i.e., the proposed approach can in a straightforward manner be used to solve for complex configurations involving deviated wellbores. This advantage is demonstrated through a realistic optimization case featuring two horizontal wellbores subject to multiple well placement constraints. In this case, the approach arrives at a practical solution following a reasonable cost function progression and termination.

The ability to derive constraint derivatives is a significant advantage of the methodology presented in this work because it allows us to use state-of-the-art gradient-based constraint-handling techniques. In general, an efficient constraint-handling capability is critical to retain search efficiency and avoid poor solution sampling. This is particularly pertinent for the well placement problem since this problem typically involves computationally expensive cost functions and nonconvex solution spaces caused by nonlinear geological and engineering restrictions. In terms of further work, additional tests of the adjoint-based gradient approximation methodology can be performed both using models with economic production constraints and in the joint solution of problems involving well placement, control, type (i.e., injector, producer, or shut), and/or completion design. The further study and testing of the proposed methodology in several of these research areas is the subject of ongoing work, in addition to work involving benchmarking the adjoint-based procedure against derivative-free methodologies and the development of hybrid algorithms.

References

[1] Wolfgang Bangerth, Hector Klie, Vincent Matossian, Manish Parashar, and Mary F. Wheeler. An autonomic reservoir framework for the stochastic optimization of well placement. Cluster Computing, 8(4):255–269, 2005. ISSN 1386-7857. doi: 10.1007/s10586-005-4093-3.

[2] Melvyn S Berger. Nonlinearity and functional analysis: lectures on nonlinear problems in mathematical analysis. Academic Press, 1977.

[3] Richard G Carter. Numerical experience with a class of algorithms for nonlinear optimization using inexact function and gradient information. SIAM Journal on Scientific Computing, 14(2): 368–388, 1993.

[4] David Castineira and Faruk Omer Alpak. Automatic well placement optimization in a channelized turbidite reservoir using adjoint based sensitivities. In SPE Reservoir Simulation Symposium, 2-4 February, The Woodlands, TX, 2009. ISBN 978-1-55563-209-0. doi: 10.2118/119156-MS. SPE-119156-MS.

[5] David H Eberly. 3D game engine design: a practical approach to real-time computer graphics. CRC Press, 2006.

[6] Fahim Forouzanfar and A.C. Reynolds. Well-placement optimization using a derivative-free method. Journal of Petroleum Science and Engineering, 109(C):96–116, 2013. ISSN 0920-4105. doi: 10.1016/j.petrol.2013.07.009.

[7] Fahim Forouzanfar and A.C. Reynolds. Joint optimization of number of wells well locations and controls using a gradient-based algorithm. Chemical Engineering Research and Design, 92

(7):1315–1328, 2014. ISSN 0263-8762. doi: 10.1016/j.cherd.2013.11.006.

[8] TT Garipov, P Tomin, R Rin, DV Voskov, and HA Tchelepi. Unified thermo-compositional-mechanical framework for reservoir simulation. Computational Geosciences, pages 1–19, 2018.

[9] Philip E. Gill, Walter Murray, and Michael A. Saunders. SNOPT: An SQP algorithm for large-scale constrained optimization. SIAM Journal on Optimization, 12(4):979–1006, 2002. ISSN 1052-6234. doi: 10.1137/S1052623499350013.

[10] Jocelyn Iott, Raphael T Haftka, and Howard M Adelman. Selecting step sizes in sensitivity analysis by finite differences. NASA Technical Memorandum 86382, 1985.

[11] B. Jafarpour and L. Li. Generalized field development optimization: Coupled well-placement and control under geologic uncertainty. In 13th European Conference on the Mathematics of Oil Recovery, 10-13 September, Biarritz, France, 2012. doi: 10.3997/2214-4609.20143198.

[12] M. Jesmani, B. Jafarpour, M.C. Bellout, R.G. Hanea, and B. Foss. Application of simultaneous perturbation stochastic approximation to well placement optimization under uncertainty. In 15th European Conference on the Mathematics of Oil Recovery, 29 August-1 September, Amsterdam, The Netherlands, 2016. doi: 10.3997/2214-4609.201601873.

