Well control optimization using a two-step surrogate treatment

Daniel U. de Brito¹,², Louis J. Durlofsky¹

1Department of Energy Resources Engineering, Stanford University, Stanford, CA 94305, USA
2Now at Petrôleo Brasileiro S.A., Petrobras

Email addresses: dubrito@stanford.edu (Daniel U. de Brito), lou@stanford.edu (Louis J. Durlofsky)

Abstract

Large numbers of flow simulations are typically required for the determination of optimal well settings. These simulations are often computationally demanding, which poses challenges for the optimizations. In this paper we present a new two-step surrogate treatment (ST) that reduces the computational expense associated with well control optimization. The method is applicable for oil production via waterflood, with well rates optimized at a single control period. The two-step ST entails two separate optimizations, which can both be performed very efficiently. In the first optimization, optimal well-rate ratios (i.e., the fraction of total injection or production associated with each well) are determined such that a measure of velocity variability over the field is minimized, leading to more uniform sweep. In the second step, overall injection and production rates are determined. The flow physics in the first step is highly simplified, while the actual physical system is simulated in the second step. Near-globally-optimal results can be determined in both cases, as the first optimization is posed as a QP problem, and the second step entails just a single optimization variable. Under full parallelization, the overall elapsed time for the ST corresponds to the runtime for 1–2 full-order simulations. Results are presented for multiple well configurations, for 2D and 3D channelized models, and comparisons with formal optimization procedures (mesh adaptive direct search or MADS, and an adjoint-gradient method) are conducted. Three different fluid mobility ratios ($M = 1, 3$ and $5$) are considered. Optimization results demonstrate that the two-step ST provides results in reasonable agreement with those from MADS and adjoint-gradient methods, with speedups of $5 \times$ or more. We also show that the ST is applicable in the inner-loop in field development optimization, where it will be especially useful since many different well configurations must be evaluated.

Keywords: Well control optimization, surrogate model, proxy model, field development optimization, reservoir simulation

1. Introduction

Well control optimization entails defining the set of controls (well rates or bottom-hole pressures) that maximize a predefined cost function, such as net present value or total oil production. Optimization computations generally rely on a reservoir simulator for the evaluation of proposed well control strategies. Since the flow models are typically time dependent and nonlinear, the solution of the resulting nonconvex optimization problem may be computationally demanding. This is even more of an issue when well locations and controls are optimized simultaneously. In such cases, the objective function evaluation for each set of well locations requires the determination of optimal well controls.

In this study, we introduce a new two-step surrogate treatment (ST) for well control optimization for oil fields producing under waterflood. In the first step of the ST, we determine the well-rate ratios (i.e., the fraction of total injection or production associated with each well), for an appropriately defined unit-mobility-ratio problem. In the second step of the ST, we fix the well-rate ratios found in the first step, and then solve a full-physics optimization problem in one variable to determine the overall injection/production rate that maximizes net present value (NPV). Although the problems in the ST are idealized, they are posed such that we can efficiently find near-global optima in both steps (though this does not mean we determine the global optimum for the actual problem).

Well control optimization problems have been addressed using both adjoint-gradient and derivative-free methods. Adjoint-gradient-based methods, which in-
clude those described in [11, 2, 3], are very efficient though they require access to reservoir simulation source code. Derivative-free optimization (DFO) methods represent a useful alternative when gradients are not available. Both stochastic and pattern search DFO methods have been applied in well control optimization problems. Stochastic methods used in this setting include genetic algorithms (GAs) and particle swarm optimization (PSO), which are both population-based evolutionary methods. These procedures were shown to be applicable for both well control [4, 5, 6, 7] and well location optimization [8, 9]. Pattern-search or stencil-based optimization techniques, such as mesh adaptive direct search (MADS) [10], have also been used for well control optimization [11, 12, 13]. In this work both adjoint-gradient-based and MADS procedures will be compared to our new two-stage surrogate treatment.

The use of proxy or surrogate models is an effective way to reduce the cost of optimization in cases when the underlying full-physics model is expensive to simulate. Surrogate treatments have been applied for both well control and well placement optimization problems. For example, in an early study involving field development optimization, Rosenwald and Green [14] used single-phase flow models, for which linear programming techniques are applicable. Subsequent studies that used surrogate models in the form of statistical proxies, for well placement optimization, include those of Guyaguler et al. [15], Yeten et al. [16] and Farmer et al. [17]. Reduced-physics surrogate models were used by Wilson and Durlofsky [18] for optimizing the development of shale gas reservoirs [18], while Aliyev and Durlofsky [19] applied a sequence of upscaled models as surrogates for joint (well location and control) optimization in waterflooding problems. Coarse models were also used as surrogates in well placement optimization by Bukhamsin [20].

Surrogate treatments for well control optimization, based on reduced-order modeling with proper orthogonal decomposition, have been developed by a number of researchers; see, e.g., [21, 22, 23], along with the review by Jansen and Durlofsky [24]. Reduced-physics models based on streamline methods were shown to provide useful surrogates in water injection optimization [25]. Well control optimization applications using streamlines have also been described by Park and Datta-Gupta [26], who used 3D streamlines to generate flood efficiency maps, and by Wen et al. [27], who applied both streamline and time-of-flight (TOF) computations for the optimization of mature fields under waterflood. Flow diagnostics tools [28, 29], typically constructed based on reduced-physics simulations, represent another effective set of surrogate treatments that can be used for optimization. For example, Møyner et al. [30] performed well control optimization for systems involving two- and three-phase flow. Lie et al. [29] used quantities derived from TOF and multiscale methods as fast proxies for well control optimization under waterflood. We note finally that He et al. [31] and Chen et al. [32] have demonstrated that surrogate models can also be applied to accelerate history matching and uncertainty quantification computations.

In this paper, we develop and apply a new two-step surrogate treatment (ST) to enable efficient well control optimization for subsurface flow problems. Although field development optimization is not the focus of this paper, a key applications for our ST is the joint optimization of well locations and controls, where the ST can be used as a fast optimization method for the (inner-loop) well control optimization. The computational savings achieved by ST derive from solving two approximate subproblems rather than the actual full-physics problem. In the first subproblem we simplify the physics, and in the second subproblem we consider the actual physics but greatly simplify the optimization problem. An important advantage of the two-step ST is that near-globally optimum solutions can be obtained for both subproblems at a cost of 1–2 full-order runs under full parallelization. The method provides well controls for only a single control period, but this solution can then be used as an initial guess for cases involving multiple control periods. Geological uncertainty is not considered, though this could be addressed by optimizing over multiple realizations.

