Comput Geosci (2017) 21:47–74
DOI 10.1007/s10596-016-9594-z

ORIGINAL PAPER

A unified formulation for generalized oilfield development optimization

Shiva Navabi[1] · **Reza Khaninezhad**[1] · **Behnam Jafarpour**[1]

Received: 7 March 2016 / Accepted: 27 September 2016 / Published online: 22 November 2016

Abstract Oilfield development involves several key decisions, including the number, type (injection/production), location, drilling schedule, and operating control trajectories of the wells. Without considering the coupling between these decision variables, any optimization problem formulation is bound to find suboptimal solutions. This paper presents a unified formulation for oilfield development optimization that seeks to simultaneously optimize these decision variables. We show that the source/sink term of the governing multiphase flow equations includes all the above decision variables. This insight leads to a novel and unified formulation of the field development optimization problem that considers the source/sink term in reservoir simulation equations as optimization decision variables. Therefore, a single optimization problem is formulated to simultaneously search for optimal decision variables by determining the complete dynamic form of the source/sink terms. The optimization objective function is the project net present value (NPV), which involves discounted revenue from oil production, operating costs (e.g. water injection and recycling), and capital costs (e.g., cost of drilling wells). A major difficulty after formulating the generalized field development optimization problem is finding an efficient solution approach. Since the total number of cells in a reservoir model far exceeds the number of cells that are intersected by wells, the source/sink terms tend to be sparse. In fact, the drilling cost in the NPV objective function serves as a sparsity-promoting penalty to minimize the number of wells while maximizing the NPV. Inspired by this insight, we solve the optimization problem using an efficient gradient-based method based on recent algorithmic developments in sparse reconstruction literature. The gradients of the NPV function with respect to the source/sink terms is readily computed using well-established adjoint methods. Numerical experiments are presented to evaluate the feasibility and performance of the generalized field development formulation for simultaneous optimization of the number, location, type, controls, and drilling schedule of the wells.

Keywords Field development optimization · Well placement · Well control optimization · Drilling schedule · Sparsity-promoting solution

1 Introduction

Traditionally, field development planning, including well locations and their control settings, has been accomplished by integrating geologic and geophysical data with expert knowledge as well as practical considerations and constraints. The advent of reservoir simulation technology introduced a predictive tool to evaluate and support/reject proposed development plans. With increased complexity of models that are developed for simulation, the choice of a sound development strategy becomes too difficult to pin down. In many cases, to identify an optimal development plan, an intractable number of strategies may have to be evaluated, making the process computationally impractical, if not infeasible. A more systematic and efficient approach to field development is using numerical optimization in

✉ Behnam Jafarpour
jafarpou@usc.edu

[1] University of Southern California, 925 Bloom Walk, HED 313, Los Angeles CA, USA

which a proposed development strategy is gradually updated and improved based on well-established optimization techniques.

Several methods have been proposed and implemented for solving a host of field development optimization problems in the past. Some of the important development decisions include the type (injection/production), number, and location of the wells, as well as their dynamic control settings, and their drilling schedule. The majority of the existing algorithms consider separate (independent) optimization problems for each individual decision variables without accounting for the interplay among them. Optimizing an individual variable type (e.g. well locations) while disregarding its coupling with other variable types (e.g. well controls) leads to solutions that are inherently suboptimal as they do not exploit the full range of variability in the disregarded decision variables [32]. In other words, fixing the values of other variables when optimizing a certain variable introduces strong constraints that often result in significant performance loss in the optimization problem. Examples of approaches that have been used to optimize individual decision variables are well placement optimization to identify the optimal well locations given predetermined and fixed well controls [3, 19, 36, 41, 58], and well control optimization in which the operational settings of the wells (rates and pressures) are optimized for a fixed well configuration [1, 9, 21, 47].

Both gradient-based local search methods and gradient-free global search algorithms have been applied to field development optimization problems. Gradient-based methods are computationally efficient and monotonically improve the objective function. However, they easily get trapped in local solutions and are sensitive to initialization. In addition, gradient calculation either requires development of complex adjoint models or numerical approximation techniques that are computationally demanding. Global search methods are a class of optimization techniques that do not require gradient information. Some of these methods are based on evolutionary programming techniques that adopt a population-based search mechanism to explore the space of feasible solutions. There are a number of well-known algorithms that belong to this category of heuristic methods that are suitable for problems with objective functions that have complex response surfaces. Examples of application of this class of methods to field development optimization include [7, 12, 22, 26, 27, 30, 36]. In general, these methods are easy to implement and do not have the requirements of gradient-based optimization methods. However, they are heuristic in nature and are computationally demanding, making them prohibitive for large-scale applications.

Recent studies in the literature have attempted to include multiple variable types in the field development optimization problem. As an initial step towards such integrative solution approaches, Li and Jafarpour [32] proposed a sequential optimization scheme in which a coupled well placement and well control problem is decoupled and solved sequentially. In this approach, well locations and their control settings are optimized separately; however, the solution of each optimization problem is used to re-initialize the other optimization problem. These sequential optimizations are repeated until no further improvement in the NPV objective function is observed. The results in Li and Jafarpour [32] show significant improvement in the NPV values when the two optimizations are combined. Li et al. [33] applied a generalized version of the SPSA algorithm [50, 51] to simultaneously optimize well locations and controls and also reported major improvement in production performance. Humphries et al. [28] also propose a sequential scheme by decoupling well placement and control optimization problems and report possible improvement over the joint optimization approach. They hypothesize that this might be due to judiciously chosen control solutions that are fixed when the well placement problem is solved; they then argue that a proper control initialization leads to a reduction in the solution space to be explored during well placement optimization which can increase the likelihood of finding the globally optimal well locations. While the local nature of the optimization algorithm can result in scenarios where sequential solutions outperform coupled solutions, those outcomes cannot be generalized. Theoretically, sequential methods provide approximate solutions to joint optimization problems that are typically applied to reduce complexity and improve computational efficiency (see Shu et al. [49] for a discussion). Other methods have also been proposed to solve joint field development optimization problems [24, 27, 29]. Isebor et al. [30] propose a hybrid technique based on particle swarm optimization (PSO) to solve the well placement and well control optimization problems simultaneously. Humphries et al. [27] combine the PSO algorithm with the generalized pattern search (GPA) technique to jointly optimize locations and controls of the wells. They reported improved solutions when the joint problem is decoupled into consecutive stages of well placement and control optimization and solved sequentially. In Forouzanfar and Reynolds [24], a formulation is proposed for simultaneously optimizing the number, location and rate allocations of the wells. Their solution approach is initialized by assigning a well to each cell in the reservoir domain and improving the NPV by iteratively eliminating non-profitable wells using a gradient-based algorithm. Unlike our proposed framework, the formulation in Forouzanfar

and Reynolds [24] does not include well control trajectories and assumes that all the wells are drilled initially and have time-invariant controls. In our approach, a discounted capital cost of drilling wells is included in the objective function that enables drilling schedule optimization, by favoring later drilling times.

While the previous studies have taken different approaches to include additional decision variables in the optimization problem, a fully generalized field development optimization formulation that seamlessly incorporates all the important decision variables into a single optimization problem has not yet been presented. In this paper, we show that the entire decision variables of the generalized field development problem are collectively captured in the source/sink term of the discretized governing flow equations. Inspired by this insight, we formulate a novel approach for the generalized field development optimization problem as a continuous optimization problem. Since such a generalized problem involves several coupled variables, finding an efficient solution approach poses another challenge. Noting that the total number of cells in a reservoir model far exceeds the number of cells that are intersected by wells, the source/sink terms in a model are extremely sparse (contain many zeros that correspond to cells that do not contain any wells). As discussed in the next part, this implies that the solution of the resulting generalized field development optimization problem is sparse. In fact, we show that in maximizing the NPV objective function, the well drilling cost essentially serves as a sparsity-promoting penalty to minimize the number of wells. The inherently sparse nature of the decision variables in the proposed generalized field development optimization formulation inspires the use of sparsity-promoting optimization methods. Using recent algorithmic developments in sparse reconstruction literature [4, 11, 20], we apply an efficient gradient-based algorithm to solve the resulting optimization problem.

The remainder of this paper is organized as follows. Section 2 presents the subsurface flow governing equations to illustrate that all the decision variables of interest in a generalized field development optimization are indeed captured by the source/sink terms; a discussion on the sparse nature of the source/sink terms in the discretized form of the flow equations is presented next, followed by the formulation of the generalized field development optimization problem. The similarities between the solution of the generalized field development optimization problem and sparse reconstruction techniques is then exploited to develop a sparse optimization algorithm to solve it. Section 3 presents a set of numerical experiments to evaluate the performance of the proposed solution approach. The paper is concluded in Section 4 with a discussion on the implications of this work and its possible future extensions.

2 Generalized field development optimization

We motivate the problem formulation by presenting the governing equations of multiphase flow in the subsurface environment [31] and discussing the source/sink terms as decision variables of a generalized field development optimization formulation. The proposed formulation is presented after introducing and discussing the NPV objective function and its sparsity-promoting property.

2.1 Governing flow equations

In this section, to introduce the decision variables of our generalized field development optimization, we briefly present the governing flow equations for a two-phase (oil-water) slightly-compressible 2-dimensional flow system. Our presentation follows that of Jansen [31]. For simplicity, and without loss of generality, we only consider a system with constant compressibilities, ignore capillary pressure and gravity forces. With these assumptions, the governing equations for phase n (n = oil,water) can be expressed as

$$\underbrace{-\frac{h}{\mu_n}\left[\frac{\partial}{\partial x}\left(kk_{rn}\frac{\partial p}{\partial x}\right)+\frac{\partial}{\partial y}\left(kk_{rn}\frac{\partial p}{\partial y}\right)\right]}_{\text{flux term}}$$
$$+\underbrace{h\left[\phi S_n(c_n+c_r)\frac{\partial p}{\partial t}+\phi\frac{\partial S_n}{\partial t}\right]}_{\text{accumulation term}}-\underbrace{hq_n'''}_{\text{source term}}=0 \tag{1}$$

From the physical saturation constraint, we have $S_w + S_o = 1$. The finite difference discretized version of the above equations can be written as

$$\begin{aligned}&V\left[\phi S_n(c_n+c_r)\frac{\partial p}{\partial t}+\phi\frac{\partial S_n}{\partial t}\right]_{i,j}-(T_n)_{i-\frac{1}{2},j}\ p_{i-1,j}\\&-(T_n)_{i,j-\frac{1}{2}}\ p_{i,j-1}+\Big[(T_n)_{i-\frac{1}{2},j}+(T_n)_{i,j-\frac{1}{2}}\\&+(T_n)_{i,j+\frac{1}{2}}+(T_n)_{i+\frac{1}{2},j}\Big]p_{i,j}-(T_n)_{i,j+\frac{1}{2}}p_{i,j+1}\\&-(T_n)_{i+\frac{1}{2},j}p_{i+1,j}=V(q_n''')_{i,j}\end{aligned} \tag{2}$$

where, the discretized transmissibilities are defined as

$$(T_n)_{i,j}\triangleq\frac{\Delta y}{\Delta x}\frac{h}{\mu_n}(kk_{rn})_{i,j} \tag{3}$$

The matrix form of the resulting discretized equations can be expressed as

$$\underbrace{\begin{bmatrix} \mathbf{V}_{wp} & \mathbf{V}_{ws} \\ \mathbf{V}_{op} & \mathbf{V}_{os} \end{bmatrix} \begin{bmatrix} \dot{\mathbf{p}} \\ \dot{\mathbf{s}} \end{bmatrix}}_{\text{accumulation term}} + \underbrace{\begin{bmatrix} \mathbf{T}_w & 0 \\ \mathbf{T}_o & 0 \end{bmatrix} \begin{bmatrix} \mathbf{p} \\ \mathbf{s} \end{bmatrix}}_{\text{flux term}} = \underbrace{\begin{bmatrix} \mathbf{q}_w \\ \mathbf{q}_o \end{bmatrix}}_{\text{source term}} \tag{4}$$

In Eq. 4, the saturation and pressure values in all cells are denoted by the vectors $\mathbf{p}$ and $\mathbf{s}$. Vectors $\mathbf{q}_w$ and $\mathbf{q}_o$ in the source/sink term of Eq. 4 denote, respectively, the flow rates of water and oil phases in all the cells with their entries arranged in the following format

$$\begin{aligned} \mathbf{q}_w^T &\triangleq [\, \cdots \; (q_w)_{i,j} \; \cdots \,] \\ \mathbf{q}_o^T &\triangleq [\, \cdots \; (q_o)_{i,j} \; \cdots \,] \end{aligned} \tag{5}$$

For any given time, the entries of the source/sink vector are zero for cells that do not contain a well. Hence, the source/sink vectors in Eq. 4 contain the information about well locations, types, and control settings. Therefore, these vectors constitute the decision variables of the generalized field development optimization problem as described in the sequel. Figure 1 shows the correspondence between the well location in a typical reservoir and the entries of the source/sink term. As illustrated in Fig. 1, in a typical oilfield, only a small number of grid blocks are intersected by wells and, thus, the vast majority of entries in the source/sink vector ($\mathbf{q}^T = [\mathbf{q}_w^T \;\; \mathbf{q}_o^T]$) are 0s; that is, $\mathbf{q}$ is sparse. This interpretation highlights the similarity between the field development optimization problem and sparse reconstruction, where a sparse solution to an optimization problem is sought. The optimal solution of the source/sink vector should minimize the number of wells while maximizing the production or the net present value (NPV). Solution sparsity is promoted due to the presence and significance of the drilling cost in the objective function. Note that the entries in the source/sink vector ($[\mathbf{q}_w^T \;\; \mathbf{q}_o^T]^T$) in Eq. 4 are expressed for a single time step (Fig. 2a). In practice, wells are controlled dynamically and have time-varying control trajectories. Typically, the control trajectory is discretized into a set of control steps for which the control settings are optimized. Examples of such multistep control trajectories are displayed in Fig. 2b.

Concatenation of the source/sink vectors corresponding to successive time steps leads to a representation that incorporates temporal variation in well controls. As illustrated in Fig. 2c, in addition to location, type, number and time-varying control values of the operating wells, from this spatiotemporal description one can also deduce the drilling schedule. The temporal resolution of the drilling schedule depends on the selected control time steps, with smaller control time steps resulting in more frequent changes in development plans (control adjustment and drilling schedule). To better illustrate the dynamic well controls and drilling schedule, we expand the source/sink *vector* into a source/sink *matrix* $\mathbf{Q}$ as follows (note that we use a matrix form to facilitate the discussion, but for optimization purpose this matrix is vectorized).

$$\mathbf{Q} = \begin{bmatrix} q_1^{t_1} & q_1^{t_2} & \cdots & q_1^{t_T} \\ q_2^{t_1} & q_2^{t_2} & \cdots & q_2^{t_T} \\ \vdots & \vdots & \ddots & \vdots \\ q_N^{t_1} & q_N^{t_2} & \cdots & q_N^{t_T} \end{bmatrix} \tag{6}$$

The rows of $\mathbf{Q}$ correspond to the indices of the grid blocks (that can contain wells) and its columns consist of well control trajectories in time (Fig. 3). In Eq. 6, T denotes

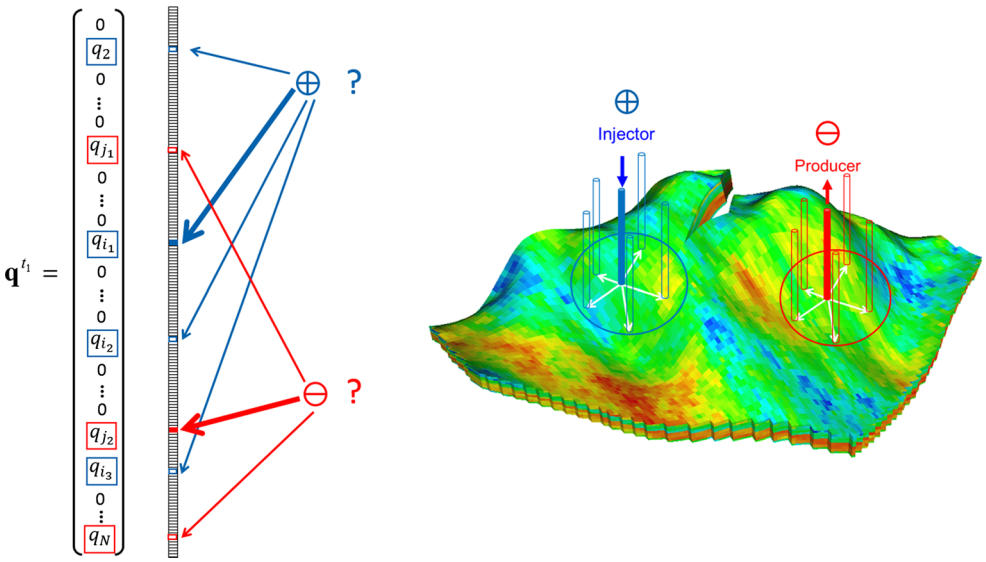


Fig. 1 Source/sink vector as the main decision variable of the generalized field development optimization problem