[13] M. Jesmani, B. Jafarpour, M.C. Bellout, and B. Foss. A reduced random sampling strategy for fast robust well placement optimization. Journal of Petroleum Science and Engineering, 25(3-4): In review., 2017.

[14] Mansoureh Jesmani, Mathias C. Bellout, Remus Hanea, and Bjarne Foss. Well placement optimization subject to realistic field development constraints. Computational Geosciences, 20(6):1185–1209, 2016. ISSN 1573-1499. doi: 10.1007/s10596-016-9584-1.

[15] Olwijn Leeuwenburgh, Paul J. P. Egberts, and Oscar A. Abbink. Ensemble methods for reservoir life-cycle optimization and well placement. In SPE/DGS Saudi Arabia Section Technical Symposium and Exhibition, 4-7 April, Al-Khobar, Saudi Arabia, 2010. ISBN 978-1-55563-848-1. doi: 10.2118/136916-MS. SPE-136916-MS.

[16] Lianlin Li and Behnam Jafarpour. A variable-control well placement optimization for improved reservoir development. Computational Geosciences, 16(4):871–889, 2012. ISSN 1420-0597. doi: 10.1007/s10596-012-9292-4.

[17] Lianlin Li, Behnam Jafarpour, and M.Reza Mohammad-Khaninezhad. A simultaneous perturbation stochastic approximation algorithm for coupled well placement and control optimization under geologic uncertainty. Computational Geosciences, 17(1):167–188, 2013. doi: 10.1007/s10596-012-9323-1.

[18] Hilmar Magnusson. Development of Constraint Handling Techniques for Well Placement Optimization in Petroleum Field Development. MSc Thesis, Department of Industrial Mathematics, NTNU, 2016.

[19] D. W. Peaceman. Interpretation of well-block pressures in numerical reservoir simulation. SPE Journal, 18:183–194, 1978. ISSN 0197-7520. doi: 10.2118/6893-PA. SPE-6893-PA.

[20] E. Peters, Y. Chen, O. Leeuwenburgh, and D.S. Oliver. Extended Brugge benchmark case for history matching and water flooding optimization. Computers and Geosciences, 50:16–24, 2013.

[21] R-E Plessix. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. Geophysical Journal International, 167(2):495–503, 2006.

[22] Pallav Sarma. Efficient Closed-Loop Optimal Control of Petroleum Reservoirs Under

Uncertainty. PhD Thesis, Stanford University, 2006.

[23] Pallav Sarma and Wen H. Chen. Efficient well placement optimization with gradient-based algorithms and adjoint models. In Intelligent Energy Conference and Exhibition, 25-27 February, Amsterdam, The Netherlands, 2008. ISBN 978-1-55563-166-6. doi: 10.2118/112257-MS. SPE-112257-MS.

[24] Mehrdad Gharib Shirangi. Advanced Techniques for Closed-Loop Reservoir Optimization Under Uncertainty. PhD thesis, Stanford University, 2017.

[25] Jones Shu. Comparison of Various Techniques for Computing Well Index. MSc Thesis, Stanford University, 2005.

[26] Stijn Vlemmix, Gerard J. P. Joosten, Roald Brouwer, and Jan-Dirk Jansen. Adjoint-based well trajectory optimization. In EUROPEC/EAGE Conference and Exhibition, 8-11 June, Amsterdam, The Netherlands, 2009. ISBN 978-90-7378-167-2. doi: 10.2118/121891-MS. SPE-121891-MS.

[27] Oleg Volkov and Mathias C. Bellout. Gradient-based production optimization with simulation-based economic constraints. Computational Geosciences, 21(5):1385–1402, 2017. ISSN 1573-1499. doi: 10.1007/s10596-017-9634-3.

[28] Chunhong Wang, Gaoming Li, and Albert Coburn Reynolds. Optimal well placement for production optimization. In SPE Eastern Regional Meeting, 2007. ISBN 978-1-55563-227-4. doi: 10.2118/111154-MS. SPE-111154-MS.

[29] Maarten Zandvliet, Martijn Handels, Gijs van Essen, Roald Brouwer, and Jan-Dirk Jansen. Adjoint-based well-placement optimization under production constraints. SPE Journal, 13(4): 392–399, 2008. SPE-105797-PA.