This paper is organized as follows. In Sect. 2 we present the basic oil-water flow equations and the general well control optimization problem. In Sect. 3, the two-step surrogate treatment is described in detail. A procedure for handling BHP constraints is also discussed, and the general workflow is explained. In Sect. 4, we present ST well control optimization results for 40 different (waterflood) well scenarios in a 2D model. Well control optimization for additional mobility ratios, and comparisons with optimization results using adjoint-gradient and MADS procedures, are also provided. Results for a 3D case are then presented. Finally, in Sect. 5, we summarize this work and provide suggestions for future research in this area.

2. Problem statement

In this section, we present the equations governing reservoir flow, and we provide a formal statement of the general optimization problem.
2.1. Oil-water flow equations

For two-phase oil-water flow, the governing equations, which derive from Darcy’s law for each phase and mass conservation statements, can be expressed in terms of the so-called pressure and saturation equations. In the absence of capillary pressure, compressibility, and gravity effects, these equations are as follows:

\[ \nabla \cdot (\lambda_t (S_w) \, k \nabla p) = -\tilde{q}_w^w, \]  \hspace{1cm} (1)

\[ \phi \frac{\partial S_w}{\partial t} + \nabla \cdot (f(S_w) \, u_t) = \tilde{q}_w^w. \]  \hspace{1cm} (2)

Here subscripts \( o \) and \( w \) refer to oil and water phases, \( \lambda_t \) is the total mobility, with \( \lambda_t = \lambda_o + \lambda_w \), where \( \lambda_o = k_{ro}/\mu_o \) and \( \lambda_w = k_{rw}/\mu_w \), \( k_{ro} \) and \( k_{rw} \) are relative permeability to the oil and water phases respectively, and \( \mu_o \) and \( \mu_w \) are phase viscosities. Additional variables are water saturation \( S_w \), absolute permeability tensor \( k \), pressure \( p \), total source/sink \( \tilde{q}_w^w \) (the superscript \( w \) here refers to wells), porosity \( \phi \), and time \( t \). The fractional flow of water \( f(S_w) \) is defined as \( f(S_w) = \lambda_w/\lambda_t \), the total velocity as \( u_t = u_o + u_w \) (\( u_o \) and \( u_w \) are the Darcy velocities for oil and water), and \( \tilde{q}_w^w \) is the water source/sink term. We additionally have the saturation constraint \( S_o + S_w = 1 \), which completes the problem statement.

In the discretized system, the source terms \( (\tilde{q}_w^w \) and \( \tilde{q}_w^p \)) are represented using the Peaceman well model. These terms are nonzero only in grid blocks containing wells that are open to flow. The well model relates flow rate to bottom-hole pressure (BHP), and either well rates or BHPs can be specified as the control variables. In this study we consider vertical wells, completed over the entire reservoir thickness, with flow rates as the control parameter. Stanford’s Automatic Differentiation-based General Purpose Research Simulator (AD-GPRS) \[33\] is used for all flow simulations.

2.2. General optimization problem

The general field development optimization problem involves determination of the well types, locations and controls, with the goal of minimizing a cost function \( J \). Following \[34\], the optimization problem can be stated as follows:

\[ \min_{x \in X, u \in U, z \in Z} \quad J(x, u, z), \quad \text{subject to} \begin{cases} g(p, x, u, z) = 0, \\ c(p, x, u, z) \leq 0 \end{cases} \]  \hspace{1cm} (3)

The vectors \( x \) and \( u \) indicate integer well location variables and continuous well control variables, respectively, while \( z \) are categorical variables, which indicate whether the well is an injector \((z_k = -1)\), a producer \((z_k = 1)\), or not drilled at all \((z_k = 0)\). Here \( g = 0 \) denotes the flow simulation equations, \( p \) represents the solution unknowns, which in our system are the pressure and saturation in every grid block, and \( c \) defines any nonlinear constraints. The spaces \( X \) and \( U \) are defined to include bound constraints, which can be expressed as \( x_l \leq x \leq x_u \), \( u_l \leq u \leq u_u \), where subscripts \( l \) and \( u \) denote lower and upper bounds.

In this work the objective is to maximize net present value (NPV); i.e., we set \( J = -\text{NPV} \), with NPV given by:

\[ \text{NPV}(x,u,z) = \sum_{k=1}^{n_w} \sum_{i=1}^{n_t} \Delta t_i \left( p_{iw} \tilde{q}_w^w(x,u) - c_{pw} \tilde{q}_w^w(x,u) \right) \left( 1 + \frac{d}{365} \right), \]  \hspace{1cm} (4)

\[ - \sum_{k=1}^{n_i} \sum_{i=1}^{n_t} \Delta t_i c_{iw} \tilde{q}_w^w(x,u) \left( 1 + \frac{d}{365} \right) - \sum_{k=1}^{n_w} |z_k| c_w. \]  \hspace{1cm} (4)

Here \( n_i \) is the number of injection wells, \( n_p \) is the number of production wells, \( n_w = n_i + n_p \) is the total number of wells, \( n_t \) is the number of simulation time steps, \( t_s \) and \( \Delta t_i \) are the time and time step size at time step \( s \), and \( d \) is the annual discount rate. The price of oil, the cost for handling produced water, and the cost of injected water are, respectively, \( p_{iw}, c_{pw} \) and \( c_w \). The oil and water production rates and the water injection rate, for well \( k \) at time step \( s \), are denoted \( q_{iw}^{k,s}, q_{pw}^{k,s} \) and \( q_{iw}^{k,s} \). The variable \( t_s \) represents the time at which well \( k \) is drilled, and the per-well drilling cost is denoted by \( c_w \).

In the optimizations performed in this study, the well types and locations are fixed (though in one set of examples we extract the well configurations from a field development optimization run). Then, with \( J = -\text{NPV}(u) \), Eq. \[5\] can be expressed as:

\[ \max_{u \in U} \quad \text{NPV}(u), \quad \text{subject to} \begin{cases} g(p, u) = 0, \\ c(p, u) \leq 0 \end{cases} \]  \hspace{1cm} (5)

The space \( U \) again includes the bound constraints for the continuous well control variables, and \( c \) specifies any nonlinear constraints.