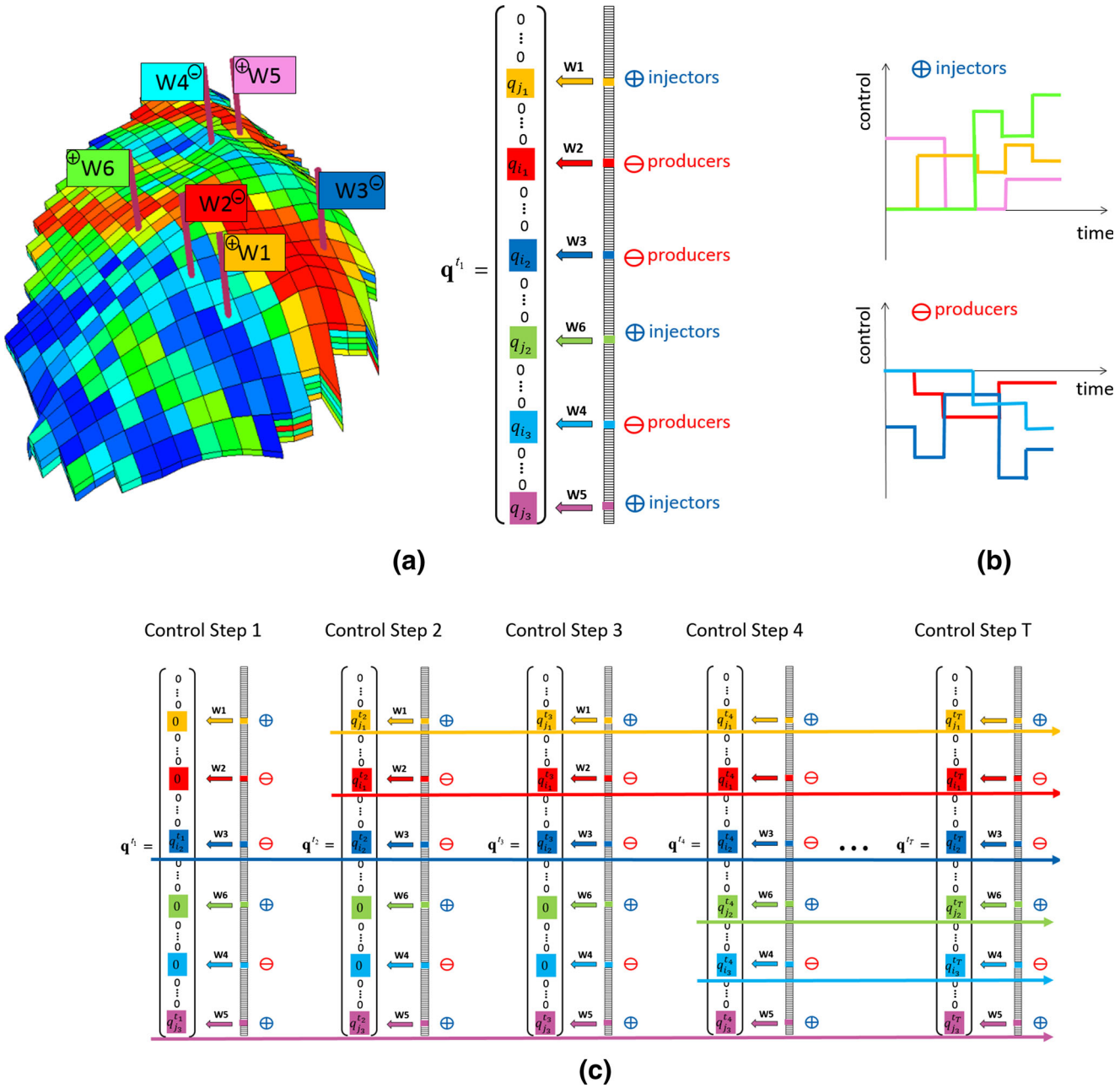


Fig. 2 **a** Source/sink vector for single control step problem; **b** Time-varying control trajectories for multiple wells; **c** Concatenation of source/sink vectors to incorporate information about dynamic well controls and drilling schedule

the total number of control steps and N is the total number of grid blocks in the discretized domain. The entry $q_i^{t_j}$ indicates the *total* fluid flow rate of the well located in grid block i at control time step t_j. The well type is determined by the *sign* of each entry in $\mathbf{Q}$; by convention, injector $(+)$ and producer $(-)$. It is important to note that the sparse nature of the solution does not depend on the type of well controls used (total fluid flow rates or bottom-hole pressure). While the above equations use rate-controlled injection and production wells, the formulation can be generalized to problems that include a mixture of BHP-controlled and rate-controlled wells.

The matrix form of the two-phase flow equations can be described as

$$\underbrace{\begin{bmatrix} \mathbf{V}_{wp}(s) & \mathbf{V}_{ws} \\ \mathbf{V}_{op}(s) & \mathbf{V}_{os} \end{bmatrix} \begin{bmatrix} \dot{\mathbf{p}} \\ \dot{\mathbf{s}} \end{bmatrix}}_{\text{accumulation term}} + \underbrace{\begin{bmatrix} \mathbf{T}_w(s) & 0 \\ \mathbf{T}_o(s) & 0 \end{bmatrix} \begin{bmatrix} \mathbf{p} \\ \mathbf{s} \end{bmatrix}}_{\text{flux term}} = \underbrace{\begin{bmatrix} \mathbf{F}_w(s) \\ \mathbf{F}_o(s) \end{bmatrix}}_{\text{source term}} \mathbf{q}_t \tag{7}$$

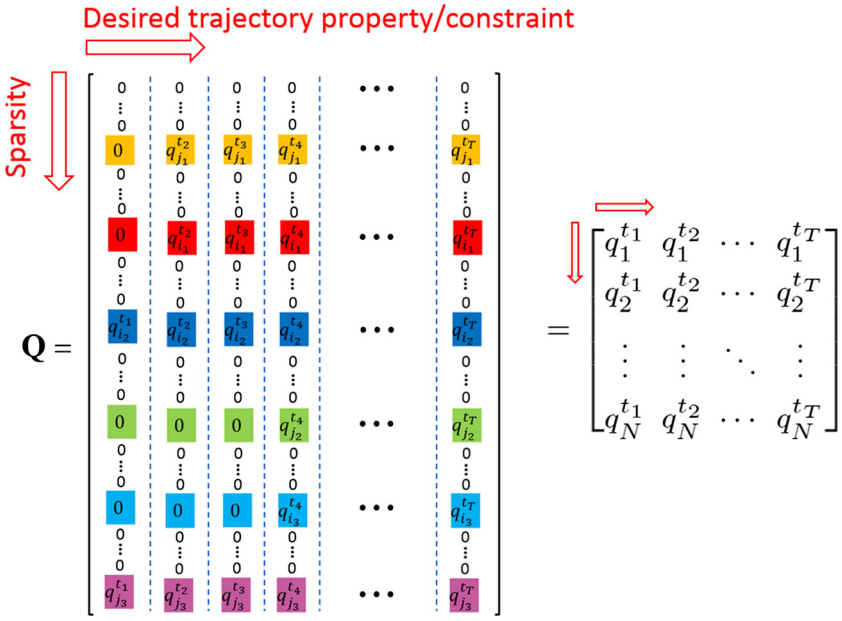


Fig. 3 Structure of the controls allocation matrix; sparsity is imposed along the rows while variation of dynamic control values is regulated by posing proper constraints along the columns

where $\mathbf{q}_t$ represents the total fluid flow rate and $\mathbf{F}_o$ and $\mathbf{F}_w$ are diagonal matrices whose entries contain fractional flows and are, thus, mostly zeros, except for the ones corresponding to the grid cells that contain a well:

$$\begin{aligned} \mathbf{F}_w &\triangleq [0 \cdots 0\ (f_w)_{i,j}\ 0 \cdots 0] \\ \mathbf{F}_o &\triangleq [0 \cdots 0\ (f_o)_{i,j}\ 0 \cdots 0] \end{aligned} \tag{8}$$

where,

$$f_o = \frac{\lambda_o}{\lambda_o + \lambda_w} \quad , \quad f_w = \frac{\lambda_w}{\lambda_o + \lambda_w} \tag{9}$$

in which λ_o and λ_w represent oil and water mobilities, respectively. The proportion of oil and water in the total fluid flow rate q_t can then be computed as:

$$q_o = f_o q_t \quad , \quad q_w = f_w q_t \tag{10}$$

The general form of the matrix representation for the two-phase flow, where some wells have total fluid flow rate controls and some are controlled by prescribed BHBs, can be written in the following partitioned format [31, Chapter 2]:

$$\underbrace{\left[\begin{array}{ccc|ccc} \mathbf{V}_{wp,11} & \mathbf{0} & \mathbf{0} & \mathbf{V}_{ws,11} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_{wp,22} & \mathbf{0} & \mathbf{0} & \mathbf{V}_{ws,22} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{V}_{wp,33} & \mathbf{0} & \mathbf{0} & \mathbf{V}_{ws,33} \\ \hline \mathbf{V}_{op,11} & \mathbf{0} & \mathbf{0} & \mathbf{V}_{os,11} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_{op,22} & \mathbf{0} & \mathbf{0} & \mathbf{V}_{os,22} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{V}_{op,33} & \mathbf{0} & \mathbf{0} & \mathbf{V}_{os,33} \end{array}\right] \left[\begin{array}{c} \dot{\mathbf{p}}_1 \\ \dot{\mathbf{p}}_2 \\ \dot{\mathbf{p}}_3 \\ \hline \dot{\mathbf{s}}_1 \\ \dot{\mathbf{s}}_2 \\ \dot{\mathbf{s}}_3 \end{array}\right]}_{\text{accumulation term}}$$

$$+ \underbrace{\left[\begin{array}{ccc|ccc} \mathbf{T}_{w,11} & \mathbf{T}_{w,12} & \mathbf{T}_{w,13} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{T}_{w,21} & \mathbf{T}_{w,22} & \mathbf{T}_{w,23} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{T}_{w,31} & \mathbf{T}_{w,32} & \mathbf{T}_{w,33} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{T}_{o,11} & \mathbf{T}_{o,12} & \mathbf{T}_{o,13} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{T}_{o,21} & \mathbf{T}_{o,22} & \mathbf{T}_{o,23} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{T}_{o,31} & \mathbf{T}_{o,32} & \mathbf{T}_{o,33} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{array}\right] \left[\begin{array}{c} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \\ \hline \mathbf{s}_1 \\ \mathbf{s}_2 \\ \mathbf{s}_3 \end{array}\right]}_{\text{flux term}}$$

$$= \underbrace{\left[\begin{array}{ccc|ccc} \mathbf{0} & \mathbf{0} & \mathbf{0} & & & \\ \mathbf{0} & \mathbf{F}_{w,22} & \mathbf{0} & & & \\ \mathbf{0} & \mathbf{0} & \mathbf{F}_{w,33} & & & \\ \hline & & & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ & & & \mathbf{0} & \mathbf{F}_{o,22} & \mathbf{0} \\ & & & \mathbf{0} & \mathbf{0} & \mathbf{F}_{o,33} \end{array}\right] \left[\begin{array}{c} \mathbf{0} \\ \breve{\mathbf{q}}_{well,t} \\ \mathbf{J}_3\,(\breve{\mathbf{p}}_{well} - \mathbf{p}_3) \end{array}\right]}_{\text{source term}} \tag{11}$$

Here, $\mathbf{p}_1$ is a vector containing the pressure values in grid blocks that do not contain a well, $\mathbf{p}_2$ denotes a vector whose entries are the pressure values in well grid blocks that intersect a well controlled by prescribed total fluid flow rate (and, hence, have $\breve{\mathbf{q}}_{well,t}$ as their source term) while $\mathbf{p}_3$ vector has pressure values in cells with BHP controlled wells (with the corresponding source/sink term $\mathbf{J}_3$ ($\breve{\mathbf{p}}_{well} - \mathbf{p}_3$)). The oil and water flow rate proportions are obtained from Eq. 10. In multi-layer formations q_o is the sum of oil flow rates in all the grid cells perforated by the production well: $q_o = \sum_{j\in\mathcal{I}} q_{o,j}$ where $\mathcal{I}$ is the set of indices of the perforated grid blocks. The notation $\breve{\mathbf{p}}_{well}$ refers to the vector whose entries are pressure values in BHP-controlled wells, where $\mathbf{J}_3$ is the diagonal matrix with well productivity indices as its diagonal entries.

In Eq. 11, very few entries of the source/sink term are non-zeros. In practice, only a small subset of grid cells in a reservoir model are intersected with wells, resulting in the vector representing the source/sink terms to be sparse. This practical observation is also supported by the sparsity-promoting nature of the well drilling cost term (l_0 norm of the source and sink vector). The matrix form in Eq. 6 can be generalized to a multistep control matrix $\mathbf{U}$ with a mixture of pressure (BHP) and rate controlled modes as follows:

$$\mathbf{U} = \begin{bmatrix} q_1^{t_1} & q_1^{t_2} & \cdots & q_1^{t_T} \\ p_2^{t_1} & p_2^{t_2} & \cdots & p_2^{t_T} \\ \vdots & \vdots & \vdots & \vdots \\ q_k^{t_1} & q_k^{t_2} & \cdots & q_k^{t_T} \\ \vdots & \vdots & \vdots & \vdots \\ p_m^{t_1} & p_m^{t_2} & \cdots & p_m^{t_T} \\ \vdots & \vdots & \ddots & \vdots \\ p_N^{t_1} & p_N^{t_2} & \cdots & p_N^{t_T} \end{bmatrix} \tag{12}$$

where the m^{th} grid block contains a well controlled by BHP control ($p_m^{t_j}$) and the k^{th} grid cell is penetrated by a rate-controlled well with total rate ($q_k^{t_j}$). The above matrix offers a simple intuition where the rows represent the grid cells and their associated source/sink terms, and the columns have the control steps (trajectory) in time. For optimization, $\mathbf{U}$ is vectorized to form a vector of decision variables $\mathbf{u} = [\mathbf{U}_{1,-}, \mathbf{U}_{2,-}, \ldots, \mathbf{U}_{N,-}]^T$.

Considering the rate-controlled problem with the rate allocations matrix $\mathbf{U}$, the drilling cost applies a sparsity-promoting penalty to the rows of $\mathbf{U}$ in Eq. 6 to minimize the number of wells to drill. Furthermore, other constraints may be imposed along the columns of $\mathbf{U}$ to impart a desirable behavior on the control trajectory. Such constraints may be enforced to, for instance, minimize a particular norm of the control trajectory (total water injected) or impose certain structure on the control solution [2]. Next, we formulate the generalized field development optimization problem.

2.2 Optimization formulation

The multivariate optimization problem formulation for the generalized field development can be expressed as

$$\begin{aligned} \mathbf{u}^* &= \arg\min \quad J(\mathbf{u}) \\ \textit{subject to} \quad & \mathbf{u}_{min} \le \mathbf{u} \le \mathbf{u}_{max} \\ & g_i(\mathbf{u}) = 0 \quad , \quad i = 1, 2, \cdots, n \\ & f_j(\mathbf{u}) \le 0 \quad , \quad j = 1, 2, \cdots, m \end{aligned} \tag{13}$$

The decision variables are of dimension $N \times T$. In general, the equality $g_i(.)$ and inequality $f_j(.)$ constraints are problem-specific and user-defined. A general equality constraint is the mass balance equations, which is satisfied by running the flow simulation to compute the objective function. Typical inequality constraints include economic water-cut or bubble point pressure constraints. A common objective function in field development optimization is the net present value of the project. The NPV function for a reservoir asset typically incorporates the discounted version of production revenues and the operating and capital costs, mathematically expressed as:

$$\begin{aligned} NPV =& O(\mathbf{u}) - C(\mathbf{u}) \\ =& \int_0^T \Bigg[\underbrace{\sum_{i=1}^{N_{\text{prod}}} r_o(t)\, q_{o,i}(t)}_{\text{oil revenue}} - \underbrace{\sum_{i=1}^{N_{\text{prod}}} r_{w,disp}(t)\, q_{w,disp,i}(t)}_{\text{water disposal cost}} \\ & \underbrace{- \sum_{i=1}^{N_{\text{inj}}} r_{w,inj}(t)\, q_{w,inj,i}(t)}_{\text{water injection cost}} \Bigg] \underbrace{\frac{1}{(1+d(t))^t}}_{\text{discount factor}} dt \\ & - \sum_{i=1}^{N_{\text{well}}} \frac{\gamma}{(1+d(t))^{b_i}}. \end{aligned} \tag{14}$$

Note that in Eq. 14, the production quantities depend on the decision variables. The integral term represents the operating costs, $C(\mathbf{u})$ (last sum) denotes the drilling cost, N_{prod} and N_{inj} are the number of production and injection wells, respectively, and $r_o(t)$, $r_{w,disp}(t)$ and $r_{w,inj}(t)$ stand for the price of oil, the cost of water disposal/recycling and the injection cost ($/stb), respectively. The annual discount factor is $d(t)$, which is applied to both operating and capital

expenditures, and γ denotes the cost of drilling a single well ($). Since the number of wells and their drilling times are unknown, the objective function includes an associated capital cost, which is written separately (last term). Note that the objective function is particularly sensitive to the number of wells and the drilling times due to the high cost associated with drilling.

To proceed, we seek to express the capital cost as a function of the decision variables. Because the drilling cost is not a function of well control values, we present a manipulation of the decision variables to express this cost in terms of the decision variables. To mathematically express the drilling cost in terms of $\mathbf{U}$, consider a variant of matrix $\mathbf{U}$ in which for each row (possible well), all entries after the *first* nonzero element, which represents the introduction of a new well, are zeroed out. The resulting matrix, $\mathbf{U}_C$, can be obtained from a linear transformation of $\mathbf{U}$:

$$\mathbf{U}_C \coloneqq \mathbf{A}\mathbf{U} \tag{15}$$

where $\mathbf{A}$ is an N^2 dimensional matrix that can be expressed as:

$$\mathbf{A} = \mathbf{U}_C\mathbf{U}^T(\mathbf{U}\mathbf{U}^T)^{-1} \tag{16}$$

In Eq. 16, the term $\mathbf{U}\mathbf{U}^T$ is not guaranteed to be invertible. Various numerical methods and preconditioning techniques exist for *numerical* solution of this system [46]. The number of nonzero entries in each column of $\mathbf{U}_C$, equals the number of wells that are drilled at the control time step corresponding to that column index.