[30] Kai Zhang, Gaoming Li, Albert C. Reynolds, Jun Yao, and Liming Zhang. Optimal well placement using an adjoint gradient. Journal of Petroleum Science and Engineering, 73(3):220–226, 2010. ISSN 0920-4105. doi: 10.1016/j.petrol.2010.07.002.

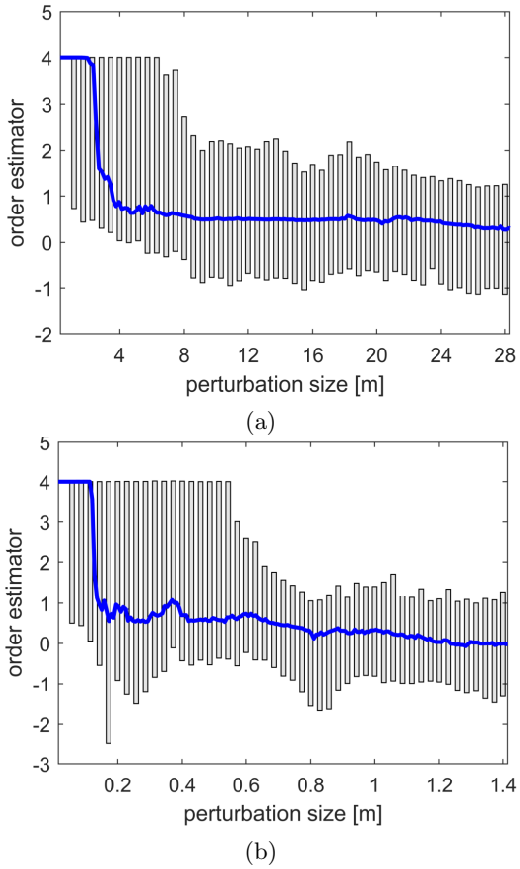


Figure 1: Statistical distribution of order estimator values as functions of perturbation size ε for perturbations performed in (a) horizontal and (b) vertical directions. The median (blue line), and quartiles P25 and P75 (gray box plots) are based on Case 4 using 150 well location samples.

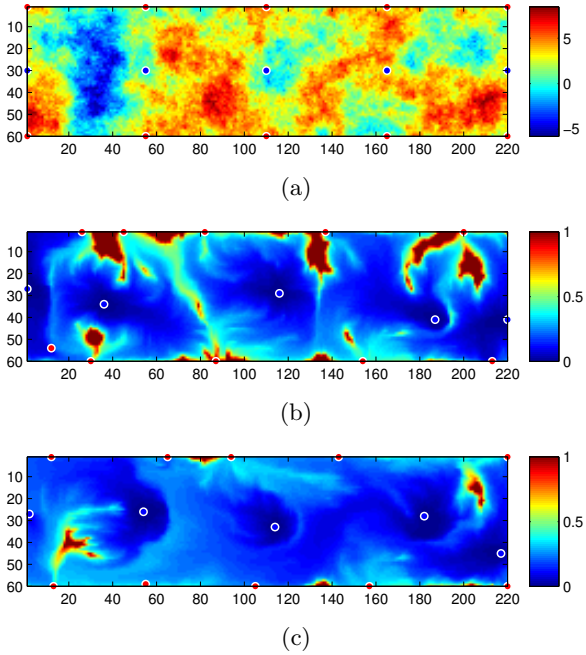
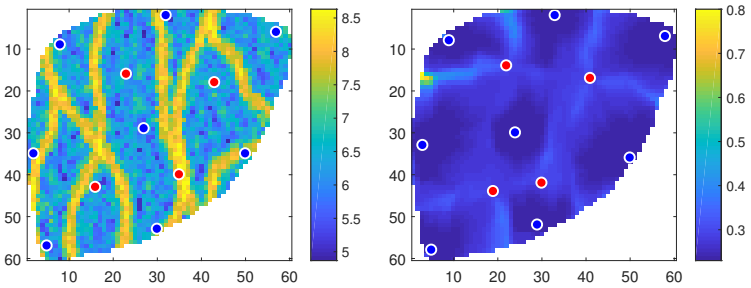
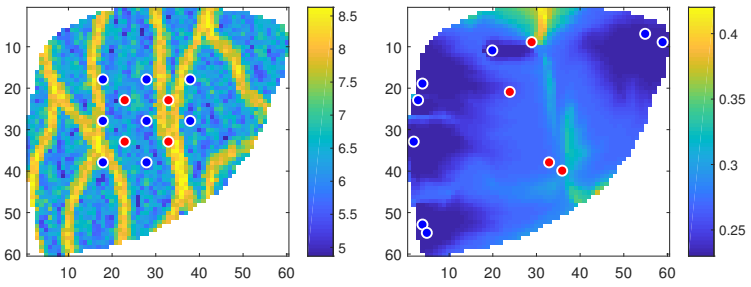