We note that, given the nature of the optimization problem, fixing the well locations and types represents a considerable simplification. This specification eliminates the integer well placement variables (these variables are integers because well locations are represented in terms of discrete grid blocks), as well as the categorical well type variables. If these variables are included in the optimization, the overall field development optimization problem (Eq. \[3\] is a mixed integer nonlinear programming (MINLP) problem, which is much more
difficult to solve than the real-variable problem defined in Eq. [5].

The MINLP field development problem has been addressed using both sequential and joint optimization procedures. In the sequential approach, the well types and locations are optimized first, with well controls, or a well control strategy, specified. Once the optimized well placement configuration is determined, the well controls are then optimized for this configuration. However, since the optimal well locations depend on how the wells are operated, the sequential approach will in general provide a suboptimal solution. In the joint optimization approach, well types, locations, and controls are optimized together. This can be accomplished either in a single loop [55], in which all variables are considered simultaneously, or in a nested fashion [56]. In the nested approach, well types and locations are defined in the outer loop, and well controls are optimized in the inner-loop. The ST developed in this study can be used directly in the inner-loop in nested (joint) field development optimization procedures. In this more challenging application, our two-step ST may provide significant computational speedup.

It is very useful for the surrogate model to be constructed such that the global optimum of the resulting optimization problem can be readily found. This will be the case, for example, if the surrogate treatment leads to a linear programming (LP) or a quadratic programming (QP) problem. For general LP or QP problems, a variety of global optimization algorithms are available. These include the Simplex algorithm and its extensions, and interior-point and conjugate-gradient methods [57]. As we will see, the two-step ST developed in this work is indeed compatible with global optimization. We reiterate that this does not mean we find the global optimum for the original full-physics problem, but rather that we can find the global optima for the subproblems considered in the ST.

3. Optimization methodology

In this section we present our two-step surrogate treatment for well control optimization of oil fields under waterflood. Both ST steps are described in detail, and the handling of BHP constraints is also discussed. This treatment considers only a single control period. Although our development here is for production via waterflood, ST is also expected to be applicable for other displacement processes, such as gas or water-alternating-gas injection, though it may be more approximate for these operations.

3.1. Two-step ST

We denote the optimization variables used in this problem as \( \mathbf{f} \in \mathbb{R}^{n_{w} \times 1} \), where \( n_{w} \) is the number of wells, and \( Q \in \mathbb{R} \). Here \( \mathbf{f} \) represents the vector of well-rate ratios, or the fraction of the field rate that is allocated to each well, and \( Q \) specifies the field rate. According to this definition, the actual well rates can be obtained from the following relationship between \( \mathbf{f} \) and \( Q \):

\[
\hat{\mathbf{q}} = \mathbf{f} Q,
\]

where \( \hat{\mathbf{q}} = [(\hat{q}_p^T, (\hat{q}_i^T)]^T \in \mathbb{R}^{n_w \times 1} \), is the vector of well rates, and \( \mathbf{f} = [(f_p^T, (f_i^T)]^T \in \mathbb{R}^{n_w \times 1} \), where superscripts \( p \) and \( i \) indicate production and injection wells.

With this parameterization, the well control optimization problem can be expressed as:

\[
[\mathbf{f}^T, Q]^T = \arg\min_{\mathbf{f} \in F, Q \in \mathbb{R}} J(\mathbf{f}, Q),
\]

where \( \sum_{k=1}^{n_w} f_k^p = \sum_{k=1}^{n_w} f_k^i = 1 \) (with all elements of \( \mathbf{f} \) between 0 and 1), and \( J \) is the objective function of interest (negative NPV in our case).

We address the optimization problem stated in Eq. [7] using a two-step optimization approach. This choice allows us to break the well control optimization problem into two subproblems involving different objective functions. In the first step, we optimize \( \mathbf{f} \):

\[
\mathbf{f}^* = \arg\min_{\mathbf{f} \in F} J_1(\mathbf{f}),
\]

where \( J_1 \) is the first objective function. In the second step we optimize \( Q \) using \( \mathbf{f}^* \) from the first step:

\[
Q^* = \arg\min_{Q \in \mathbb{R}} J_2(\mathbf{f}^*, Q),
\]

where \( J_2 \) is the second objective function.

The objective function used in the determination of optimal well-rate ratios corresponds to a measure of velocity variability. By minimizing this quantity, a more uniform sweep is achieved. This minimization is formulated as a QP problem, under the assumptions of unit mobility ratio \( (M = 1) \) and unit voidage replacement ratio \( (VRR = 1) \). Given the optimal well-rate ratios \( \mathbf{f}^* \), \( Q \) is determined in the second subproblem, by running a limited number of full-physics simulations, such that the NPV of the actual problem is maximized. We now describe the two subproblems in turn.
3.1.1. First subproblem: optimization of \( f \)

We introduce significant approximations, in the first step of the ST, relative to the (target) optimization of the full-physics problem. Specifically, our goal in the first step is to minimize a measure of velocity variation within the model, which we expect to lead to improved sweep and thus higher oil production and NPV (this expectation will be confirmed below). In this step we assume unit mobility ratio, which results in a linear flow equation. We note that other simulation-based approaches such as flow diagnostics [30, 29] could also be used in this step, though it is not clear that such approaches such as flow diagnostics [30, 29] could also accelerate considerably the calculation of such quantities when well rates are changed. Specifically, the velocity response matrices \( V_d \in \mathbb{R}^{n_b \times n_w} \), where \( n_b \) is the number of grid blocks, enable the resulting velocity fields \( \hat{u}_d \in \mathbb{R}^{n_b \times 1} \) to be written as linear combinations of the specified well rates \( \hat{q} \):

\[
\hat{u}_d = V_d \hat{q},
\]

where \( d = x, y, z \).