It should be noted that Eqs. 15–16 are written based on the assumption that $\mathbf{U}_C$ is known. Mathematically expressing $\mathbf{U}_C$ in terms of $\mathbf{U}$ allows for convenient computation of the drilling time in terms of $\mathbf{U}$. An approximate *linear* transformation for obtaining $\mathbf{U}_C$ from $\mathbf{U}$ (15–16) is also a straightforward way to facilitate the computation of gradients of the drilling cost term needed for optimization. To illustrate these steps, let us consider a simple example where the control allocation matrix $\mathbf{U}$ is given as:

$$\mathbf{U} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -0.8 \\ 0 & 0 & -0.2 \end{bmatrix} \tag{17}$$

From the above expression, it can be inferred that Wells 1, 2 and 3 are drilled in control time steps 1, 2, and 3, respectively. The resulting matrix $\mathbf{U}_C$ will have, in each row of it, zero entries except for the *first non-zero* entry. In this case, we will have:

$$\mathbf{U}_C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -0.2 \end{bmatrix} \tag{18}$$

$\mathbf{A}$ can be reconstructed from Eq. 15 as:

$$\mathbf{A} = \mathbf{U}_C\mathbf{U}^T(\mathbf{U}\mathbf{U}^T)^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{bmatrix} \tag{19}$$

It can now be verified that $\mathbf{A}\mathbf{U} = \mathbf{U}_C$. Note that the rows of $\mathbf{U}_C$ contain the drilling time information, which will be used for discounting the capital cost. Hence, the discounted capital cost can be expressed in terms of $\mathbf{U}_C$ as:

$$C(\mathbf{U}) = \sum_{t=1}^{T} \frac{\gamma}{(1+d(t))^{b(t)}} \parallel \mathbf{U}_C^t \parallel_0 \tag{20}$$

where $\mathbf{U}_C^t$ denotes the t^{th} column of $\mathbf{U}_C$ and t enumerates the column indices of $\mathbf{U}_C$ (control time steps); $b(t)$ is the number of years prior to wells counted by $\parallel \mathbf{U}_C^t \parallel_0$ coming online. Note that the discount factor in the objective function controls the drilling schedule by giving preference to wells that are drilled later. The zero "norm" (ℓ_0) of the source/sink terms at each time step (that is, $\parallel \mathbf{U}_C^t \parallel_0$) counts the number of non-zero entries within the t^{th} column of $\mathbf{U}_C$, which refers to the number of wells coming online at that time step. We also note that the ℓ_0-"norm" is not a true norm, but in the literature it is loosely referred to as a norm.

The above drilling cost behaves as a sparsity-promoting penalty that encourages solutions with minimum number of wells (drilling cost) and delayed drilling while maximizing the NPV. An extensive literature exists on solving such sparse optimization problems [52, 57, 61]. From an optimization perspective, the (ℓ_0) function is discontinuous and the exact solution of the resulting optimization becomes combinatorial and NP-hard. However, several (heuristic) approximate solution methods exist for practical applications. A common class of techniques involves approximating the (ℓ_0) function with a better-behaved function. An approximate continuous function, with a similar behavior is (see the ℓ_r-norm behavior for different values of r in Fig. 4a)

$$\hat{C}(\mathbf{U}) = \sum_{t=1}^{T} \frac{\gamma}{(1+d(t))^{b(t)}} \parallel \mathbf{U}_C^t \parallel_{r\to 0^+} \tag{21}$$

For values $0 < r \leq 1$, the ℓ_r norm function has sparsity-promoting property, which has been extensively exploited in solving ℓ_0-"norm" problems. For sparsity along the rows of $\mathbf{U}$, a sufficiently small r (i.e. $0 < r \ll 1$) has been shown [13, 14] to have a similar sparsity-promoting behavior as ℓ_0-"norm". However, given that the well cost will be affected for any norm other than ℓ_0, it is imperative to ensure that at the solution the well cost is properly reflected by using a very small r value (e.g., $r < .01$).

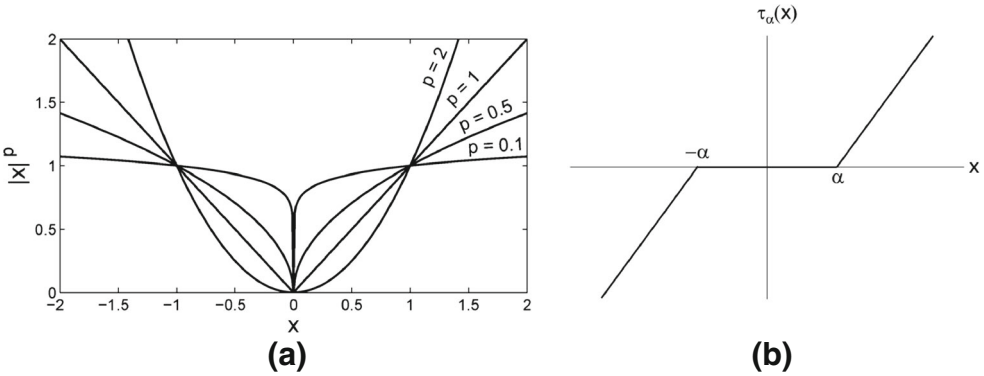


Fig. 4 **a** Illustration of the ℓ_p norm functions for different values of p; **b** the shrinkage function

While the above norm can effectively represent the drilling cost, in many cases it is also desirable to constrain the control trajectory (columns of $\mathbf{U}$). For instance, the total amount of injection can either be fixed or minimized. For practical reasons, it may be desirable to penalize abrupt or frequent changes in the control trajectory. Standard vector norm penalty functions (such as roughness penalty) may be used to achieve such goals. An alternative representation of the drilling cost function, as proposed in Navabi et al. [37], is obtained by using a mixed $\ell_{p,r}$ norm which applies a ℓ_p , $p > 1$ norm along the columns of $\mathbf{U}$ and a $\ell_r, 0 < r \ll 1$ norm along its rows. In that case, the capital (drilling) cost $C(\mathbf{U})$ can be expressed as:

$$C(\mathbf{U}) = \Big[\sum_{i=1}^{N} \frac{\gamma}{(1+d(t))^t}\Big(\epsilon + \sum_{t=1}^{T} |q_i^t|^p\Big)^{\frac{r}{p}}\Big]^{\frac{1}{r}} \simeq \| \mathbf{U} \|_{\Lambda;p,r} \tag{22}$$

The term ($\| \mathbf{U} \|_{\Lambda;p,r}$) in Eq. 22 is a *weighted mixed norm* of $\mathbf{U}$ that accounts for the drilling costs and promotes sparsity along the rows of $\mathbf{U}$. Since the entries of the nonzero rows are control allocations of the corresponding wells in time, p affects the behavior of the control trajectory. In this case, setting $p > 1$ prevents dynamic control trajectories from being sparse in time (wells being shut-off). Another effect that ensues from setting $p > 1$ is avoiding large and sudden injection/production controls; the larger the value of $p > 1$, the less likely it is for wells to have very large operation rates. The ℓ_r norm that represents the drilling cost, on the other hand, is applied to the rows of $\mathbf{U}$ to promotes sparsity (i.e. minimize the number of wells). The constant ϵ is a small number introduced to avoid matrix singularity during optimization iterations. The weights $\frac{\gamma}{(1+d(t))^{b_i}}$ incorporate the discounting effect pertaining to drilled wells.

The drilling cost is also accounted for in the formulation proposed by Forouzanfar and Reynolds [24] where given the rate of the k^{th} well at the l^{th} and $(l-1)^{\text{th}}$ iterations, denoted by u_k^l and u_k^{l-1}, respectively, they define the drilling cost function as $f_k^l(u^l) = (\frac{u_k^l}{u_k^{l-1}})^\beta$ where $\beta \in (0, 1]$. To compute the drilling costs, $f_k^l(u^l)$ is multiplied by C_{inj} or C_{prd}, the cost of drilling one injection or production well, depending on whether the well is an injector or producer. As the authors discuss, the drilling cost functional defined this way is continuously differentiable and drives the rates of negligible wells to zero while wells with higher rates tend to grow in their control allocation during the optimization iterations. At the end of the optimization, $f_k^l(u) = 1.0$ for all the wells with $u_k^l > 0$ to compute the corresponding net present value. In [24], the authors have studied the effect of choosing different values for the parameter β and conclude that for different values of β the resulting net present values are very close, although the well configurations and the required simulation runs may vary for different values of β.

In the formulation developed by Forouzanfar and Reynolds [24], it is assumed that all the wells are drilled at the beginning of the operation and once wells are eliminated they cannot be added at the later iterations. In our proposed formulation, wells with dynamic control trajectories are included and discounting is applied to the drilling term in the objective function (21) to optimize the drilling schedule.

Shirangi et al. [48] also consider the optimization of well drilling sequence in a closed-loop field development scheme. In their approach, the drilling time for a new well is determined after fixing the drilling times of previously added wells. In our framework, the drilling times of all the wells are inferred from their control trajectories, which are optimized for all wells *simultaneously*, along with their locations, number, and type. Furthermore, multiple wells are allowed to be drilled at the same time (assuming drilling resources allow for it). While a closed-loop implementation of our framework is straightforward (by adding a dynamic data feedback mechanism), we have focused on an open-loop formulation to introduce our approach.

2.3 Solution method: iterative shrinkage thresholding algorithm

A main difficulty after formulating the above optimization problem is solving it with an efficient and reliable

algorithm. In particular, the well drilling cost is non-differentiable, and standard optimization algorithms with smooth functions cannot be used to minimize it. A general class of methods that is suitable for a mixture of differentiable and nondifferentiable functions is "proximal splitting", where splitting refers to separating the differentiable and nondifferentiable parts while proximal reference is because the nonsmooth functions are minimized using proximity operators. In this section, we develop a special case of proximal splitting methods, known as iterative shrinkage thresholding algorithm, which we use to solve the generalized field development optimization problem. We first describe the proximal gradient optimization techniques [42] with properties that are suitable for optimization of non-smooth functions in large-scale problems. Proximal splitting optimization methods [15, 42] use the concept of proximal mapping to find the solution to a convex optimization problem more effectively than classical approaches. The proximity operator of a convex function is in fact a generalization of the notion of a projection operator and can be defined as:

$$\text{prox}_f(\mathbf{x}) = \arg\min_{\mathbf{u}} \left(f(\mathbf{u}) + \frac{1}{2} \parallel \mathbf{u} - \mathbf{x} \parallel_2^2 \right) \tag{23}$$

Suppose f is a convex differentiable function with continuous gradient ∇f and that g is a convex non-smooth function. To solve the following optimization problem

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \quad O(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x}) \tag{24}$$

an iterative proximal forward-backward splitting method can be formulated (see Appendix A) as:

$$\begin{aligned} \mathbf{x}^{k+1} &= \text{prox}_{\alpha_k g}(\mathbf{x}^k - \alpha_k \nabla f(\mathbf{x}^k)) \\ &= \arg\min_{\mathbf{u}} \left(g(\mathbf{u}) + \frac{1}{2\alpha_k} \parallel \mathbf{u} - (\mathbf{x}^k - \alpha_k \nabla f(\mathbf{x}^k)) \parallel_2^2 \right) \end{aligned} \tag{25}$$

where, α_k is the step size with values in a suitable bounded interval (see [15] for a discussion on theoretical requirements for the choice of appropriate step-size for different classes of proximal splitting methods). As discussed in Parikh et al. [43, Section 1.2], from the definition in Eq. 23 it can be inferred that $\text{prox}_f(\mathbf{x})$ is a point that provides a trade-off between minimizing f and staying *in the proximity of* $\mathbf{x}$; hence, $\text{prox}_f(\mathbf{x})$ is sometimes referred to as a *proximal point* of $\mathbf{x}$ with respect to f. In $\text{prox}_{\alpha_k g}$, as used in Eq. 25, the parameter α_k acts as a relative weight or trade-off parameter between the two terms involved. This simple framework can be generalized to handle as many non-smooth constraint functions as desired, implying that a complicated objective function can be decomposed into the sum of its separate *simpler* components.

The problem in Eq. 24 is similar to the field development optimization problem in Eqs. 13–14 where the functions f and g correspond to the operating cost $-O(.)$ (typically differentiable) and drilling cost $C(.)$ (nondifferentiable) in Eq. 14, respectively. Proximal splitting methods are developed to separate these differentiable and non-differentiable components of the objective function to develop more effective solution algorithms.

As discussed in Section 2.2, the drilling cost function $C(.)$, defined and approximated in Eqs. 20 and 21 respectively, is not smooth because of the term $\ell_p \ , \ p \ll 1$, making the proximal splitting a suitable solution approach. It should be noted that for non-convex functions proximal splitting methods are gradient-based search techniques that can be trapped in local minima.

The iterative shrinkage algorithms [18, 25, 56] provide a straightforward solution approach to problems of the form given in Eq. 24. Recently, efficient extensions of iterative shrinkage thresholding algorithm (ISTA) with enhanced convergence rates have been developed [6, 38] for solving the problem in Eq. 24. Here, we use the fast iterative shrinkage thresholding algorithm (FISTA) of Beck and Teboulle [6] to solve the field development optimization problem. The pseudocode for the FISTA with gradient projection is outlined in Table 1. In Table 1, $\tau_\alpha(.)$ is a shrinkage operator for the updated matrix $(\mathbf{Q}_C^{k+1})$, according to the specified threshold $\alpha(= \frac{\gamma}{L})$. FISTA uses soft thresholding as the shrinkage mechanism that is implemented using the shrinkage function shown in Fig. 4b and defined as

$$\tau_\alpha(x) = \begin{cases} x - \alpha & x \geq \alpha \\ 0 & -\alpha \leq x \leq \alpha \\ x + \alpha & x \leq \alpha \end{cases} \tag{26}$$

The term $\nabla_{\mathbf{u}} J(\mathbf{U}^k)$ in Table 1 is the gradient of the objective function $J(\mathbf{u})$ $(= -NPV)$ in Eq. 14 with respect to the rate allocations at the k^{th} iteration. It comprises the gradients of both OPEX $(O(\mathbf{u}))$ and CAPEX $(C(\mathbf{u}))$ terms with respect to the rate allocations (detailed expression for the

Table 1 Iterative shrinkage thresholding pseudocode

Iterative optimization algorithm
Initialization $\quad t_1 = 1 \quad , \quad \mathbf{x}_1 = \mathbf{u}^0$
for$k \geq 1$ (while convergence criteria not satisfied)
$t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2}$
$\mathbf{u}^{k+1} = \mathbf{u}^k - \frac{1}{L} P_k \nabla_{\mathbf{u}} J(\mathbf{u}^k)$
$\mathbf{x}^{k+1} = \arg\min_{\mathbf{x}} \left\{ \frac{L}{2} \parallel \mathbf{x} - \mathbf{u}^{k+1} \parallel_2^2 + C(\mathbf{u}^k) \right\}$
$= \tau_{\frac{\gamma}{L}}(\mathbf{u}_C^{k+1}) = \left(\mid \mathbf{u}_C^{k+1} \mid - \frac{\gamma}{L} \right)^+ \text{sign}(\mathbf{u}_C^{k+1})$
$\mathbf{u}^{k+1} = \mathbf{x}^{k+1} + \frac{t_k - 1}{t_{k+1}} (\mathbf{x}^{k+1} - \mathbf{x}^k)$
$\mathbf{u}^{k+1} = P_\Omega(\mathbf{u}^{k+1})$
End

gradient of $C(\mathbf{u})$ is provided in Appendix B). The positive part operator $(\cdot)^+$ in Table 1 is defined as $(x)^+ = \max(x, 0)$. The sign$(\cdot)$ function operates element-wise on matrix inputs. The notation L represents the gradient Lipschitz constant of $\nabla_{\mathbf{u}} J(.)$ that sets an upper bound on the magnitude of the slope of $\nabla_{\mathbf{u}} J(.)$ within its feasible domain, that is

$$L = \sup \left\{ \frac{\| \nabla f(\mathbf{x}_1) - \nabla f(\mathbf{x}_2) \|}{\| \mathbf{x}_1 - \mathbf{x}_2 \|} \right\} \tag{27}$$

For the numerical experiments in this study, at each iteration, we approximate the Lipschitz constant of $\nabla_{\mathbf{U}} J(.)$ using

$$L^k \simeq \max_{\substack{m,n=1 \\ m \neq n}}^{k} \left\{ \frac{\| \nabla O(\mathbf{u}_m) - \nabla O(\mathbf{u}_n) \|}{\| \mathbf{u}_m - \mathbf{u}_n \|} \right\} \tag{28}$$

where $\mathbf{u}$ denotes the column-wise vectorized version of $\mathbf{U}$. The Lipschitz constant is approximated at each iteration as the maximum value among all computed Lipschitz constants using all pairs of the gradients calculated up to the current iteration. The gradient projection matrix $\mathbf{P}_k$ is applied to the gradient to modify the update direction to ensure that the updated solution lies within the feasible set specified by the constraints.

A simple field development optimization problem with rate-controlled wells can be formulated as a nonlinear programing problem with box and linear equality constraints to balance field injection and production rates:

$$\begin{aligned} \min \quad & O(\mathbf{u}) \\ \text{subject to} \quad & \mathbf{u} \leq \mathbf{1} \ , \ -\mathbf{u} \leq \mathbf{1} \\ & \mathbf{1}^T \mathbf{u}_i = 0 \quad , \quad i = 1, 2, \cdots, T \end{aligned} \tag{29}$$

Since entries of $\mathbf{u}$ denote the control *allocations* for each well at the specified control step, they should neither exceed the normalized upper bound (the maximum control allocation for an injector, normalized to $+1$) nor the normalized lower bound (the maximum control allocation for a producer, normalized to -1); hence, the inequality constraints in Eq. 29 are designed to enforce the required box constraints on entries of $\mathbf{u}$. To balance the input/output mass into the reservoir, a set of equality constraints, ($1^T \ \mathbf{u}_i = 0 \ , \ i = 1, 2, \cdots, T$) is used, where $\mathbf{1}$ is the all-ones vector of the same size as $\mathbf{u}_i$'s.