Figure 2: Case 1 results using layer from SPE10 model. (a) Model log-permeability in x- and y-directions with initial well configuration; injectors (blue dots), producers (red dots). (b) Oil saturation and optimal well placement configuration after optimization of well location. (c) Oil saturation and optimal well placement configuration after optimization of both well placement and rate controls.



(a) (b)

Figure 3: Case 2: (a) Log-permeability and well location for standard initial configuration. (b) Corresponding well placement solution. Production and injector wells drawn as red and blue circles, respectively.



(a) (b)

Figure 4: Case 2: (a) Log-permeability and well location for mini initial configuration. (b) Corresponding well placement solution. Production and injector wells drawn as red and blue circles, respectively.

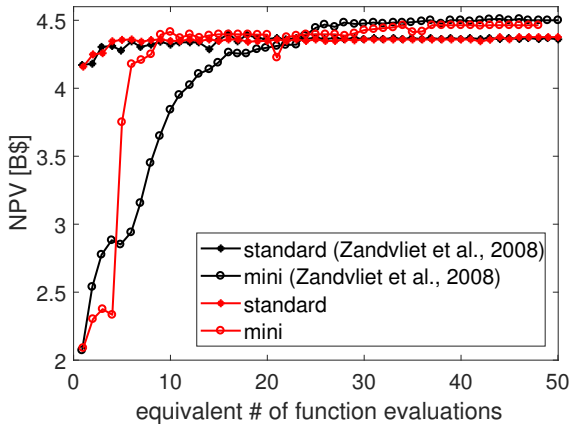


Figure 5: Case 2 results (Egg model): NPV function evolution curves as a function of (equivalent) number of function evaluations (reservoir simulations). Black curves correspond to values reported by Zandvliet et al. [29]. Red curves correspond to values obtain using the methodology described in this paper.

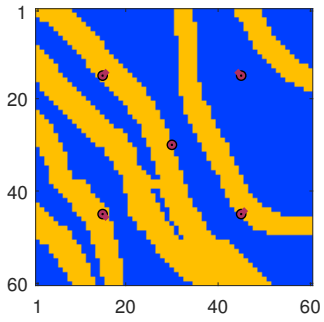


Figure 6: Case 3: Permeability field with orange representing sand facies (500mD) and blue representing shale facies (10mD). The initial well configuration is shown with a vertical injector (in the middle) and two-block trajectories of four horizontal producers.

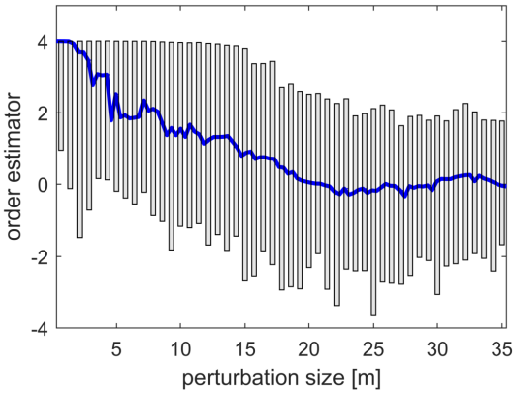


Figure 7: Case 3: Statistical distributions of values of the order estimator R as functions of perturbation size ε for perturbations performed in horizontal directions. The median (blue line), and quartiles P25 and P75 (gray box plots) are obtained using 80 well location samples.