The methodology to build the response matrices consists of running a number of single-phase flow simulations to obtain a series of velocity coefficients. These coefficients are simply the velocity field components at every grid block of the model resulting from a single well producing at a unit flow rate. For well control optimization problems, where the well number, type, and locations are fixed, the number of single-phase flow simulations required is equal to the number of wells \( n_w \). In each simulation, one of the wells operates individually, at a constant flow rate, until the reservoir pressure reaches pseudo-steady state (PSS); i.e., until \( \frac{\partial p}{\partial t} \) becomes a constant. When this occurs, the velocity components are saved as columns in the velocity response matrices.

Note that it is also possible (and faster) to compute the PSS solution directly rather than through the time-integration of Eq. 10. The direct PSS solution can be computed by replacing the right-hand side of Eq. 10 by \( \phi K \), where \( K \) is an arbitrary constant. We use time integration here since AD-GPRS is applied for these computations, and this simulator does not currently provide the PSS solution directly.

Figure 1 depicts the assembly of the velocity response matrix for a \( 3 \times 2 \) model with wells located at grid blocks 2 and 4. The two model setups in Fig. 1,b indicate the well configurations for the two simulation runs required to build the response matrices. We denote \((v_d)_{ij} \), an element of \( V_d \), to be the velocity component

The velocity components at grid block centers can be obtained by averaging the velocity components at the faces:

\[
(u_x)_{i,j,k} = \frac{(u_x)_{i-\frac{1}{2},j,k} + (u_x)_{i+\frac{1}{2},j,k}}{2}.
\]

We denote the velocity components for grid block \( i, j, k \) as \((u_d)_{i,j,k} \), with \( d = x, y, z \).

The linearity of Eq. 10 implies that both the pressure and the velocity fields can be expressed as a superposition of the responses of individual wells [38, 39]. This also allows the use of response matrices for pressure and velocity computations [14, 38], which can accelerate considerably the calculation of such quantities when well rates are changed. Specifically, the velocity response matrices \( V_d \in \mathbb{R}^{n_b \times n_w} \), where \( n_b \) is the number of grid blocks, enable the resulting velocity fields \( \hat{u}_d \in \mathbb{R}^{n_b \times 1} \) to be written as linear combinations of the specified well rates \( \hat{q} \):

\[
\hat{u}_d = V_d \hat{q},
\]

where \( d = x, y, z \).

The methodology to build the response matrices consists of running a number of single-phase flow simulations to obtain a series of velocity coefficients. These coefficients are simply the velocity field components at every grid block of the model resulting from a single well producing at a unit flow rate. For well control optimization problems, where the well number, type, and locations are fixed, the number of single-phase flow simulations required is equal to the number of wells \( n_w \). In each simulation, one of the wells operates individually, at a constant flow rate, until the reservoir pressure reaches pseudo-steady state (PSS); i.e., until \( \frac{\partial p}{\partial t} \) becomes a constant. When this occurs, the velocity components are saved as columns in the velocity response matrices.

Note that it is also possible (and faster) to compute the PSS solution directly rather than through the time-integration of Eq. 10. The direct PSS solution can be computed by replacing the right-hand side of Eq. 10 by \( \phi K \), where \( K \) is an arbitrary constant. We use time integration here since AD-GPRS is applied for these computations, and this simulator does not currently provide the PSS solution directly.

Figure 1 depicts the assembly of the velocity response matrix for a \( 3 \times 2 \) model with wells located at grid blocks 2 and 4. The two model setups in Fig. 1,b indicate the well configurations for the two simulation runs required to build the response matrices. We denote \((v_d)_{ij} \), an element of \( V_d \), to be the velocity component

\[
\hat{u}_d = V_d \hat{q},
\]

where \( d = x, y, z \).

The methodology to build the response matrices consists of running a number of single-phase flow simulations to obtain a series of velocity coefficients. These coefficients are simply the velocity field components at every grid block of the model resulting from a single well producing at a unit flow rate. For well control optimization problems, where the well number, type, and locations are fixed, the number of single-phase flow simulations required is equal to the number of wells \( n_w \). In each simulation, one of the wells operates individually, at a constant flow rate, until the reservoir pressure reaches pseudo-steady state (PSS); i.e., until \( \frac{\partial p}{\partial t} \) becomes a constant. When this occurs, the velocity components are saved as columns in the velocity response matrices.

Note that it is also possible (and faster) to compute the PSS solution directly rather than through the time-integration of Eq. 10. The direct PSS solution can be computed by replacing the right-hand side of Eq. 10 by \( \phi K \), where \( K \) is an arbitrary constant. We use time integration here since AD-GPRS is applied for these computations, and this simulator does not currently provide the PSS solution directly.

Figure 1 depicts the assembly of the velocity response matrix for a \( 3 \times 2 \) model with wells located at grid blocks 2 and 4. The two model setups in Fig. 1,b indicate the well configurations for the two simulation runs required to build the response matrices. We denote \((v_d)_{ij} \), an element of \( V_d \), to be the velocity component

\[
\hat{u}_d = V_d \hat{q},
\]

where \( d = x, y, z \).

The methodology to build the response matrices consists of running a number of single-phase flow simulations to obtain a series of velocity coefficients. These coefficients are simply the velocity field components at every grid block of the model resulting from a single well producing at a unit flow rate. For well control optimization problems, where the well number, type, and locations are fixed, the number of single-phase flow simulations required is equal to the number of wells \( n_w \). In each simulation, one of the wells operates individually, at a constant flow rate, until the reservoir pressure reaches pseudo-steady state (PSS); i.e., until \( \frac{\partial p}{\partial t} \) becomes a constant. When this occurs, the velocity components are saved as columns in the velocity response matrices.

Note that it is also possible (and faster) to compute the PSS solution directly rather than through the time-integration of Eq. 10. The direct PSS solution can be computed by replacing the right-hand side of Eq. 10 by \( \phi K \), where \( K \) is an arbitrary constant. We use time integration here since AD-GPRS is applied for these computations, and this simulator does not currently provide the PSS solution directly.

Figure 1 depicts the assembly of the velocity response matrix for a \( 3 \times 2 \) model with wells located at grid blocks 2 and 4. The two model setups in Fig. 1,b indicate the well configurations for the two simulation runs required to build the response matrices. We denote \((v_d)_{ij} \), an element of \( V_d \), to be the velocity component

\[
\hat{u}_d = V_d \hat{q},
\]

where \( d = x, y, z \).