In this paper, we have used the iterative gradient projection optimization algorithm to solve the resulting bound-constrained optimization problem [35, Chapter 12]. In each iteration of the algorithm, the negative of the gradient of the objective function is projected onto the feasible set to define the direction of iterative updates. The gradient projection matrix $\mathbf{P}_k$, which modifies the direction of change in the NPV response surface at iteration k, is formed as [35, Chapter 12],

$$\mathbf{P}_k = [\mathbf{I} - \mathbf{A}_{\mathbf{u}}^T (\mathbf{A}_{\mathbf{u}} \, \mathbf{A}_{\mathbf{u}}^T)^{-1} \mathbf{A}_{\mathbf{u}}] \tag{30}$$

in which $\mathbf{A}_u$ is a $M \times NT$-dimensional matrix whose rows represent the subspace formed by the set of active constraints. In each iteration of the optimization, the set of all equality and inequality constraints in Eq. 29 is divided into active and inactive constraints. Active constraints include the T equality constraints in Eq. 29 as well as the inequality constraints for which $u_j = 1$ or $u_j = -1$, where u_j is the j^{th} entry in the vector $\mathbf{u}$. As outlined in Table 1, at the end of each iteration, the mass balance constraint is enforced by ensuring that the sum of injection and production allocations, with positive and negative signs, respectively, is equal to 0. This constraint is satisfied by defining the feasible set as

$$\Omega = \left\{ \mathbf{X} \in [-1, 1]^{N \times T} : \mathbf{l} \leq \mathbf{x}_i \leq \mathbf{u}, \text{ and} \sum_{j=1}^{N} x_{ji} = 0, \forall \, i = 1, 2, \cdots, T \right\} \tag{31}$$

where $\mathbf{u}$ and $\mathbf{l}$ correspond to the normalized upper and lower bounds on the maximum control allocations for injection and production wells, respectively: $\mathbf{u} = \mathbf{1}$ and $\mathbf{l} = -\mathbf{1}$. $\mathbf{x}_i \ , \ i = 1, 2, \cdots, T$ are the columns of the $N \times T$ dimensional control allocation matrix $\mathbf{X}$. Denoting $y_{ji} = x_{ji} - \frac{\sum_{j=1}^{N} x_{ji}}{N}$, $\forall j = 1, 2, \cdots, N$, $\forall i = 1, 2, \cdots, T$, it can be verified that the entries of the vector $\mathbf{y}$ defined as the shifted version of $\mathbf{x}$ sum up to 0. The projection operator $P_\Omega(y_{ji}) = \frac{y_{ji}}{\|\mathbf{y}\|}$, $\forall j = 1, 2, \cdots, N$, $\forall i = 1, 2, \cdots, T$ performs the desired mapping, where Ω is defined in Eq. 31. The projection operator defined this way comprises a normalization factor to satisfy the inequality bound constraints in Eq. 29.

3 Numerical experiment

In this section, we examine the performance of the proposed algorithm in a number of numerical experiments with two-dimensional and three-dimensional two-phase (oil-water) reservoir models.[1]

3.1 Example 1: synthetic 2D, two-phase reservoir

The first example involves a synthetic 2D model with heterogeneous permeability (Fig. 5). The model is discretized into $50 \times 50 \times 1 = 2500$ grid blocks. The simulation parameters for this experiment are summarized in Table 2. In this example, the oil price and the cost of water injection and disposal were set to \$80/stb, \$10/stb and \$10/stb, respectively.

[1] Numerical experiments were implemented and simulated using MATLAB Reservoir Simulation Toolbox (MRST) developed by SINTEF Applied Mathematics Group, [34].

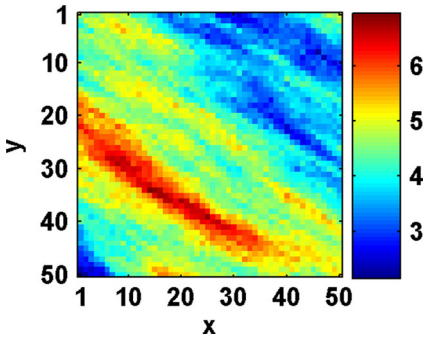


Fig. 5 The log-permeability model in Example 1

The annual discount rate was assumed to be 10 %. The well control trajectories are divided into 5 evenly spaced increments (control time steps) over the simulation time of 10 years.

While, in general, the solution procedure can be initialized with any feasible starting point, we assign very small rates to a large number of randomly placed producers and injectors to give each grid block the same chance of containing a well. Because of the sparsity-promoting nature of the drilling cost in the objective function, we apply the FISTA method (with adjoint-based sensitivities with respect to well controls) to solve the problem. Using this algorithm, the wells with insignificant contributions to production performance are successively removed to decrease drilling cost and to allow for active wells to assume higher rate allocations. More precisely, once a well control rate falls bellow the shrinkage threshold (denoted as α in the shrinkage function $\tau_\alpha(\cdot)$ in Eq. 26), the well drilling cost is eliminated from the objective function. It is important to note that while the drilling cost for inactive wells is removed, those wells are not eliminated from the optimization process and can assume significant control values (i.e. greater than α) in later iterations of the optimization. At convergence, this procedure results in simultaneous identification of the number and type of wells, well locations, well control trajectories, and their drilling schedule. We note that the idea of surrounding the main wells with pseudo wells that have small rates for deriving gradient information and improving well locations have been used by others in the past (e.g. [54, 55, 59, 60]). Our approach is different in that the size of the decision vector throughout the optimization remains unchanged while the values of the decision variables determine the well cost in the objective function.

The optimized well configuration for this example is shown in Fig. 6, which shows that the algorithm has placed two injectors (marked with red crosses) and one producer (marked with a black circle) to start reservoir operation in control time step 1. More wells are added as reservoir operation proceeds to later control steps; for instance, from Fig. 6, the algorithm adds a second producer in the bottom left corner of the field in control time step 2 (beginning of the third year of production) and activates the third injector near the left-hand side of the field in control time step 3 (beginning of the fifth year of production). The difference in drilling times of the wells in the optimized drilling schedule is due to the discounting effect in the well cost term of the NPV objective function, which favors fewer and later drilled wells to reduce the capital cost. The optimized

Table 2 Simulation parameter setup used for the numerical experiments

Parameters	Synthetic 2D Reservoir	PUNQ-S3 Reservoir (Ex. 1)	PUNQ-S3 Reservoir (Ex. 2)
Reservoir grid dimensions	$50 \times 50 \times 1 = 2500$	$28 \times 19 \times 5 = 2660$	$28 \times 19 \times 5 = 2660$
Number of active cells	2500	1761	1761
Physical cell dimension	$40 \times 40 \times 40$ ft^3	$40 \times 40 \times 40$ ft^3	$40 \times 40 \times 40$ ft^3
Rock porosity	30 %	30 %	30 %
Simulated reservoir life cycle	10 yrs	10 yrs	10 yrs
Number of control steps	5	5	5
Fluid phases	Oil-water (2-phase)	Oil-water (2-phase)	Oil-water (2-phase)
Initial water saturation	0.05	0.05	0.05
Injection volume	1 PV	1 PV	1 PV
Perforation direction	Vertical	Vertical	Vertical
Injectors control mode	Total water injection rate	Total water injection rate	BHP
Producers control mode	Total fluid production rate	Total fluid production rate	BHP
Well drilling cost	\$10 M	\$10 M	\$10 M
Oil price	\$80/stb	\$80/stb	\$80/stb
Water injection cost	\$10/stb	\$10/stb	\$10/stb
Water disposal cost	\$10/stb	\$10/stb	\$10/stb
Annual discount rate	10 %	10 %	10 %

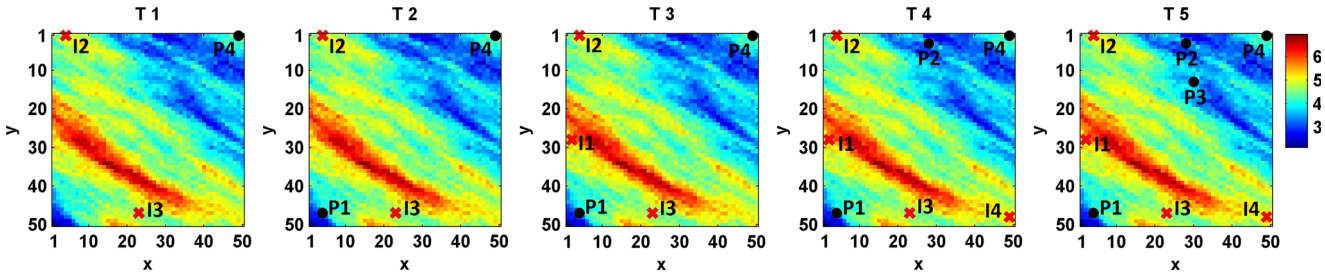


Fig. 6 Optimized well configuration within the control time steps 1–5; the injectors and producers are marked with (×) and (•), respectively. **Ti** is the i^{th} control time step

control time trajectories for producers and injectors are displayed in Fig. 7. From Fig. 7, Producer 4 starts production in the beginning of water flooding whereas Producers 1, 2, and 3 are drilled in second, fourth and fifth control time steps, respectively. Figure 7 also shows that Injectors 2 and 3 start injection in the first control time step and Injectors 1 and 4 are added in the third and fourth control time steps, respectively. The algorithm has placed a total of 8 wells in the field, which has resulted in an NPV value of $ 263.19 M after 10 years of operation. The oil saturation profile of the reservoir after 10 years of operation with the optimized well configuration and control settings is shown in Fig. 8. The evolution of the objective function (negative NPV), including the OPEX ($O(\mathbf{U})$) and CAPEX ($C(\mathbf{U})$) terms, throughout the optimization iterations is shown in Fig. 9. The algorithm has converged after 64 iterations. The current implementation of the algorithm uses a backtracking line search method [8, Section 9.2] to update the solution; hence, each iteration requires a few reservoir simulation forward runs to properly adjust the step size. In the present experiment, for about the first 40 iterations, the number of full reservoir simulations needed ranged between 1 to 3 runs. In the intermediate iterations, it took more simulation runs to calculate the appropriate step size (1∼6 runs). Towards the end of the optimization (iteration 60 onward), it could take up to 10 simulation runs for some iterations to reduce the objective function. In total, about 160 reservoir simulation forward runs were needed before the algorithm converged. This performance can be improved by using more efficient step size selection methods [5, 17, 23, 45].

While the decline in the objective function value and the capital expenditures has a monotonically decreasing trend

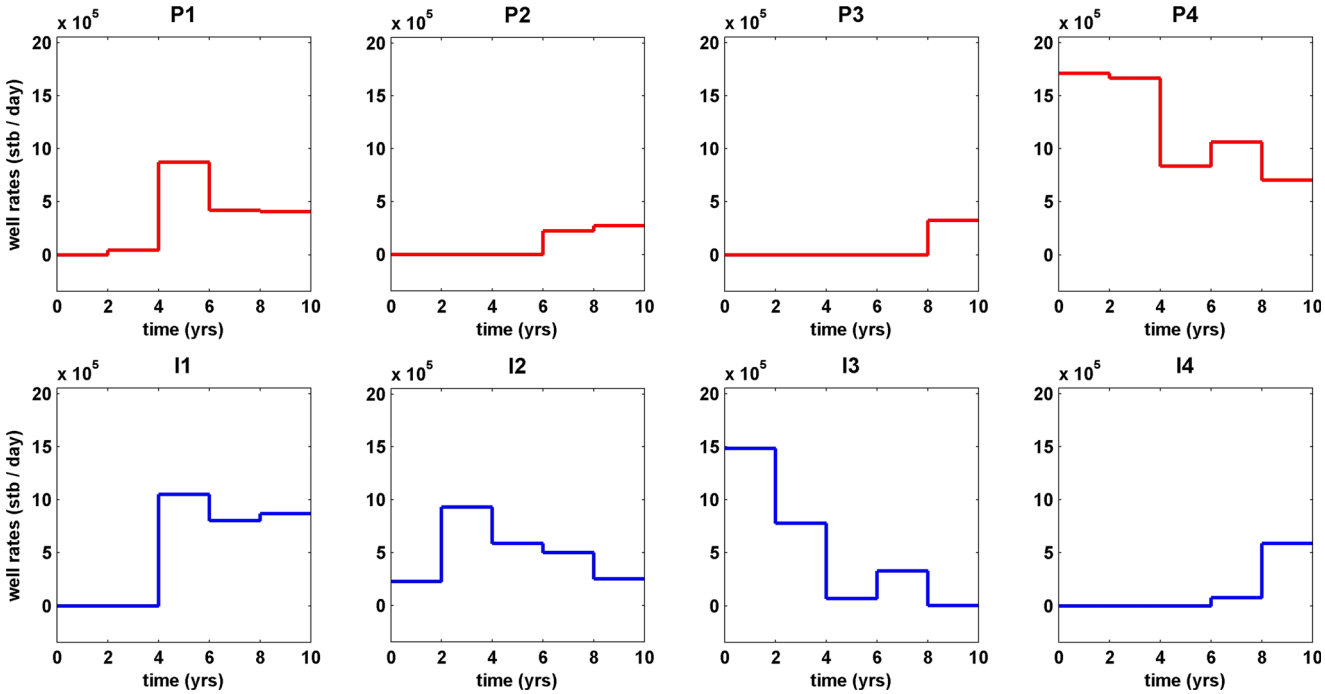


Fig. 7 Optimized 5-step control trajectories of the producers (*top*) and injectors (*bottom*) in Example 1

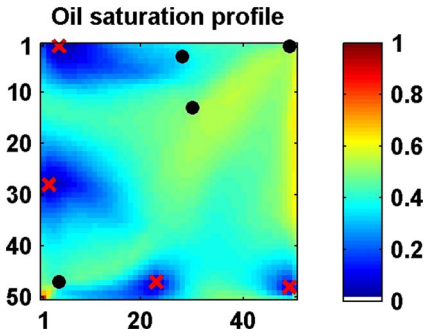


Fig. 8 Final oil saturation profile of the reservoir after 10 years of operation with optimized well configuration and control settings in Example 1

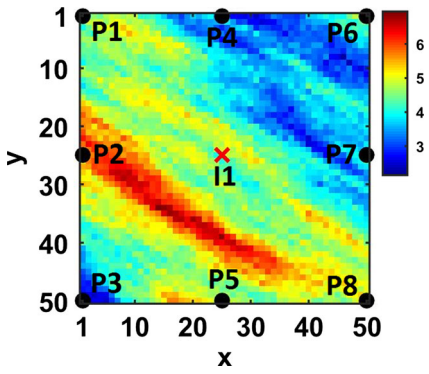


Fig. 10 9-Spot well pattern as initial well configuration for Example 2

(as expected), the evolution of the OPEX term exhibits a fluctuating behavior, especially in the intermediate iterations. The abrupt changes in the operating costs during the intermediate iterations are resulted from the removal of several wells that can have a significant (sometimes unfavorable) impact on the OPEX term, but with a net favorable effect on the overall objective function. These fluctuations tend to disappear as the algorithm approaches its convergence (after about iteration 50), where the dominant wells are mostly selected and fewer wells need to be removed.