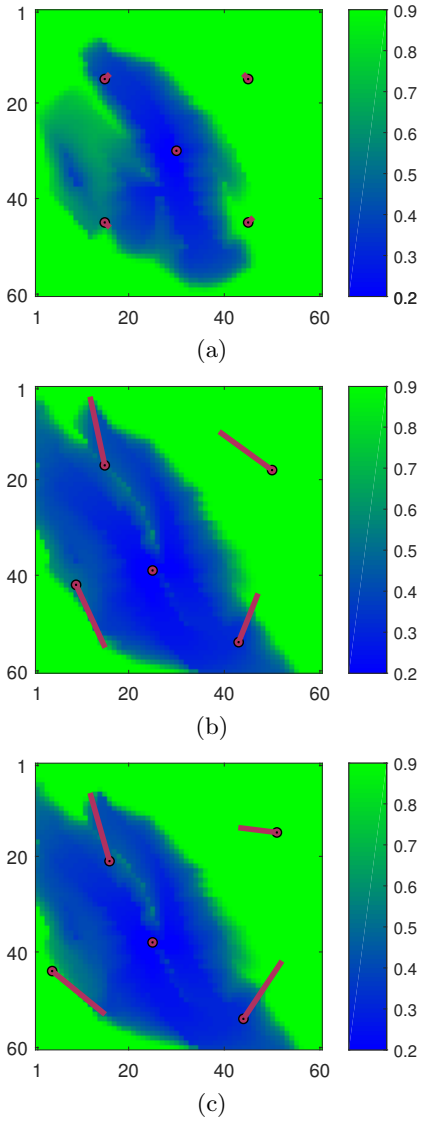


Figure 8: Case 3: Oil (green) and water (blue) saturations at the end of production time corresponding to (a) initial well configuration and final well configuration for (b) $\varepsilon = 15$m and (c) $\varepsilon = 24$m. The well locations are marked by pink trajectories.

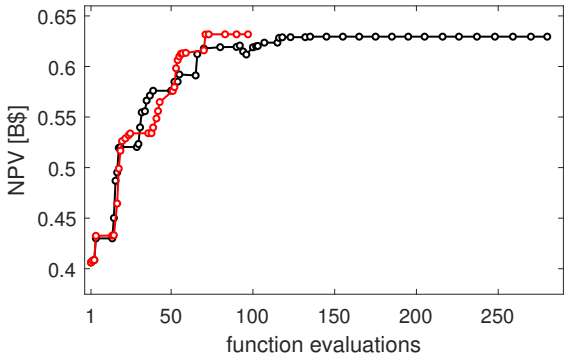


Figure 9: Case 3: Cost function evolution including the function evaluations during line search. Circles mark the gradient evaluations requested by the SQP solver for $\varepsilon = 15$m (blue line) and $\varepsilon = 24$m (black line).

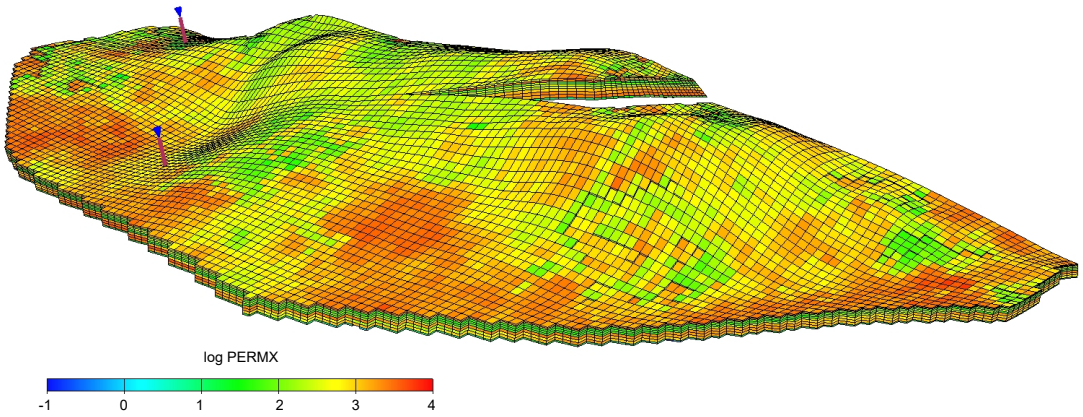
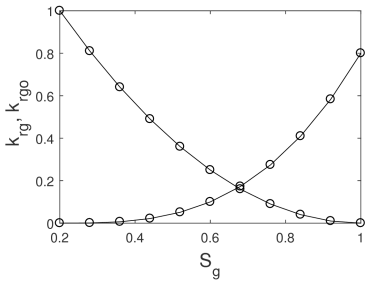
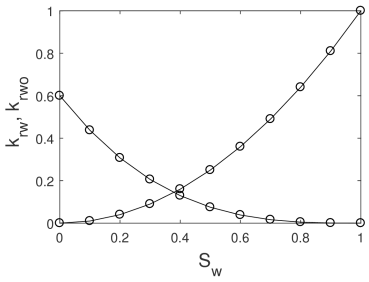