The methodology to build the response matrices consists of running a number of single-phase flow simulations to obtain a series of velocity coefficients. These coefficients are simply the velocity field components at every grid block of the model resulting from a single well producing at a unit flow rate. For well control optimization problems, where the well number, type, and locations are fixed, the number of single-phase flow simulations required is equal to the number of wells \( n_w \). In each simulation, one of the wells operates individually, at a constant flow rate, until the reservoir pressure reaches pseudo-steady state (PSS); i.e., until \( \frac{\partial p}{\partial t} \) becomes a constant. When this occurs, the velocity components are saved as columns in the velocity response matrices.

Note that it is also possible (and faster) to compute the PSS solution directly rather than through the time-integration of Eq. 10. The direct PSS solution can be computed by replacing the right-hand side of Eq. 10 by \( \phi K \), where \( K \) is an arbitrary constant. We use time integration here since AD-GPRS is applied for these computations, and this simulator does not currently provide the PSS solution directly.

Figure 1 depicts the assembly of the velocity response matrix for a \( 3 \times 2 \) model with wells located at grid blocks 2 and 4. The two model setups in Fig. 1,b indicate the well configurations for the two simulation runs required to build the response matrices. We denote \((v_d)_{ij} \), an element of \( V_d \), to be the velocity component

\[
\hat{u}_d = V_d \hat{q},
\]

where \( d = x, y, z \).

The methodology to build the response matrices consists of running a number of single-phase flow simulations to obtain a series of velocity coefficients. These coefficients are simply the velocity field components at every grid block of the model resulting from a single well producing at a unit flow rate. For well control optimization problems, where the well number, type, and locations are fixed, the number of single-phase flow simulations required is equal to the number of wells \( n_w \). In each simulation, one of the wells operates individually, at a constant flow rate, until the reservoir pressure reaches pseudo-steady state (PSS); i.e., until \( \frac{\partial p}{\partial t} \) becomes a constant. When this occurs, the velocity components are saved as columns in the velocity response matrices.

Note that it is also possible (and faster) to compute the PSS solution directly rather than through the time-integration of Eq. 10. The direct PSS solution can be computed by replacing the right-hand side of Eq. 10 by \( \phi K \), where \( K \) is an arbitrary constant. We use time integration here since AD-GPRS is applied for these computations, and this simulator does not currently provide the PSS solution directly.
Figure 1: Velocity response matrix for a $3 \times 2$ simulation model with two wells located at grid blocks 2 and 4.

Even though the $V_d$ matrices are constructed from pseudo-steady state solutions considering one well at a time, the multwell response obtained through superposition will reproduce the steady-state solution as long as $V_{RR} = 1$. This is now demonstrated by comparing the velocity fields generated through direct simulation (i.e., by solution of Eq. (10) until steady state is reached) and by superposition for a $100 \times 100$ model containing four production and two injection wells. The channelized model and well locations are shown in Fig. 2. The resulting velocity fields in the $x$ and $y$-directions, as well as cross plots comparing the two solutions, are presented in Fig. 3. The velocity fields constructed via direct simulation and by superposition coincide very closely. The average absolute errors for velocities computed through superposition are about $7 \times 10^{-5}$ for both components. Although the concept and use of velocity response matrices is not new [38], we are not aware of this approach being used before to construct steady-state solutions from the superposition of individual well responses obtained at pseudo-steady state.

The fact that the velocity field can be expressed as a linear combination of the well rates, using Eq. (14), means that the minimization of the (weighted) squared velocities can be formulated as a QP problem. As indicated earlier, through use of such a minimization, we can find the set of well rates $f$ that result in a more uniform sweep and thus (expected) higher oil recovery and/or NPV. The general QP minimization can be stated as follows:

$$\min_{f \in \mathbb{F}} \sum_{k=1}^{n_p} f_k^p \mathbf{W}_x \mathbf{u}_x + \sum_{k=1}^{n_i} f_k^i \mathbf{W}_y \mathbf{u}_y + \sum_{k=1}^{n_z} f_k^z \mathbf{W}_z \mathbf{u}_z,$$

subject to $\sum_{k=1}^{n_p} f_k^p = \sum_{k=1}^{n_i} f_k^i = \sum_{k=1}^{n_z} f_k^z = 1, \quad f_k^p, f_k^i, f_k^z \geq 0, \quad f \in \mathbb{R}$,

(15)

where $\mathbf{W}_d \in \mathbb{R}^{n_x \times n_y}$ with $d = x, y, z$ is a diagonal matrix whose elements are positive weighting factors. The field rate $Q$ is arbitrary at this point, and here we set $Q = 1$. In the results below we specify $\mathbf{W}_d = I$ for simplicity. However, by setting the elements of $\mathbf{W}_d$ more generally, we can emphasize velocity uniformity in particular areas of interest in the reservoir. This could be
beneficial, for example, when new wells are introduced into a reservoir that has already been under production for some time.

Solution of the optimization problem in Eq. 15 provides well-rate ratios that minimize the sum of the resulting squared velocities over the entire model domain. This QP problem involves a quadratic objective function subject to linear constraints, and it can be solved efficiently using well-established optimization techniques that guarantee that a global optimum is found (as discussed in Sect. 2.2). The use of the response matrix approach enables the very fast computation of \( \hat{u} \), i.e., we do not need to solve any flow equations, thus accelerating the determination of \( \mathbf{f}^\ast \). Note that the minimization of the actual velocity variance, rather than the squared velocity, also corresponds to a QP problem. We implemented both of these minimizations, and found that the minimization of squared velocity provided better results for the cases evaluated. It is possible that squared velocity penalizes high velocities more strongly, which leads to more uniform sweep and thus higher NPVs.

The computational cost for constructing the response matrix and solving the optimization problem in Eq. 15 is around 30–40% of a single full-physics AD-GPRS simulation for the cases considered here. However, as noted earlier, the PSS problem could be solved directly, without integrating in time. Such a treatment, though not implemented in this work, will decrease the cost of this step to just a few percent of a full-physics AD-GPRS run.

3.1.2. Second subproblem: optimization of \( Q \)

Once the optimal well-rate ratios \( \mathbf{f}^\ast \) have been determined, the actual well rates become a linear function of the field rate \( Q \), as indicated by Eq. 6. Thus, the optimization now involves only a single variable. At this stage we consider the actual objective function (in this case NPV as defined in Eq. 4) and we simulate the full-physics system.