3.2 Example 2: initializations with conventional 9-spot pattern

In this example, the same optimization problem as in Section 3.1 is considered. However the solution procedure is initialized with a 9-spot well pattern (1 injector in the center and 8 producers located in the corner points and near the boundaries of the reservoir) as displayed in Fig. 10; each well was surrounded by a local set of wells of the same type with very small rates. The single injector was assigned maximum injection rate while the 8 producers were assigned uniform control rate allocations. Other than the initialization of the algorithm, the simulation experiment setup is exactly the same as Example 1 in Section 3.1. Similar to Example 1, the algorithm simultaneously determines the optimal solution to the configuration in terms of number, locations and types of the wells, their dynamic controls and drilling times. The optimized wells configuration for this example is shown in Fig. 11 where it can be seen that the algorithm has reduced the number of initial wells to a total of 6 wells: 5 producers and 1 injector. Comparing the optimized wells locations (Fig. 11) with the initial 9-spot pattern (Fig. 10) it is observed that producers 2, 4, and 5 are removed in the optimized configuration. The locations of the remaining wells (I1, P1, P3, P6, P7, and P8) have undergone some local changes in their nearby zones after optimization and more importantly, their drilling times are altered in the optimized schedule. While in the initial 9-spot pattern all the wells are drilled in the beginning of reservoir operation, as inferred from the optimized dynamic controls shown in Fig. 12 wells P1, P3, P6, P7, and P8 are drilled

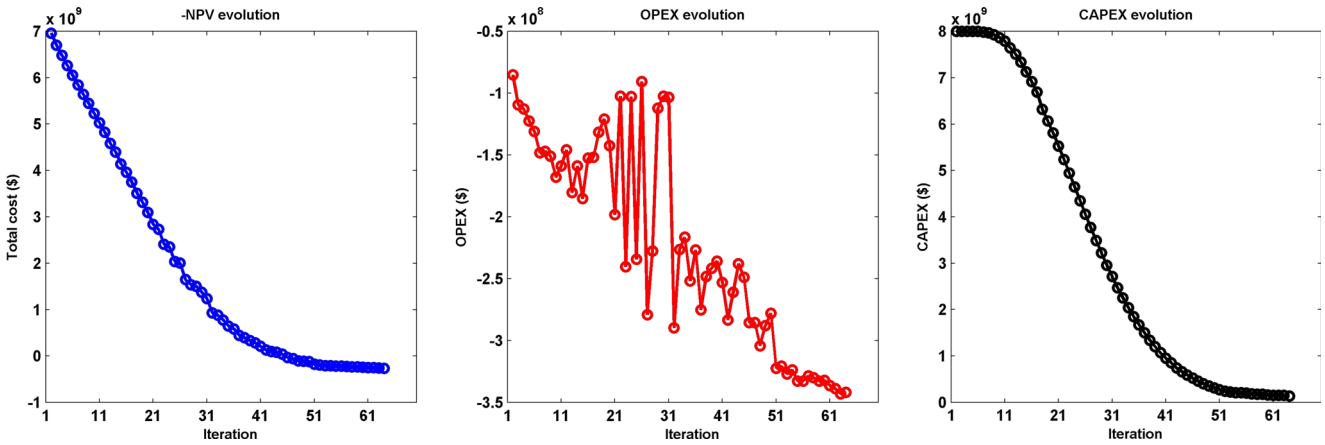


Fig. 9 Evolution of objective function (*left*), the operating cost function (*middle*) and capital expenditures (*right*) in Example 1

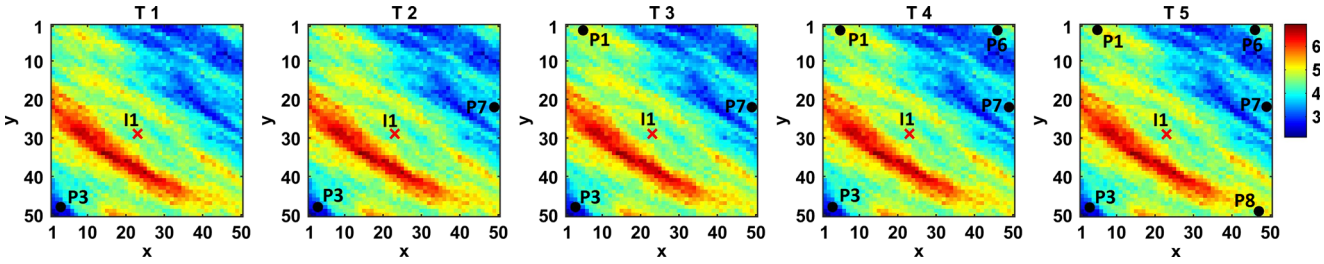


Fig. 11 Optimized well configuration within the control time steps 1-5 for Example 2

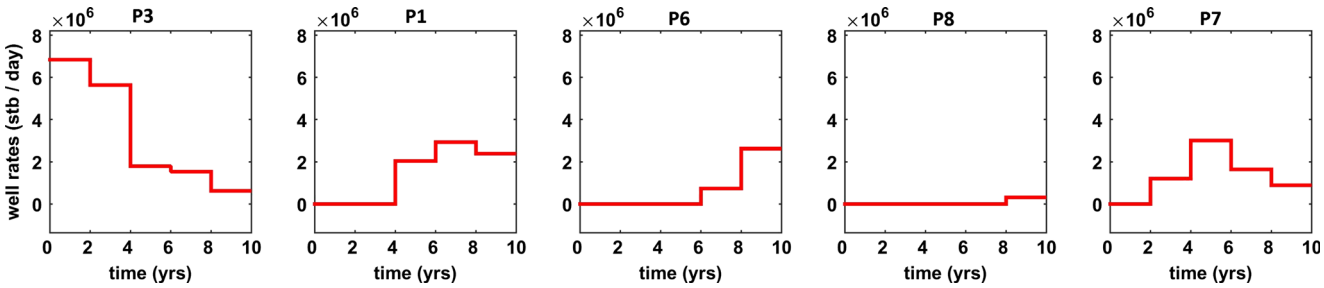
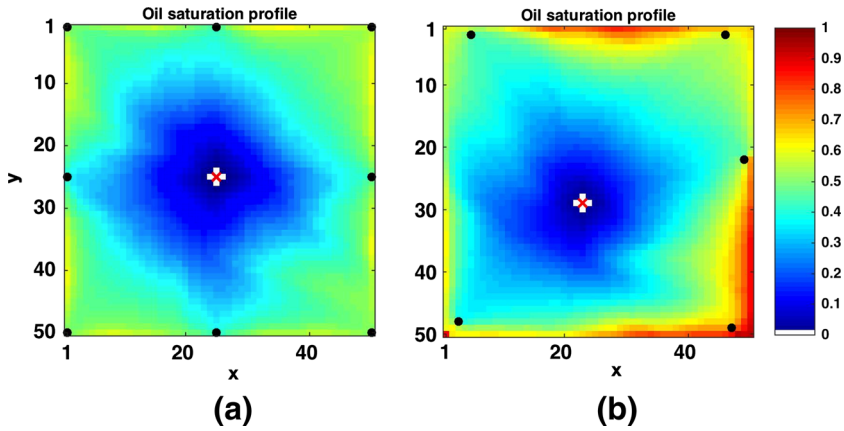


Fig. 12 Optimized 5-step control trajectories of the producers in Example 2

Fig. 13 Final oil saturation profile of the reservoir after 10 years of operation with 9-Spot well pattern (**a**) and with optimized well configuration and control settings (**b**) in Example 2



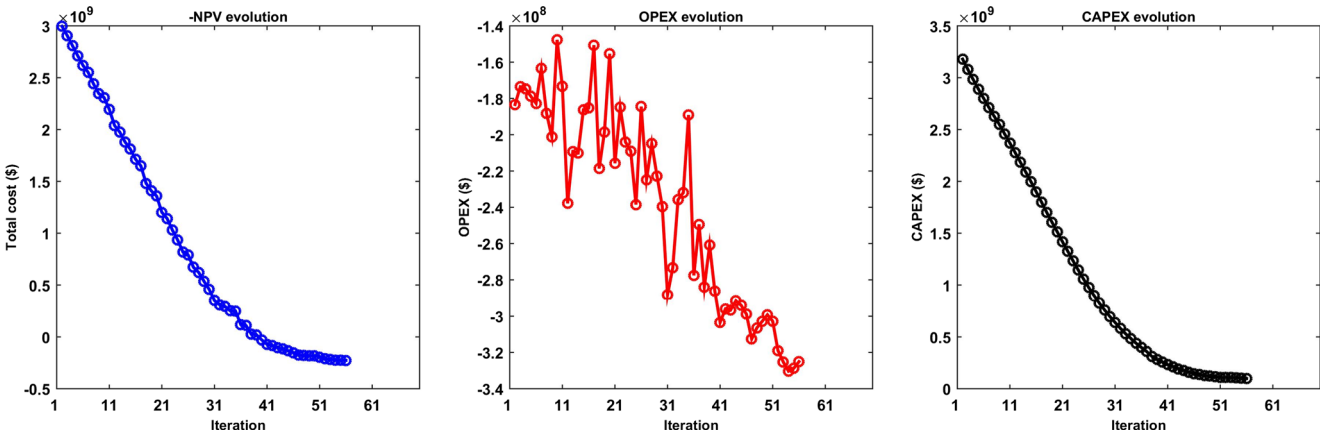


Fig. 14 Evolution of objective function (*left*), the operating cost function (*middle*) and capital expenditures (*right*) in Example 2

in the third, first, fourth, second, and last control time steps respectively. Injector I1 is drilled in the first control time step and operates with maximum injection rate throughout reservoir life time. The final oil saturation profiles of the reservoir after 10 years of operation with both conventional 9-spot well pattern and uniform control settings and the

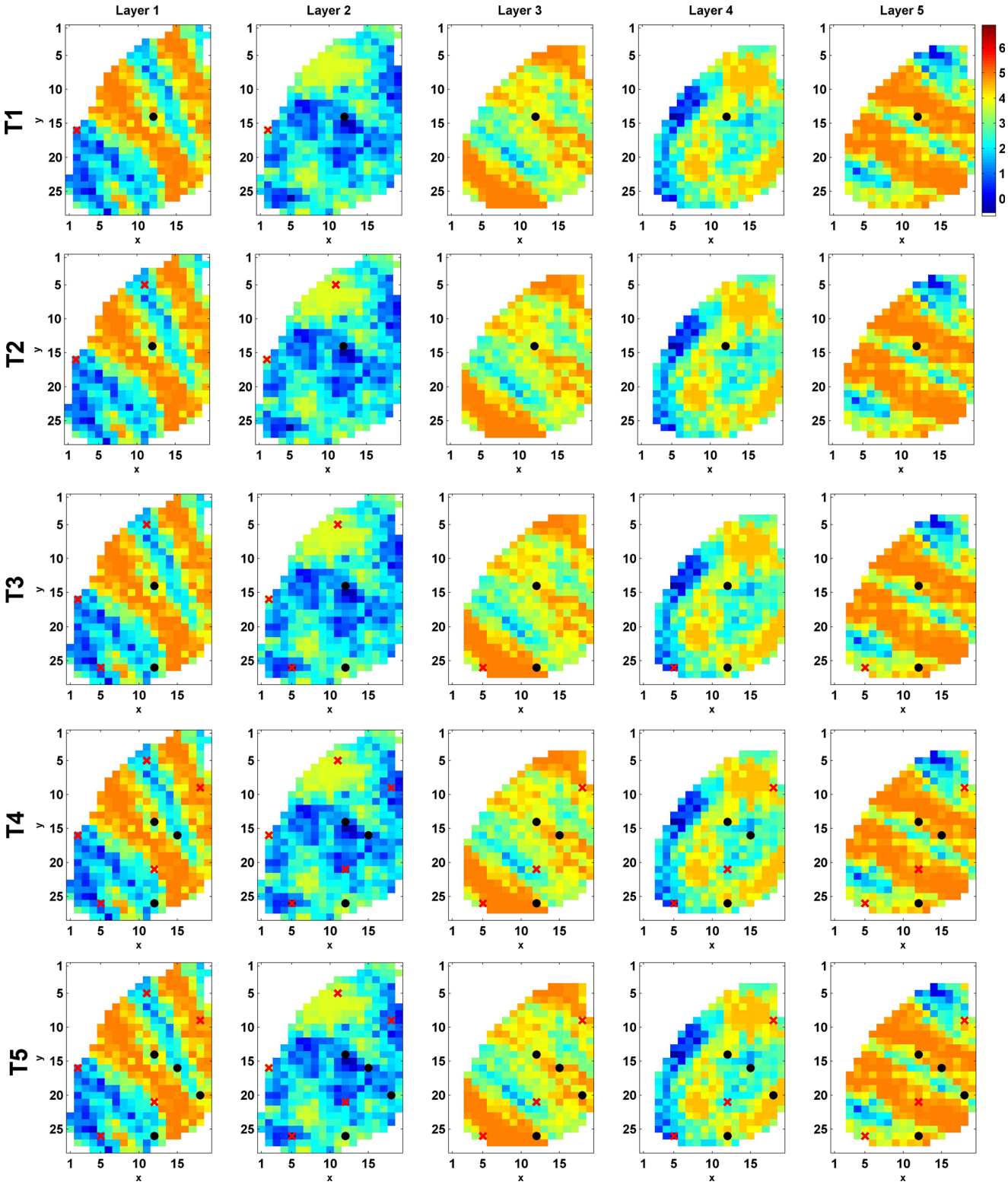


Fig. 15 Optimized well locations within control time steps 1–5 in Example 3; the injectors and producers are marked with (×) and (•), respectively

Table 3 Optimization results

Experiment	N_{inj}	N_{prod}	NPV
Synthetic 2D reservoir (random initialization)	4	4	$ 263.19 M
Synthetic 2D reservoir (9-Spot initialization)	1	5	$ 240.84 M
PUNQ-S3 3D reservoir with rate controls	5	4	$ 101.21 M
PUNQ-S3 3D reservoir with BHP controls	5	5	$ 104.62 M

optimized well configuration and control rates are depicted in Fig. 13a, b. While the reduction in the number of wells for the optimized configuration has left some portions of the oil bypassed, the NPV of the project is higher for the optimized settings ($ 240.84 M) compared with the NPV that results under the conventional 9-spot well pattern with uniform well control settings ($ 174.36 M). The evolution of the objective function along with the changes in the CAPEX and OPEX portions of the NPV during optimization iterations are shown in Fig. 14. The algorithm has converged after 56 iterations.

The resulting NPV of the project with the optimized settings in Example 1 ($ 263.19 M) where the solution was initialized by randomly placing a large number of injection and production wells with small rates in the reservoir, is higher than the final NPV in this Example ($ 240.84 M) where a 9-spot well pattern was used as the initial guess. While a good choice for the initial solution can potentially improve the final results, the comparison between the results of Examples 1 and 2 implies that in the absence of such good initial guesses the algorithm can produce reasonable solutions when all cells are given an equal opportunity to contain a well.

3.3 Example 3: 3D 2-phase PUNQ-S3 reservoir

We applied the proposed algorithm to optimize well configurations and control settings in a 3D heterogeneous model with geologic structure borrowed from the PUNQ-S3 benchmark reservoir model. In our example, the gas-cap and the aquifer in the original PUNQ-S3 model are not included. The optimization was implemented over a cycle of 10 years. The log-permeability maps of the 5 layers in the PUNQ-S3 model are shown in Fig. 15. The field contains $28 \times 19 \times 5 = 2660$ discretized grid blocks, many of which are inactive as displayed in Fig. 15. The simulation parameters for this example are summarized in Table 2. All the wells are under fluid flow rate controls. Similar to Example 1, initially a large number of production and injection wells with small rates are randomly positioned throughout the reservoir. The wells are assumed to be vertical and perforated in all 5 layers of the PUNQ reservoir model except for the ones placed in the grid cells that are only active within the first 2 layers of the formation. The simulation accounts for the hydrostatic pressure gradient in vertical wells. The total fluid flow rate for each well is computed as the sum of the flow rates in all the grid cells perforated by that well:

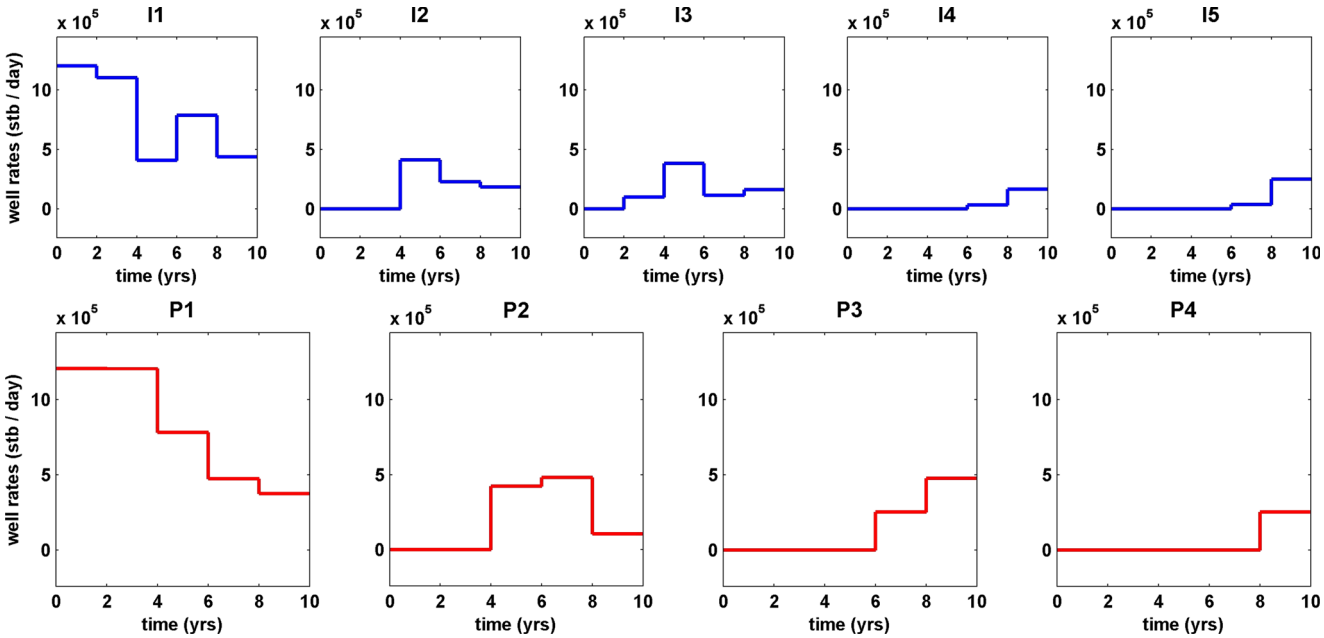


Fig. 16 Optimized control time trajectories for injectors (*top*) and producers (*bottom*) in Example 3

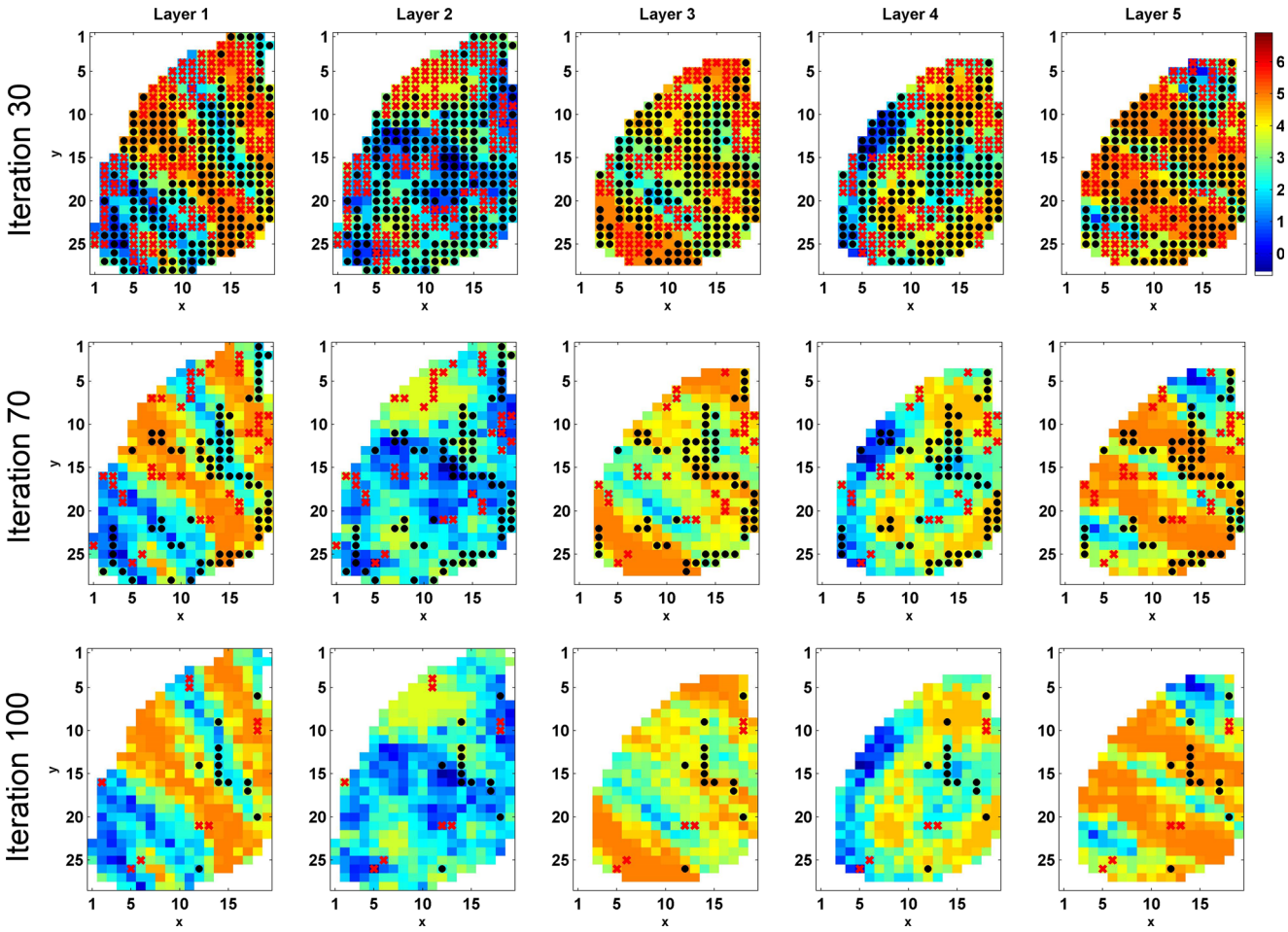


Fig. 17 Optimization solution in iterations 30 (Row 1), 70 (Row 2), and 100 (Row 3) in Example 3

$q_i^t = \sum_{j \in \mathcal{I}_i} q_{ij}^t$ where $\mathcal{I}_i$ includes the indices of all the grid cells perforated by well i and q_i^t is the total fluid flow rate of well i at control time step t.