Figure 10: Case 4: Permeability in x-direction and two injectors whose locations are kept fixed.



(a)



(b)

Figure 11: Relative permeability for (a) gas-oil and (b) water-oil systems in Case 3.

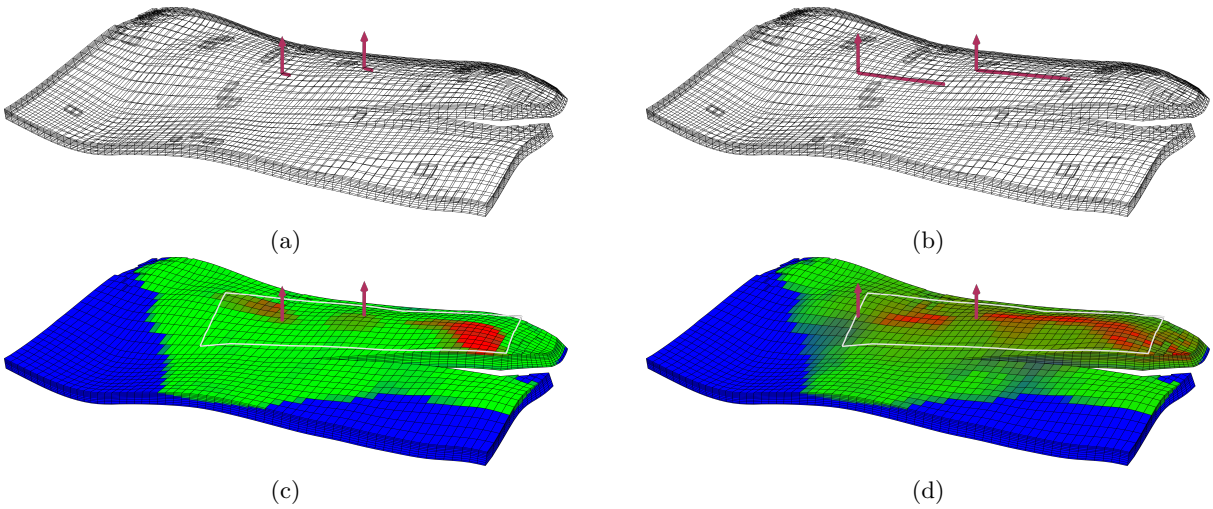


Figure 12: Case 4: (a) Initial well configuration. (b) Final well configuration. Oil (green), gas (red) and water (blue) saturations at end of production time corresponding to (c) initial well configuration and (d) final well configuration. The location of the two horizontal wells subject to optimization are marked by pink wellheads and trajectories. Upper and lower bounds for x and y well coordinates are shown as white border.

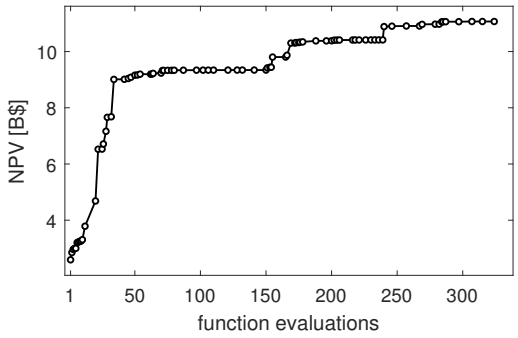


Figure 13: Case 4: Cost function evolution including the function evaluations during line search. Circles mark the gradient evaluations requested by the SQP solver.