The second-stage optimization problem can be writ-
ten as:

$$\max_{Q \in \mathbb{R}} \text{NPV}(\mathbf{f}^*, Q), \quad \text{subject to } Q > 0. \quad (16)$$

At this stage we do not require VRR = 1 or the assumption of slight compressibility, though it is expected that the overall ST will perform better when these conditions are approximated. We treat the problem defined in Eq. (16) very simply. Specifically, we define a range for $Q$ such that, over the simulation time frame, between 0.5 pore volumes injected (PVI) and 2.5 PVI are introduced into the model from all of the wells. We compute even increments in PVI based on $n_{\text{proc}}$, where $n_{\text{proc}}$ is the number of available processors, and then simulate each $Q$ value on a different processor. For example, for $n_{\text{proc}} = 41$, we simulate cases corresponding to 0.5, 0.55, 0.60, ..., 2.45, 2.5 PVI. This provides the green points shown in Fig. 4 in an elapsed (wall-clock) time corresponding to the time required for one full-physics simulation. We then take the $Q$ that provides the best NPV to be $Q^*$. We could of course iterate this procedure further, by running additional cases in the most promising PVI regions, to provide the global optimum (red point in Fig. 4). Given the level of approximation inherent in the overall surrogate treatment, however, this additional iteration may not be necessary, assuming a sufficient number of processors are available.

At this point it is useful to review the treatments and approximations introduced into our two-stage ST. In the first stage, we minimize a measure of velocity variability under the assumption of $M = 1$ and VRR = 1. The objective function used in this minimization is not the actual objective function, though it does correspond to a quantity that is expected to impact displacement efficiency. Then, given the optimized well-rate ratios from this first step, in the second step we maximize the actual objective function (NPV) using the full-physics model. Global optima can be determined for both problems – the first-stage optimization is a QP problem, and the second-stage optimization involves only a single variable. Since the second-stage problem can be solved in an elapsed time corresponding to a single full-physics run, and given that the first-stage problem can be solved even faster, the elapsed time for an optimization using our ST is the time required for 1–2 full-physics runs.

### 3.2. BHP constraint handling techniques

The two-step ST described above did not include the treatment of any constraints. In the case of well-rate specifications, as are applied here, the most important constraint is expected to be a BHP constraint. We now discuss how this constraint is handled.

We are concerned mainly with cases for which $\nu_w \gtrsim \mu_o$, and thus $M \gtrsim 1$. In such cases, less viscous (more mobile) fluid is injected into the reservoir, so the differences in pressure between injectors and producers tend to decrease with time. Therefore, if BHP constraint violations do occur, this will often be at the start of the simulation. These constraint violations will be detected in the second stage of the ST, where full-physics simulations are performed to determine $Q^*$. In such cases, we specify the BHP for the affected well(s), at the maximum or minimum, and then reapportion the remaining injection and production. Specifically, well-rate ratios in this case are assigned as:

$$\left(f_k^p\right)^* = \frac{f_k^p}{\sum_{k=1}^{n_p} f_k^p}, \quad (17)$$

where $\left(f_k^p\right)^*$ is the updated well-rate ratio for production well $k$, and $n_p$ is the number of production wells that operate under rate control. An analogous procedure is applied to the injection wells when a BHP constraint violation occurs at the start of the simulation.

The well-rate distributions now correspond to rescaled $\mathbf{f}^*$ for those wells that remain under rate control. BHP constraints are satisfied for the other wells. We again vary rate $Q$ and compute NPV in the second
ST stage. The resulting curve, which differs from that of the unconstrained case, is shown in Fig. 5. Note that the optimal rate $Q^*$ calculated in this step is not the actual total rate of the field. Rather, it is the total rate of the wells operating under rate control. The total rate of the field is the sum of $Q^*$ and the rate injected or produced by the wells operating under BHP control. This treatment does not impact the efficiency of the second-stage optimization – it can still be achieved, in an elapsed time of 1–2 full-order simulations, as described above.

![Figure 5: Relationship between NPV and PVI, used to determine the optimal field rate $Q^*$ in the second subproblem, for unconstrained (upper curve) and BHP-constrained (lower curve) cases. The points denote the optima.](image)

BHP violations may also occur later in the simulation run, though this is less of an issue for $M \gtrsim 1$, as noted earlier. If this happens, the well is switched to BHP control, and it remains on BHP control for the rest of the run. No additional corrective measures are taken, and the other wells are kept on rate control.

We note finally that other treatments for satisfying BHP constraints could also be evaluated. One potential approach is to decrease the $f$ value associated with wells that violate the constraint until the constraint is satisfied. The fraction of injection or production ‘removed’ from these wells would be apportioned to the other wells. We have not evaluated additional treatments for BHP constraint handling since we have found that, within the context of field development optimization, these constraint violations often occur when wells are located in poor reservoir regions, which means that such solutions will usually not be considered further by the optimizer.

### 3.3. General workflow

The two-step ST formulation entails the assembly of the velocity response matrices followed by the two optimization procedures. The Stanford Smart Fields Unified Optimization Framework (UOF) \cite{41,42} is used to test and compare the different optimization methods and treatments.

As explained earlier, for the construction of the velocity response matrices, a single-phase flow simulation is performed for each well, with the well specified to operate at constant flow rate. These simulations are run until the reservoir reaches pseudo-steady state. Once this is achieved, the pressure field is recorded, Darcy velocities are computed, and the velocity response matrices are assembled.

Algorithm 1 below describes the determination of $F^*$ and $Q^*$ using the two-stage ST. We first solve the squared-velocity-minimization problem (Eq. 15) under the assumption of $M = 1$ and VRR = 1. Then, NPV for the full-physics problem is maximized (Eq. 16). The QP problem in the first step is solved in Matlab \cite{43} using CVX, a package for solving convex optimization problems \cite{44}. The second optimization step requires running $n_{proc}$ full-physics simulations, where $n_{proc}$ is the number of computing nodes available (thus achieving full parallelization). If the full-physics model corresponds to $M = 1$, MRST can be used to accelerate the second-stage computations.