The optimized well configurations for this case are displayed in Fig. 15. The number of wells and final NPV values are listed in Table 3. In this case, the algorithm has placed 5 injectors and 4 producers, with varying drilling times as illustrated in Fig. 15. The optimized well trajectories of the injection and production wells are shown in Fig. 16, showing that some of the wells are drilled in later years to improve the objective function. The optimization solutions within some intermediate iterations are displayed in Fig. 17. In early iterations (Iteration 30, Fig. 17, Row 1), the algorithm tends to remove the wells that have insignificant contributions, thereby allowing the optimal *regions* for

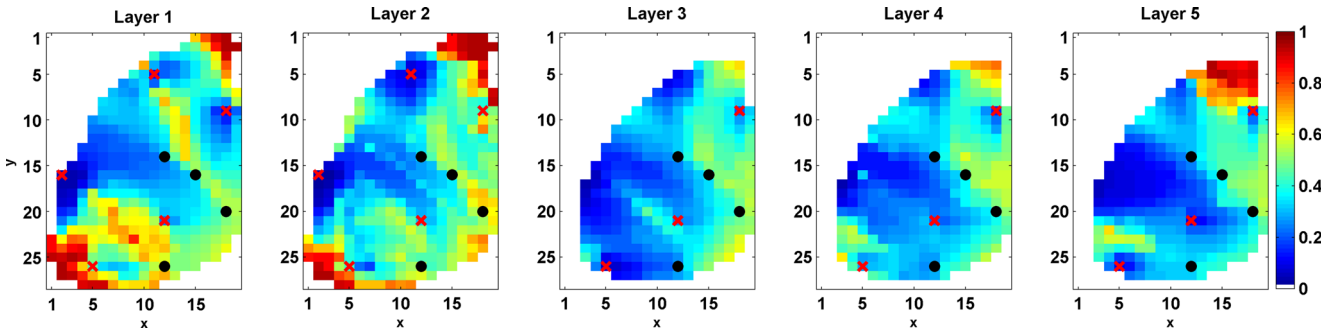


Fig. 18 Final oil saturation profile with optimized configuration and control settings of the wells in Example 3

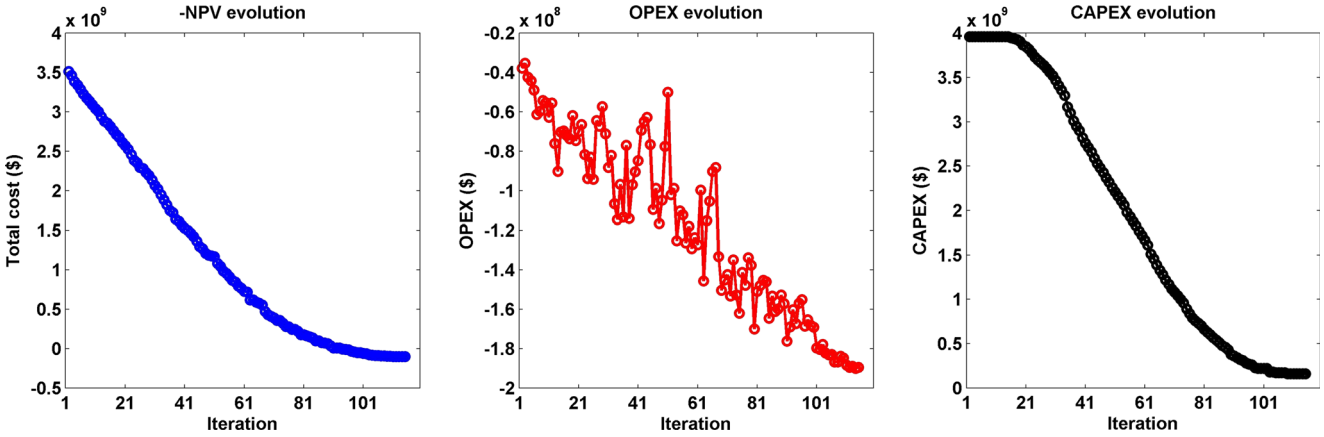


Fig. 19 Evolution of the objective function (*left*), the operating cost function (*middle*), and the capital expenditures (*right*) in Example 3

placement of production and injection wells to emerge gradually. At this initial stage, the algorithm removes several producers that are located at the proximity of the injectors to avoid unfavorable early watercuts. Further refinement of these regions is carried out during the intermediate iterations, when the potentially high-pay regions for placement of candidate wells are identified. At this stage, the algorithm begins to further shrink the size of the selected zones by increasing the control allocations of more dominant wells at the cost of removing the wells with less significant contributions (Iteration 70, Fig. 17, Second Row). This gradual improvement in the solution through successive removal of wells with less significant contributions is achieved through the iterative shrinkage thresholding mechanism that was discussed in the previous section. Note that the iterative scheme is implemented using an approximation of ℓ_0 norm minimization with ℓ_p-norm ($p < .01$ at convergence). That is, during the iterations, to have a better-behaved sparsity-promoting function, the value of $p \in (0, 1]$ is successively reduced to its very small values ($p \longrightarrow 0$) at convergence. In the terminating iterations (Iteration 100, Fig. 17, Third Row), the optimal well configuration is almost identified with some well clusters that need to be further reduced to single wells. The oil saturation profiles of the optimized configuration after 10 years are shown in Fig. 18. The ultimate NPV is $ 101.21 M. From Fig. 18, some bypassed oil is observed in the top-right and bottom-left corners of the field, which are primarily active within the first two layers of the formation and have relatively low permeability values. While it might seem that addition of more wells in these regions can increase the NPV, the revenue from the produced oil from those regions does not offset the cost of drilling new wells. The evolution of the objective function and the CAPEX and OPEX terms in the 115 iterations of the algorithm are displayed in Fig. 19. From Fig. 19, the decline in the cost function is faster during the initial and intermediate iterations (Iterations $1 \sim 80$) where numerous candidate wells still exist in the field; each time a *group* of these insignificant wells is removed, the cost function undergoes considerable reduction. When the algorithm proceeds to later iterations, the majority of insignificant wells are already removed and it takes several iterations for the algorithm to finalize the control allocations of the remaining wells within the smaller well clusters. In this example, the algorithms takes 15 additional iterations to reduce the population of potential injection wells from 9 (Iteration 100) to 5 (Iteration 115).

Table 4 Optimization results for multiple initializations

Initialization No.	N_{inj}	N_{prod}	NPV
1	5	4	$ 101.21 M
2	4	5	$ 98.57 M
3	4	5	$ 91.82 M
4	4	4	$ 106.23 M
5	4	4	$ 98.93 M

An important point to discuss is the local nature of the gradient-based optimization algorithm. The well placement and control optimization problems are known to have several local solutions with relatively close final performance (NPV) outcomes. This is mainly related to the nature of the problem where a certain volume of water is injected to displace a finite amount of producible oil from the reservoir. As can be expected, this goal can be achieved through several optimized scenarios with comparable outcomes. While global optimization techniques exist that heuristically search through all (local) solutions to identify the global solution, these methods are computationally demanding and do not render themselves to solving optimization problems with high-dimensional decision variables (as is the case in this formulation). To test the sensitivity of the solutions obtained from the proposed gradient-based

algorithm, the method was applied with five different initializations. The number of injection and production wells and the final NPV values for these test cases are summarized in Table 4. Figure 20 shows the final well configuration for these solutions, indicating that the optimized configurations for different initializations vary, which is expected

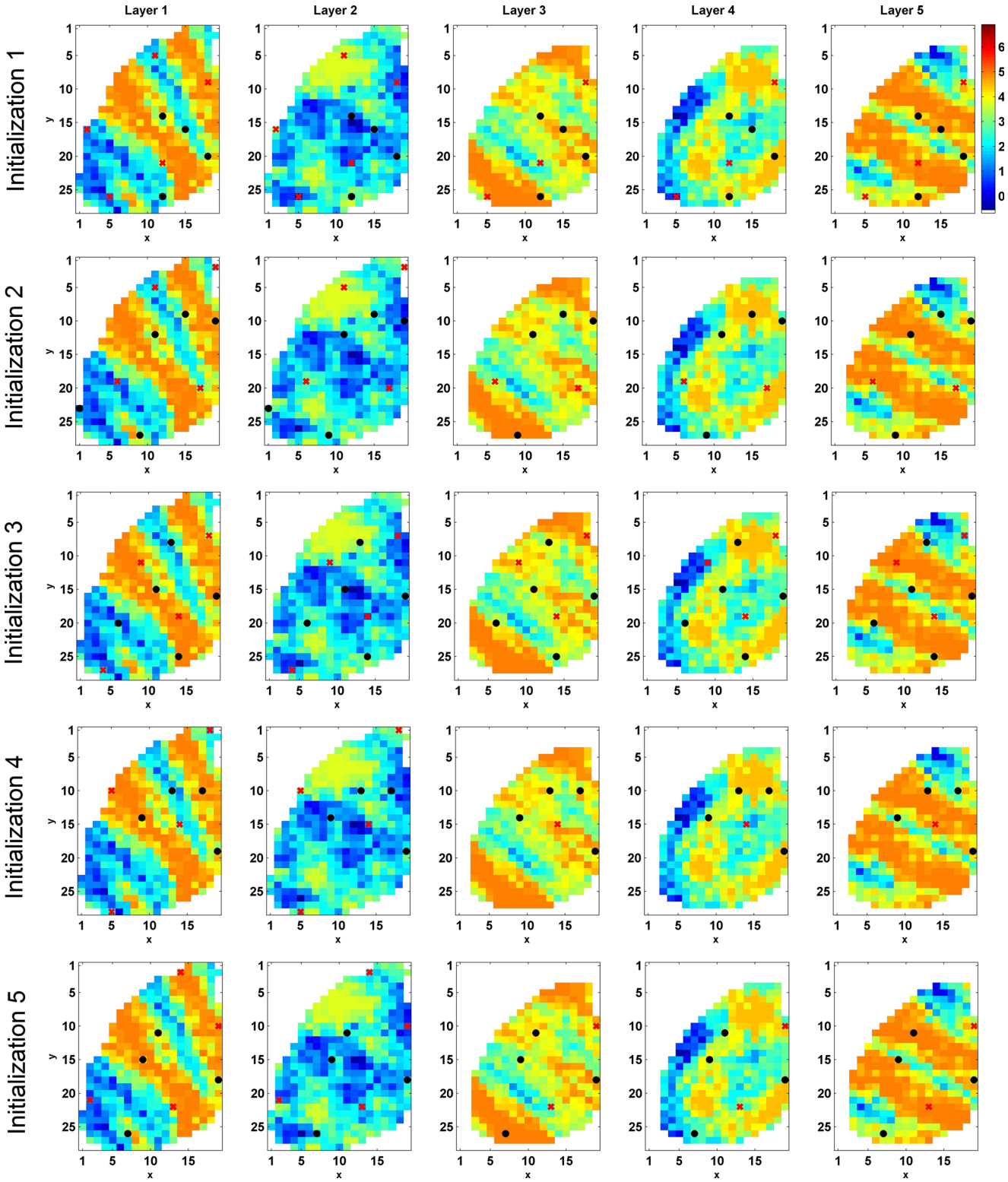


Fig. 20 Optimized well configurations for five different initializations in Example 3

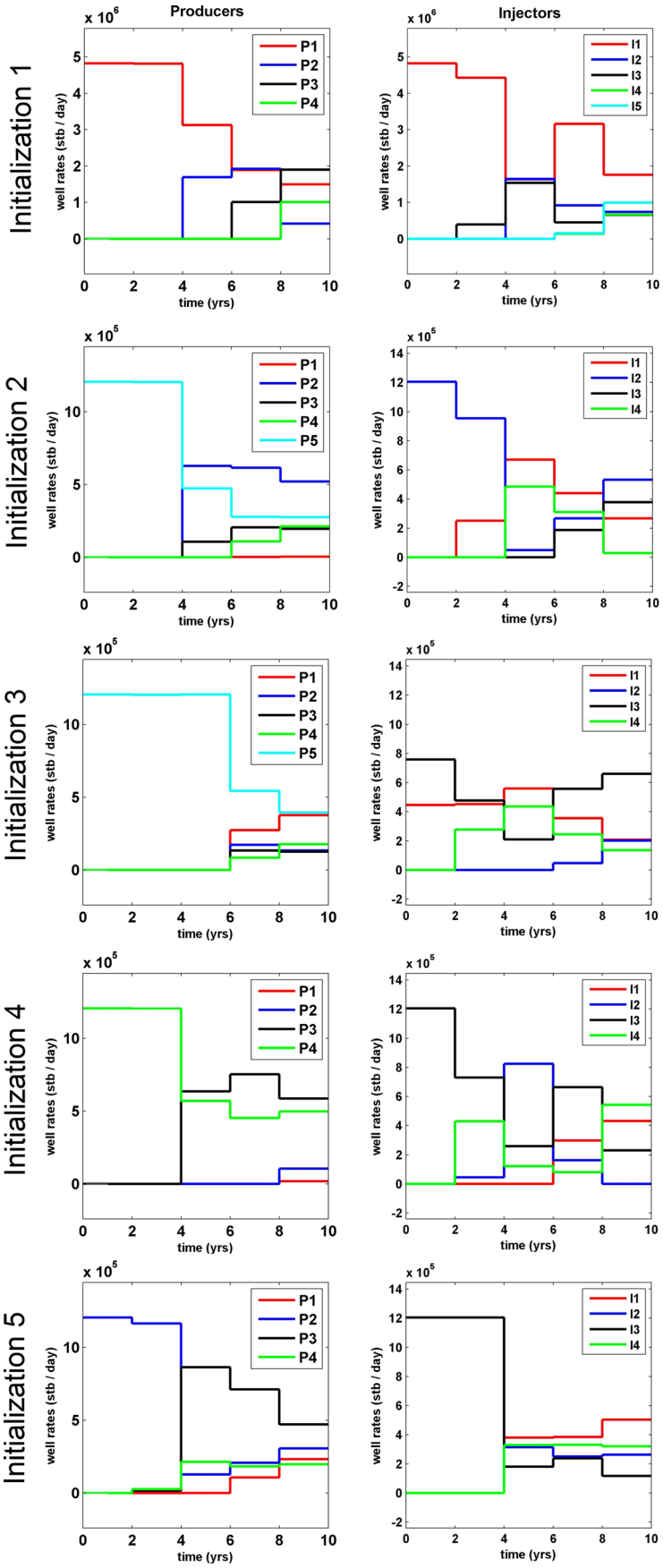


Fig. 21 Optimized control time trajectories in Example 3 with five different initializations

given the local nature of gradient-based algorithms and the number (and type) of variables that are included. Figure 21 shows the final control trajectories for different initializations, which are quite different. From Table 4 and Fig. 22, we also observe that the optimized NPV values are relatively close to each other, and fluctuate around $\sim$ \$ 100 M. This result is consistent with the observation by Van Essen et al. [53], that in large-scale production optimization problems, the NPV objective function typically consists of long connected plateaus on which the value of the objective function remains roughly the same. A similar observation was also made in history matching literature by Oliver et al. [40] and Oliver and Chen [39].