### 4. Example cases

In this section, we apply the two-step ST described in Sect. 3 to optimize well controls for 2D and 3D reservoir models. Performance is assessed for cases involving both single and multiple control periods. In the latter case, the single control period solution is used as the initial guess. The examples consider $M = 1, 3$ and $5$, where $M = \mu_0/\mu_v$. For $M = 3$ and $5$, we use the relative permeability curves shown in Fig. 6.

For $M = 1$, we specify $k_{ro} = S_o$, $k_{rw} = S_w$ and $\mu_w = \mu_v = 1$. Thus the pressure and velocity fields for $M = 1$ cases correspond exactly to those for single-phase flow. BHP constraints are imposed in these examples, but they are not violated at the start of the simulation in the cases presented. As noted earlier, Stanford’s Automatic Differentiation-based General Purpose Research Simulator (AD-GPRS) \cite{43} is used for all simulations.

Conventional optimization, using adjoint-gradients with SNOPT \cite{45} and pattern search with MADS \cite{10}, is also performed in this study to enable comparison with ST results. In nonconvex problems, such as the well control optimization considered here, these procedures are only guaranteed to find local optima, with the result depending on the initial guess. For this reason, five runs with different initial guesses are performed for adjoint-gradient and MADS optimizations in all cases where these methods are applied.
Algorithm 1. Well-rate ratios and field rate optimization

1. Solve first-step optimization problem to provide \( f^* \) (Matlab/CVX) (Eq. 15)
2. Solve second-step optimization problem to provide \( Q^* \) (Eq. 16):
   
   for number of computing nodes available do
   4. Run full-physics simulation with different field rates (AD-GPRS)
   5. Treat BHP constraints as described in Sect. 3.2
   end for
   7. Choose field rate associated with largest NPV

Figure 6: Relative permeability curves used for \( M = 3 \) and \( M = 5 \) simulations. Straight-line relative permeabilities are used for \( M = 1 \) cases.

In this study, adjoint-gradient optimizations are terminated when the gradient satisfies a threshold of 0.001, and the optimization converges to a local optimum, or after a maximum of 15 major iterations have been performed. MADS is terminated if the stencil size decreases beyond 0.01, or after a maximum of 15 MADS iterations (for 2D models) or 20 MADS iterations (for 3D models) have been performed. Each MADS iteration requires \( 2n_{opt} \) function evaluations (full-physics simulations in our case), where \( n_{opt} \) is the number of optimization variables [10]. Alternative stopping criteria for MADS, such as those introduced by Boch [46], could also be considered.

4.1. 2D reservoir model

Our first example involves the 2D reservoir model shown in Fig. 2 (this model was generated by [34]). This model, of dimensions 100 × 100, represents a complex channel-levee shale fluvial system. Relevant simulation and optimization parameters are provided in Table 1. In our first set of examples, only a single control period (with well rates as the control variables) is used, so we have \( n_{opt} = n_w - 1 \).

4.1.1. Results for different well scenarios

Although field development optimization is not the main focus of this work, we first illustrate the potential use of the two-step ST within this context. This enables us to assess ST performance for a wide range of well configurations. These configurations are those actually generated during the course of a field development optimization run. Specifically, we ran the particle swarm optimization (PSO) algorithm described in [34], with a PSO swarm size of 40. The goal was to maximize NPV by optimizing the number, type, locations and controls of a set of wells, with \( M = 3 \). We extracted the configurations (well locations and types) corresponding to the 40 particles at a relatively early stage of the optimization – specifically at iteration 25 (an early iteration was considered to maintain diversity within the swarm). We then applied ST for each of the 40 configurations. ST results are compared to optimization using MADS for the same sets of wells.

Figure 7 displays a cross-plot of the NPVs obtained by ST and MADS. Each point represents one of the 40 PSO well scenarios. Relatively close correspondence between the ST and MADS NPVs is observed (correla-
tion coefficient $R^2 = 0.94$), with most points falling near the 45-degree line. These results demonstrate that the two-step ST is able to generate reasonable optimization results for a broad NPV range (the NPVs span nearly an order of magnitude). In other words, the two-step ST is able to perform satisfactorily for either poor, average or favorable sets of well types and locations. This is an important capability if the method is to be used for field development optimization, in which case a large fraction of the proposed well locations are far from optimal, and it is important that they be assessed accurately.

We now select three of the well configurations for further assessment. These scenarios are represented by the red points in Fig. 7 and the corresponding well configurations are shown in Fig. 8a,b,c. It is evident from Fig. 7 that these cases cover a broad NPV range; i.e., we do not consider only the most promising well configurations.

As a means to assess the degree of improvement achieved by the two-step ST (and by MADS and adjoint-gradient-based optimization), we also evaluate NPV for three base cases for each well scenario. The base cases entail equally distributed injection and production rates; i.e., $f_k^i = 1/n_p$ for $k = 1, \ldots, n_p$ and $f_k^p = 1/n_i$ for $k = 1, \ldots, n_i$, for field rates that cover a range of PVI values. We ensure that the range includes the $Q^*$ value provided by ST.

The base-case NPVs are presented in Table 2. The $Q^*$ value for each case is also given. These results can then be compared to those in Table 3 where we show optimization results for each configuration for three different mobility ratios. This table includes ST results along with those using MADS and adjoint gradients. Results for the latter two procedures are median NPVs from five separate runs. The minimum and maximum NPVs for these methods are shown in Table 4. We reiterate that ST provides (essentially) global optima in both steps, so it does not need to be run multiple times to avoid poor local optima.

Base case and ST results are also compared, in terms of NPV versus time plots for each well configuration (for $M = 3$), in Fig. 8d,e,f. These results, along with those in Tables 2 and 3, demonstrate the clear improvement provided by our two-step ST. This improvement is due both to the optimization of well-rate ratios and to the optimization of $Q$, with the relative impact of these two steps varying from case to case. Of essential interest is the close correspondence between ST optimization results and those from MADS and adjoint gradients (Table 3). The largest relative discrepancy is for well scenario 2, $M = 5$, where we see ST underperform MADS by 5.6%. The close correspondence between results from our two-step ST and those from formal optimization methods is very encouraging, and suggests that the surrogate treatments may indeed capture important aspects of the true full-physics optimization problem.