3.4 Example 4: 3D two-phase PUNQ-S3 reservoir with BHP controlled wells

As discussed in Section 2, the problem formulation can be applied to the case where the wells are controlled through their bottom-hole pressures (BHPs). From Eq. 11, for BHP-controlled wells the source/sink components ($\mathbf{J}_3$ ($\check{\mathbf{p}}_{well} - \mathbf{p}_3$)) are represented with bottom-hole pressure values of the wells. Therefore, the pressure gradient, i.e. the difference between the well's BHP and the corresponding grid cell pressure in each grid block, can be used to identify the grid blocks that are penetrated by a well (pressure gradients ($\check{\mathbf{p}}_{well} - \mathbf{p}_3$) will be non-zero *only* for the grid cells intersected by a well). Provided with the pressure gradients corresponding to each grid cell at every time step, one can infer the type and dynamic control values for each well. In fact, similar to the rate allocations matrix $\mathbf{Q}$ (6), we can concatenate time-varying pressure gradients ($\check{\mathbf{p}}_{well} - \mathbf{p}_3$) of all the grid blocks to construct the matrix

$$\mathbf{H} = \begin{bmatrix} \Delta p_1^{t_1} & \Delta p_1^{t_2} & \cdots & \Delta p_1^{t_T} \\ \Delta p_2^{t_1} & \Delta p_2^{t_2} & \cdots & \Delta p_2^{t_T} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta p_N^{t_1} & \Delta p_N^{t_2} & \cdots & \Delta p_N^{t_T} \end{bmatrix} \tag{32}$$

that combines several decision variables into a structured spatio-temporal representation to formulate and solve the generalized field development optimization problem. In the matrix form of Eq. 32, the entry $\Delta p_i^{t_j}$ is equal to the pressure gradient in the i^{th} discretized grid block at time step t_j. Depending on its sign, $\Delta p_i^{t_j}$ indicates the type of well in the i^{th} grid block in the discretized spatial domain (with $\Delta p_i^{t_j} > 0$ and $\Delta p_i^{t_j} < 0$ indicating existence of an injector and a producer, respectively, while $\Delta p_i^{t_j} = 0$ indicating the absence of a well). The problem can then be treated in the same fashion as the case with rate-controlled wells by adjusting the total fluid flow rates. The formulation and solution approach that were discussed in Sections 2 and 2.3 can now be applied to set up and solve the problem for BHP-controlled wells where the dynamic control trajectories represent time-varying bottom-hole pressure values of the wells.

The above algorithm was implemented for a water-flooding experiment in the 3D PUNQ-S3 reservoir model. The parameters of this simulation experiment are listed in Table 2. As before, the solution is initialized by randomly placing many producers and injectors with small pressure gradients (translating into small production/injection rates) throughout the field. In each iteration of the optimization, dynamic pressure gradients in the grid cells are used to decide which wells are of relatively negligible contribution and should hence be removed. The optimized well configuration and drilling sequence are displayed in Fig. 23. The algorithm has placed 5 injectors and 5 producers with various drilling times as demonstrated in Fig. 23. Figure 24 shows the optimized BHP control trajectories for the injection and production wells along with the corresponding fluid flowrates in those wells. In control time steps before a well is activated, the bottom-hole pressure value for that well is essentially equal to the cell pressure in that grid block. The drilling cost is included only when the pressure gradient in a grid cell becomes nonzero. As shown in Fig. 25, in this case the algorithm has converged with a final net present value of \$ 104.62 M after 111 iterations, which is consistent with the results obtained in the previous example. Figure 26 shows

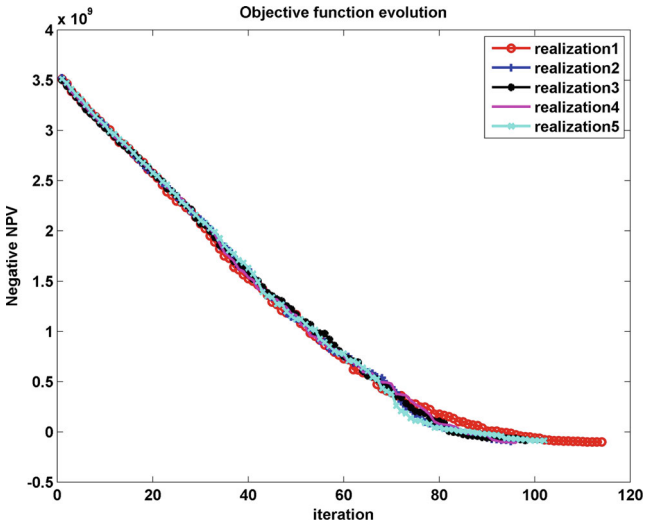


Fig. 22 Objective function evolution for five different initializations in Example 3; the plot shows different solutions with similar NPVs

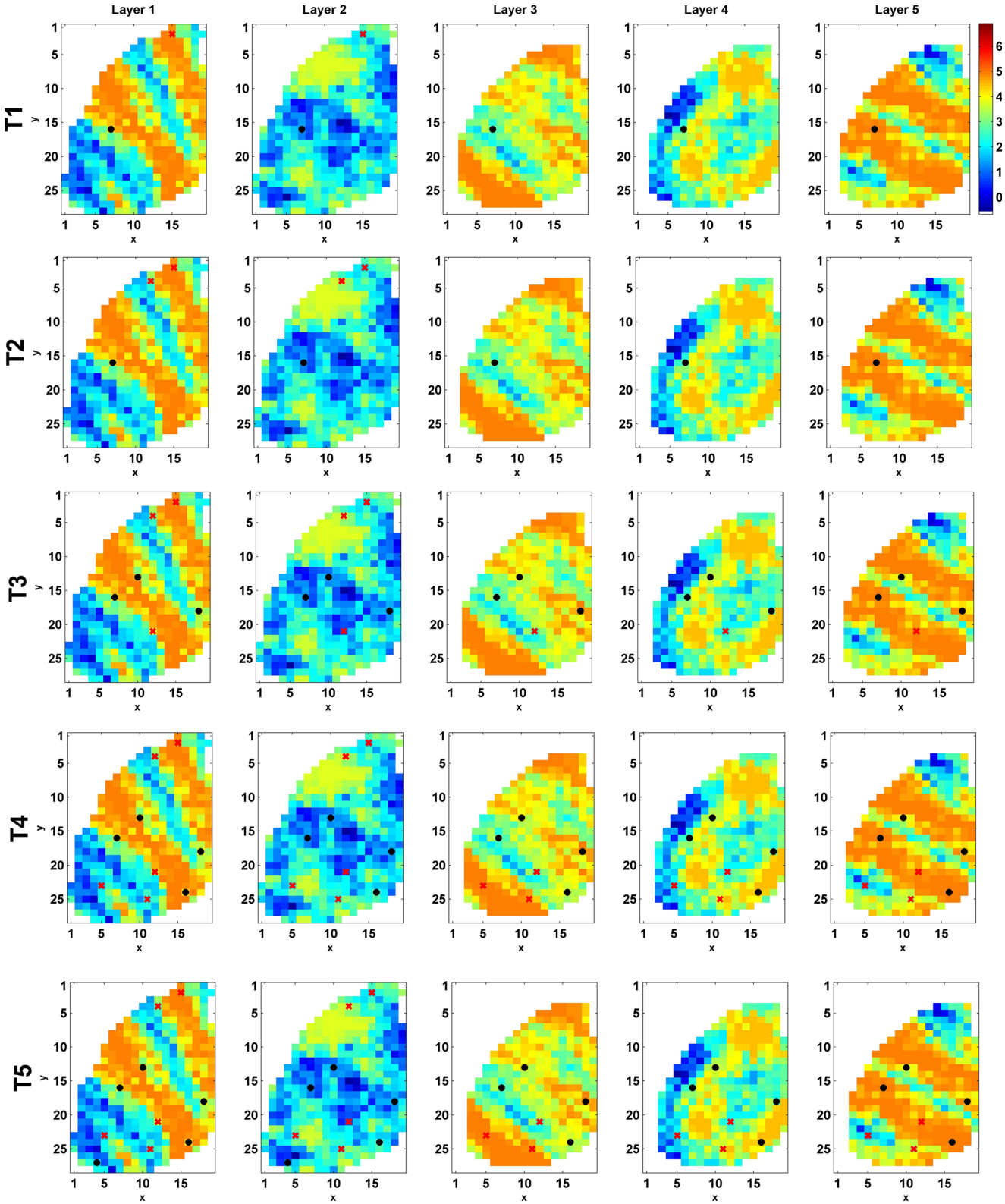


Fig. 23 Optimized well locations within control time steps 1–5 in Example 4 (note that some wells become active at later control steps)

the optimized oil saturation profile after 10 years. Although Layers 1 and 2 show some high oil saturations in the lower-left corner of the domain, the permeability and porosity values in this region are very low (unlike Layers 3-5). Producing the small volume of oil from this region requires significant amount of energy, which is not economically

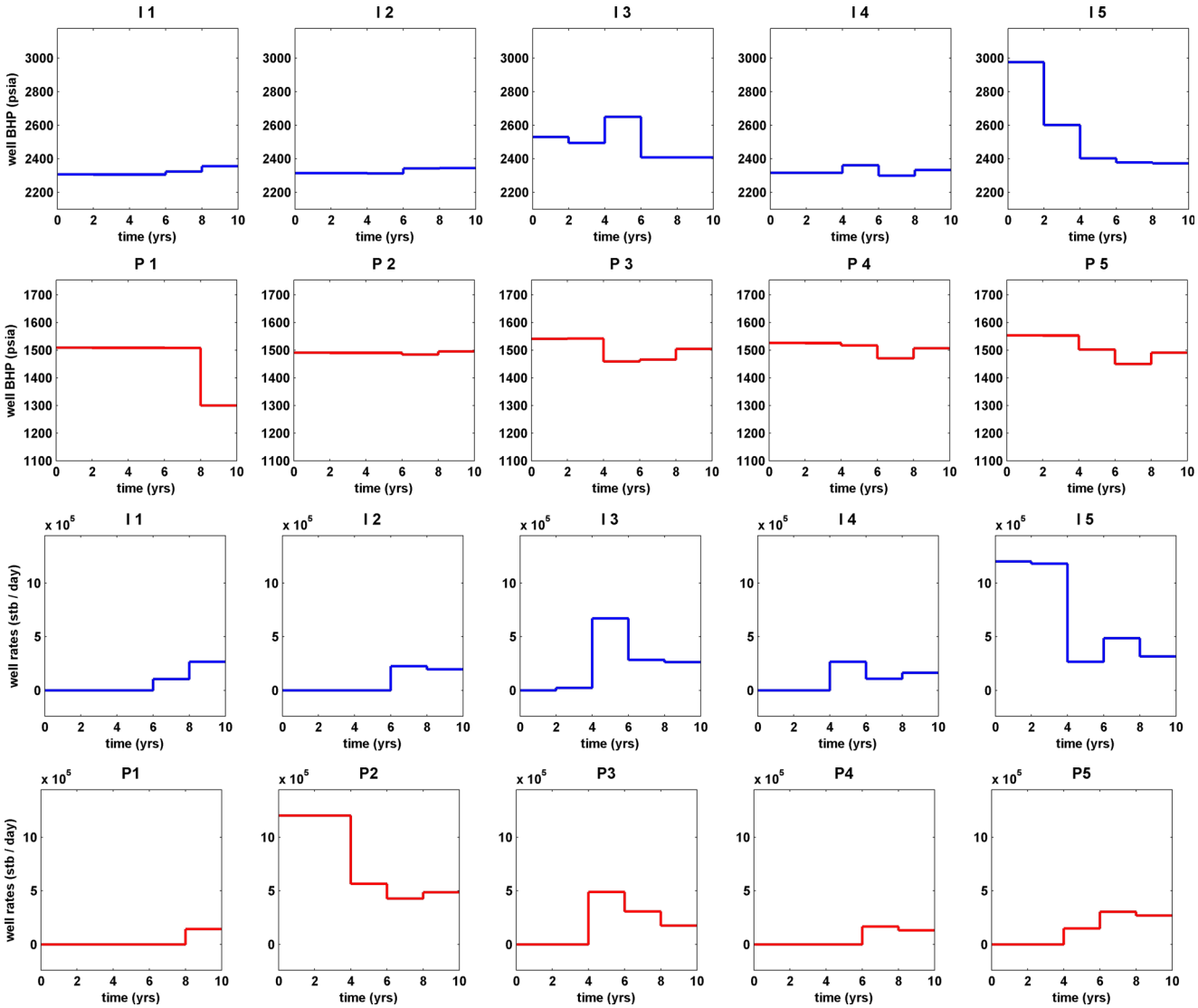


Fig. 24 Optimized injection BHPs (Row 1), production BHPs (Row 2), injection flowrate (Row 3), and production flowrate (Row 4) trajectories in Example 4

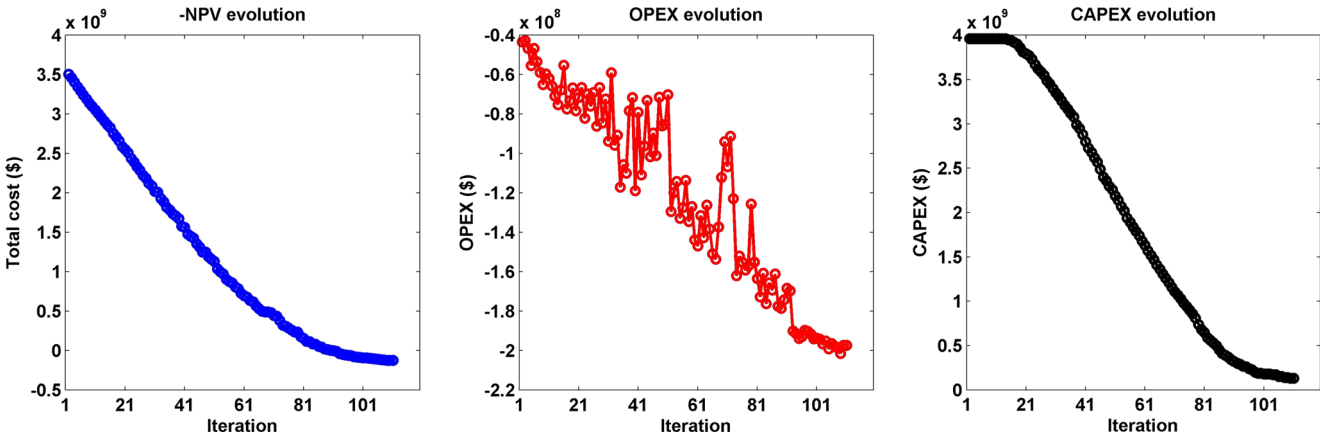


Fig. 25 Evolution of the optimization objective function (*left*), the operating cost function (*middle*) and the capital expenditures (*right*) in Example 4

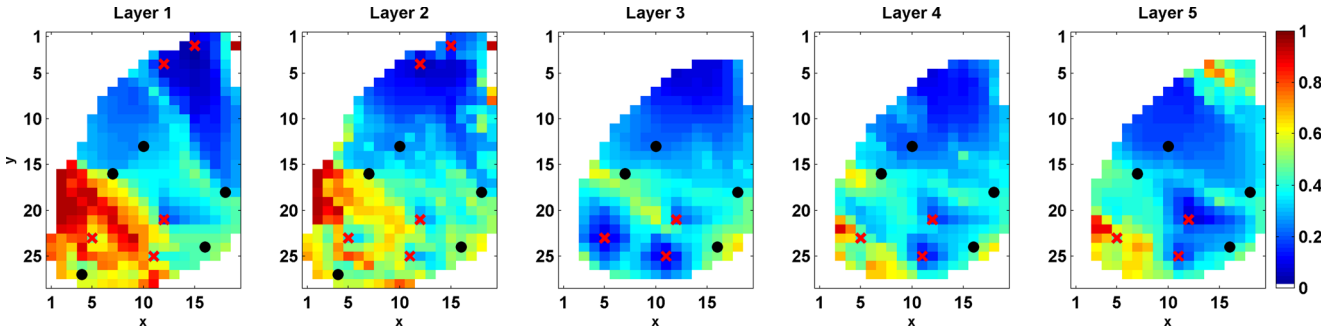


Fig. 26 Final oil saturation profile of the reservoir after 10 years of operation with optimized well configuration and control settings in Example 4

viable. Finally, in all of the examples discussed above, the control time trajectories of the wells were divided into 5 control time steps of 2 years to span a 10 year reservoir life cycle. A smaller control time step can be used to provide a higher temporal resolution for well dynamic controls, which could potentially improve the optimized NPV values. This, however, comes at increased computational cost due to the larger size of the decision variables.

4 Conclusion

We developed a novel generalized oilfield development optimization framework that can be used to optimize the number, location, and type of wells, along with their operating control trajectories in time, and their drilling schedule. The generalized problem formulation aims at jointly optimizing these design parameters, which are traditionally treated as separate optimization problems, by accounting for the coupling between them. We first showed that the source/sink terms in the discretized governing equations of the subsurface flow models include all the information about the type, number, location and control values of the existing wells in the field and hence can serve as a comprehensive generalized decision vector, which inspired the formulation of our generalized field development optimization framework. By including the drilling (capital) cost in the formulation of the NPV cost function, we showed that the resulting optimization is analogous to a sparsity-promoting joint optimization problem. In order to incorporate the time-varying form of the well control settings, we extended this vector into a matrix representation where the rows correspond to the index of grid blocks that intersect a well and the columns contain the dynamic control values of the corresponding wells in time. We employed proximal splitting optimization approach to better handle the capital cost term in the objective function that tends to have a complex behavior. We solved the resulting optimization problem by adopting efficient algorithms from the sparse reconstruction literature, specifically the iterative shrinkage thresholding scheme. The effectiveness of the generalized field development framework and the proposed solution approach was examined through a series of numerical experiments with two-phase fluid flow in heterogeneous reservoir models, including an example adapted from the three-dimensional PUNQ-S3 benchmark formation.

One of the main inspirations of this work is to develop a simple generalized formulation to solve the optimization problems that arise in oilfield development with several types of decision variable. The key in developing such a generalized framework is recognizing the importance of the source and sink vectors in the discretized form of the governing flow equations. The source and sink terms when expanded for the entire simulation time include all the decision variables of interest, i.e. the number, type, location, control trajectory, and drilling schedules of the wells, which can be optimized simultaneously. Clearly, expanding the decision vector to embed all these decision variables increases both the dimension and complexity of the optimization problem. However, it also enables solutions that can be explored to fully harness the optimization potential in field development and, at the very least, reveal the achievable performance of the field, since the generalized formulation is expected to produce a higher NPV than solving each optimization problem independently. While we proposed one approach, i.e. sparse optimization, to solve the formulated problem, alternative solution strategies may also be developed to handle the generalized problem. Our findings suggest that the current framework is worthy of further research and it can be considered as a new direction for formalizing and studying field development planning problems from a more general and unified perspective. The proposed formulation seeks to determine various field development decisions variables simultaneously and by solving a single optimization problem, as opposed to the traditional approach of solving separate optimization problems to determine different development decisions. The latter approach typically leads to suboptimal solutions as

they fail to acknowledge and properly incorporate the coupling between various decision variables that are involved in the development plan.

The proposed framework solves the generalized field development optimization problem for vertical wells. One interesting future direction is to extend the current approach to allow for separate perforations at different layers of the reservoir. Further generalization can also be made by including more complex well completions, such as slanted wells, multilateral wells, and horizontal wells. Such extensions would require effective descriptions of complex well models for optimization purpose and maintaining the feasibility of well model descriptions during optimization iterations where updates are made. Other extensions for field application include incorporation of nonlinear constraints and implementation under geologic uncertainty. The simulation experiments presented and analyzed in this work were of fairly small scales. Additional work is needed to explore the applicability of the developed framework to field-scale problems with realistic constraints.

Appendix A: Proximal forward-backward splitting algorithms

As discussed in Beck and Teboulle [6], iterative shrinkage thresholding algorithm (ISTA) is mainly devised to implement the proximal forward-backward splitting iterative scheme [10, 16, 44]. We explain the proximal forward-backward splitting approach for the following optimization problem:

$$\min_{\mathbf{x}\in\mathbb{R}^N} \quad \underbrace{f_1(\mathbf{x}) + f_2(\mathbf{x})}_{=f(\mathbf{x})} \tag{33}$$

where $f_1(\mathbf{x})$ is lower semi-continuous and convex, and $f_2(\mathbf{x}$ is convex and differentiable with a β-Lipschitz continuous gradient $\nabla f_2(\mathbf{x})$. It is shown in Combettes and Wajs [16] that one solution to the problem in Eq. 33 can be expressed as the *fixed point* equation:

$$\mathbf{x} = \text{prox}_{\beta\, f_1}\,(\mathbf{x} - \beta\nabla f_2(\mathbf{x})) \tag{34}$$

which is valid for any $\beta \in (0, +\infty)$. Equation 34 can be solved [15] iteratively through proximity operation:

$$\mathbf{x}_{n+1} = \underbrace{\text{prox}_{\beta_n\, f_1}}_{\text{backward step}}\ \underbrace{(\mathbf{x}_n - \beta_n \nabla f_2(\mathbf{x}))}_{\text{forward step}} \tag{35}$$

As illustrated in Eq. 35, this scheme can be broken into a forward (explicit) gradient descent update using the f_2 component and a backward (implicit) proximity operation step using the function f_1. Note that, for $f_1 = 0$, Eq. 35 reduces to gradient descent solution update

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \beta_n \nabla f_2(\mathbf{x}) \tag{36}$$

for minimizing a function with a Lipschitz continuous gradient, and for $f_2 = 0$, it takes the form of the iterative proximal mapping of the function f_1:

$$\mathbf{x}_{n+1} = \text{prox}_{\beta_n\, f_1}(\mathbf{x}_n) \tag{37}$$

for minimizing the nondifferentiable function f_1. The resulting forward-backward algorithm is an iterative form that combines the gradient method (forward step) and the proximal point algorithm (backward step).

Appendix B: Derivation of objective function gradient

In this appendix, we derive the expression for the gradient of the objective function $\nabla_{\mathbf{u}} J(\mathbf{u})$ in Table 1 with respect to the well control allocations (entries of matrix $\mathbf{U}$ in Eq. 6). From Eqs. 14–21 , we have:

$$J(\mathbf{u}) = -O(\mathbf{u}) + \underbrace{\sum_{t=1}^{T} \frac{\gamma}{(1+d(t))^{b(t)}} \Big(\sum_{i=1}^{N} |q_{C,i}^t|^p \Big)^{\frac{1}{p}}}_{=C(\mathbf{u})} \tag{38}$$

where $q_{C,i}^t$ is the entry in the i^{th} row and t^{th} column of the $\mathbf{U_C}$ matrix. The gradient of $J(\mathbf{u})$ with respect to each entry q_i^{tj} in matrix $\mathbf{U}$ can be written as:

$$\nabla_{q_i^{tj}} J(\mathbf{u}) = -\nabla_{q_i^{tj}} O(\mathbf{u}) - \nabla_{q_i^{tj}} C(\mathbf{u}) \tag{39}$$

Here, $\nabla_{q_i^{tj}} O(\mathbf{u})$ is obtained using an adjoint model. To get a well-behaved functional form for $C(\mathbf{u})$ that is differentiable everywhere, we use the following approximation to reformulate $C(\mathbf{u})$:

$$|q_{C,i}^t| \simeq \sqrt{(q_{C,i}^t)^2 + \epsilon}\ , \quad \epsilon \ll 1 \tag{40}$$

where ϵ is a small positive number introduced to avoid matrix singularity during optimization iterations, without affecting the numerical accuracy of the algorithm; $\nabla_{q_i^{tj}} C(\mathbf{u})$ can then be expressed as:

$$\nabla_{q_i^{tj}} C(\mathbf{u}) = \frac{\gamma}{(1+d(t))^{b(t)}} \Big(\sum_{k=1}^{N} ((q_{C,k}^t)^2 + \epsilon)^{\frac{p}{2}} \Big)^{\frac{1}{p}-1} \times . \Big(\sum_{k=1}^{N} a_{ki}\ \text{sign}(q_{C,k}^t)\ |q_{C,k}^t|\ ((q_{C,k}^t)^2 + \epsilon)^{\frac{p}{2}-1} \Big) \tag{41}$$

In the above equation, a_{ki} is the entry in the k^{th} row and i^{th} column of the matrix $\mathbf{A}$ that was defined in Eq. 16. In the

above derivation, we used the relation $\mathbf{U_C} = \mathbf{A}\,\mathbf{U}$ to obtain $\mathbf{U_C}$'s entries in terms of the elements of $\mathbf{A}$ and $\mathbf{U}$:

$$q^t_{C,i} = \sum_{k=1}^{N} a_{ik}\, q^t_k \tag{42}$$

Note that when implementing the iterative gradient-based algorithm, $\mathbf{A}$ needs to be updated in each iteration according the updated $\mathbf{U}$.

References

1. Almeida, L.F., Tupac, Y.J., Pacheco, M.A.C., Vellasco, M.M.B.R., Lazo, J.G.L., et al.: Evolutionary optimization of smart-wells control under technical uncertainties. In: Latin American & Caribbean Petroleum Engineering Conference, Society of Petroleum Engineers (2007)
2. Bach, F., Jenatton, R., Mairal, J., Obozinski, G.: Optimization with sparsity-inducing penalties. Foundations and Trends®, in Machine Learning **4**(1), 1–106 (2012)
3. Bangerth, W., Klie, H., Wheeler, M., Stoffa, P., Sen, M.: On optimization algorithms for the reservoir oil well placement problem. Comput. Geosci. **10**(3), 303–319 (2006)
4. Baraniuk, R.G.: Compressive sensing. IEEE Signal Process. Mag. **24**(4) (2007)
5. Barzilai, J., Borwein, J.M.: Two-point step size gradient methods. IMA J. Numer. Anal. **8**(1), 141–148 (1988)
6. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Imag. Sci. **2**(1), 183–202 (2009)
7. Beckner, B., Song, X.: Field development planning using simulated annealing: Optimal economic well scheduling and placement. In: Society of Petroleum Engineers. Annual Technical Conference, pp. 209–221 (1995)
8. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press (2004)
9. Brouwer, D.R.: Dynamic water flood optimization with smart wells using optimal control theory. Delft University of Technology (2004)
10. Bruck, R.E.: On the weak convergence of an ergodic iteration for the solution of variational inequalities for monotone operators in Hilbert space. J. Math. Anal. Appl. **61**(1), 159–164 (1977)
11. Candès, E.J., Wakin, M.B.: An introduction to compressive sampling. IEEE Signal Process. Mag. **25**(2), 21–30 (2008)
12. Centilmen, B., Ertekin, T., Grader, A.: Applications of neural networks in multiwell field development. In: SPE Annual Technical Conference (1999)
13. Chartrand, R.: Exact reconstruction of sparse signals via nonconvex minimization. IEEE Signal Process. Mag. **14**(10), 707–710 (2007)
14. Chartrand, R., Yin, W.: Iteratively reweighted algorithms for compressive sensing. In: IEEE International Conference on Acoustics, Speech, and Signal Processing, 2008. ICASSP 2008, pp. 3869-3872. IEEE (2008)
15. Combettes, P.L., Pesquet, J.C.: Proximal splitting methods in signal processing. In: Fixed-Point Algorithms for Inverse Problems in Science and Engineering, pp. 185–212. Springer (2011)
16. Combettes, P.L., Wajs, V.R.: Signal recovery by proximal forward-backward splitting. Multiscale Modeling & Simulation **4**(4), 1168–1200 (2005)
17. Dai, Y.H., Liao, L.Z.: R-linear convergence of the barzilai and borwein gradient method. IMA J. Numer. Anal. **22**(1), 1–10 (2002)
18. Daubechies, I., Defrise, M., De Mol, C.: An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. Commun. Pure Appl. Math. **57**(11), 1413–1457 (2004)
19. Ding, S., Jiang, H., Li, J., Tang, G.: Optimization of well placement by combination of a modified particle swarm optimization algorithm and quality map method. Comput. Geosci. **18**(5), 747–762 (2014)
20. Donoho, D.L.: Compressed sensing. IEEE Trans. Inf. Theory **52**(4), 1289–1306 (2006)
21. Echeverría Ciaurri, D., Isebor, O.J., Durlofsky, L.J.: Application of derivative-free methodologies to generally constrained oil production optimization problems. Procedia Comput. Sci. **1**(1), 1301–1310 (2010)
22. Emerick, A.A., Silva, E., Messer, B., Almeida, L.F., Szwarcman, D., Pacheco, M.A.C., Vellasco, M.M.B.R., et al.: Well placement optimization using a genetic algorithm with nonlinear constraints. In: SPE Reservoir Simulation Symposium. Society of Petroleum Engineers (2009)
23. Fletcher, R.: The Barzilai Borwein method-steepest descent resurgent. In: Report in the International Workshop on Optimization and Control with Applications. Erice, Italy (2001)
24. Forouzanfar, F., Reynolds, A.: Joint optimization of number of wells, well locations and controls using a gradient-based algorithm. Chem. Eng. Res. Des. (2013)
25. Hale, E.T., Yin, W., Zhang, Y.: A fixed-point continuation method for l1-regularized minimization with applications to compressed sensing. CAAM TR07-07, Rice University (2007)
26. Humphries, T.D., Haynes, R.D.: Joint optimization of well placement and control for nonconventional well types. arxiv preprint arxiv:14094369 (2014)
27. Humphries, T.D., Haynes, R.D., James, L.A.: Simultaneous and sequential approaches to joint optimization of well placement and control. Comput. Geosci. 1–16 (2013)
28. Humphries, T.D., Haynes, R.D., James, L.A.: Simultaneous and sequential approaches to joint optimization of well placement and control. Comput. Geosci. **18**(3-4), 433–448 (2014)
29. Isebor, O.J., Durlofsky, L.J., Ciaurri, D.E.: A derivative-free methodology with local and global search for the constrained joint optimization of well locations and controls. Comput. Geosci., pp. 1–20 (2013)
30. Isebor, O.J., Durlofsky, L.J., Ciaurri, D.E.: A derivative-free methodology with local and global search for the constrained joint optimization of well locations and controls. Comput. Geosci. **18**(3-4), 463–482 (2014)
31. Jansen, J.D.: A systems description of flow through porous media. Springer (2013)
32. Li, L., Jafarpour, B.: A variable-control well placement optimization for improved reservoir development. Comput. Geosci. **16**(4), 871–889 (2012)
33. Li, L., Jafarpour, B., Mohammad-Khaninezhad, M.R.: A simultaneous perturbation stochastic approximation algorithm for coupled well placement and control optimization under geologic uncertainty. Comput. Geosci. **17**(1), 167–188 (2013)
34. Lie, K.A., Krogstad, S., Ligaarden, I.S., Natvig, J.R., Nilsen, H.M., Skaflestad, B.: Open-source matlab implementation of consistent discretisations on complex grids. Comput. Geosci. **16**(2), 297–322 (2012)
35. Luenberger, D.G., Ye, Y.: Linear and nonlinear programming, vol. 116. Springer (2008)
36. Montes, G., Bartolome, P., Udias, A.L., et al.: The use of genetic algorithms in well placement optimization. In: SPE Latin American and Caribbean Petroleum Engineering Conference. Society of Petroleum Engineers (2001)

37. Navabi, S., Khaninezhad, R., Jafarpour, B.: A generalized formulation for oilfield development optimization (2015)
38. Nesterov, Y. et al.: Gradient Methods for Minimizing Composite Objective Function. Tech. rep., UCL (2007)
39. Oliver, D.S., Chen, Y.: Recent progress on reservoir history matching: A review. Comput. Geosci. **15**(1), 185–221 (2011)
40. Oliver, D.S., Reynolds, A.C., Liu, N.: Inverse Theory for Petroleum Reservoir Characterization and History Matching. Cambridge University Press (2008)
41. Onwunalu, J.E., Durlofsky, L.J.: Application of a particle swarm optimization algorithm for determining optimum well location and type. Comput. Geosci. **14**(1), 183–198 (2010)
42. Parikh, N., Boyd, S.: Proximal algorithms. Found. Trends Optim. **1**(3), 123–231 (2013)
43. Parikh, N., Boyd, S.P., et al.: Proximal algorithms. Found. Trends Optim. **1**(3), 127–239 (2014)
44. Passty, G.B.: Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. J. Math. Anal. Appl. **72**(2), 383–390 (1979)
45. Raydan, M.: On the Barzilai and Borwein choice of steplength for the gradient method. IMA J. Numer. Anal. **13**(3), 321–326 (1993)
46. Saad, Y., Van Der Vorst, H.A.: Iterative solution of linear systems in the 20th century. J. Comput. Appl. Math. **123**(1), 1–33 (2000)
47. Sarma, P., Durlofsky, L.J., Aziz, K., Chen, W.H.: Efficient real-time reservoir management using adjoint-based optimal control and model updating. Comput. Geosci. **10**(1), 3–36 (2006)
48. Shirangi, M.G., Durlofsky, L.J., et al.: Closed-loop field development under uncertainty by use of optimization with sample validation. SPE J. (2015)
49. Shu, T., Krunz, M., Vrudhula, S.: Joint optimization of transmit power-time and bit energy efficiency in cdma wireless sensor networks. IEEE Trans. Wirel. Commun. **5**(11), 3109–3118 (2006)
50. Spall, J.C.: Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. IEEE Trans. Autom. Control **37**(3), 332–341 (1992)
51. Spall, J.C., Hill, S.D., Stark, D.R.: Theoretical framework for comparing several stochastic optimization approaches. In: Probabilistic and Randomized Methods for Design under Uncertainty, pp. 99–117. Springer (2006)
52. Turlach, B.A., Venables, W.N., Wright, S.J.: Simultaneous variable selection. Technometrics **47**(3), 349–363 (2005)
53. Van Essen, G., Van den Hof, P., Jansen, J.D., et al.: Hierarchical long-term and short-term production optimization. SPE J. **16**(01), 191–199 (2011)
54. Vlemmix, S., Joosten, G.J., Brouwer, R., Jansen, J.D., et al.: Adjoint-based well trajectory optimization. In: EUROPEC/EAGE Conference and Exhibition. Society of Petroleum Engineers (2009)
55. Wang, C., Li, G., Reynolds, A.C., et al.: Optimal well placement for production optimization. In: Eastern Regional Meeting. Society of Petroleum Engineers (2007)
56. Wright, S.J., Nowak, R.D., Figueiredo, M.A.: Sparse reconstruction by separable approximation. IEEE Trans. Signal Process. **57**(7), 2479–2493 (2009)
57. Yuan, M., Lin, Y.: Model selection and estimation in regression with grouped variables. J. R. Stat. Soc. Ser. B Stat Methodol. **68**(1), 49–67 (2006)
58. Zandvliet, M., Handels, M., van Essen, G., Brouwer, R., Jansen, J.D., et al.: Adjoint-based well-placement optimization under production constraints. SPE J. **13**(04), 392–399 (2008)
59. Zandvliet, M.J.: Model-Based Lifecycle Optimization of Well Locations and Production Settings in Petroleum Reservoirs. Delft University of Technology, TU Delft (2008)
60. Zhang, K., Li, G., Reynolds, A.C., Yao, J., Zhang, L.: Optimal well placement using an adjoint gradient. J. Pet. Sci. Eng. **73**(3), 220–226 (2010)
61. Zhao, P., Rocha, G., Yu, B.: The composite absolute penalties family for grouped and hierarchical variable selection. Ann. Stat., 3468–3497 (2009)