The results in Table 4 demonstrate that MADS achieves very similar results in all five runs for each case. This is also generally observed for the adjoint-gradient procedure, though the well scenario 3, $M = 5$ case is an exception.

Elapsed (wall-clock) times for the optimizations are shown in Table 5. These values are median timings for the five MADS and adjoint-gradient runs (recall ST need only be run once). Individual full-physics simulation runs for these cases require $\sim$40–50 seconds. We see that ST, as noted earlier, entails an elapsed time corresponding to 1–2 full-physics runs. MADS requires five or more iterations for these cases, so the speedup using ST is fairly significant (speedup factor of 6× or more). Speedup relative to adjoint-gradient-based optimization is larger because this procedure does not read-

![Figure 7: Comparison of NPVs obtained by ST and MADS for the 40 well scenarios extracted from a PSO field development optimization run ($M = 3$). The red points indicate the three well scenarios used for further comparisons.](image-url)

### Table 2: NPVs for base cases ($M = 3$) for 2D reservoir model

| Well scenario | NPV (BN $) |
|---------------|------------|
| PVI = 0.25    | 0.54 0.48 0.15 |
| PVI = 0.75    | 1.22 1.41 |
| PVI = 1.25    | 0.81 0.89 0.62 |

1 $Q^*_{1, M = 3} = 0.49$ PVI
2 $Q^*_{2, M = 3} = 0.41$ PVI
3 $Q^*_{3, M = 3} = 0.65$ PVI
Figure 8: Well configurations and permeability field, and NPV evolution in time, for the three cases indicated by the red points in Fig. In (a, b, c), production wells are shown in red and injection wells in blue. Curves in (d, e, f) correspond to ST results (blue curves) and base-case results for 0.25 PVI, 0.75 PVI and 1.25 PVI. M = 3 in all cases.

Table 3: Optimal NPVs using different methods (2D reservoir model)

| Well scenario | Mobility ratio | NPV (BN $) | ST | MADS | Adjoint-gradient |
|---------------|----------------|------------|-----|------|-----------------|
| 1             | 1              | 1.23       | 1.24| 1.23 |
|               | 3              | 0.62       | 0.62| 0.62 |
|               | 5              | 0.47       | 0.47| 0.47 |
| 2             | 1              | 2.03       | 2.02| 2.07 |
|               | 3              | 1.03       | 1.09| 1.02 |
|               | 5              | 0.84       | 0.89| 0.89 |
| 3             | 1              | 2.55       | 2.64| 2.55 |
|               | 3              | 1.47       | 1.47| 1.47 |
|               | 5              | 1.28       | 1.29| 1.29 |

Table 4: Minimum and maximum NPVs for MADS and adjoint-gradients (2D reservoir model)

| Well scenario | Mobility ratio | NPV (BN $) | ST | MADS | Adjoint-gradient |
|---------------|----------------|------------|-----|------|-----------------|
| 1             | 1              | 1.23       | 1.24| 1.24 |
|               | 3              | 0.62       | 0.62| 0.62 |
|               | 5              | 0.47       | 0.47| 0.47 |
| 2             | 1              | 2.03       | 2.02| 2.07 |
|               | 3              | 1.03       | 1.09| 1.02 |
|               | 5              | 0.84       | 0.89| 0.89 |
| 3             | 1              | 2.55       | 2.64| 2.55 |
|               | 3              | 1.47       | 1.47| 1.47 |
|               | 5              | 1.28       | 1.29| 1.29 |

Median values

The results in this section demonstrate that our two-step surrogate treatment provides reasonable accuracy relative to formal optimization methods, and that it can lead to significant speedup. As such, it may be useful for both standalone well control optimization applications, and as an inner-loop in field development optimization.
Table 5: Elapsed time (fully parallelized for ST and MADS) for different methods for 2D reservoir model

| Well scenario | Mobility ratio | Elapsed time (s) | ST          | MADS¹        | Adjoint-gradient¹ |
|---------------|----------------|------------------|-------------|--------------|------------------|
| 1             | 1              | 52               | 389         | 1652         |                   |
| 1             | 3              | 52               | 385         | 1952         |                   |
| 5             | 51             | 389              | 3652        |              |                   |
| 5             | 60             | 389              | 652         |              |                   |
| 1             | 63             | 503              | 876         |              |                   |
| 5             | 63             | 503              | 2538        |              |                   |
| 5             | 66             | 502              | 3690        |              |                   |

¹Median values

The latter point, highlighted by the ST results for many different configurations (Fig. 7), will be exploited in a future study. We next consider the application of ST for a 3D example.

4.2. 3D reservoir model

The 3D model used in this assessment is shown in Fig. 9. This 40 \times 30 \times 9 model (10,800 total grid blocks) represents a prograding fluvial channel system, and corresponds to a portion of a synthetic geological model developed by Castro [47]. We assess ST performance for two different scenarios involving (1) three producers and three injectors, and (2) four producers and two injectors. All wells are completed over the full reservoir thickness. The two scenarios are shown in Fig. 10. Both single and multiple control periods are considered for this case.

Relevant simulation and optimization parameters appear in Table 6. We again consider \( M = 1, 3 \) and 5, and use the same relative permeability curves as in the 2D cases (the curves in Fig. 6 for \( M \neq 1 \), and straight-line relative permeabilities for \( M = 1 \)).

4.2.1. Single control period

Optimization results for the two scenarios shown in Fig. 10 will now be presented. We first consider the case where well rates are fixed for the entire simulation time frame (as was the treatment in all of the 2D runs). As in the 2D case, several base cases, with equal distribution of rates among all injection and all production wells, are assessed for \( M = 3 \).

The base-case NPVs are shown in Table 7, and the optimized NPVs using ST and MADS are presented in Table 8. The median, minimum and maximum NPVs using MADS are provided. Adjoint-gradient-based optimizations were not performed for this case, though we would expect these results to closely coincide with the MADS results, as in the 2D example.

As in the 2D cases, we again see that the NPVs provided by ST clearly exceed those for the three base cases for both scenarios. Base-case and ST results for cumulative field oil and water production and cumulative water injection, for well scenario 2 with \( M = 3 \), are shown in Fig. 11. Results are also presented for total injection \( Q^* \), as determined by ST, but with equal injection and production between wells. The ST results display the most oil production and, interestingly, no water production, in contrast to results for the other cases. These results highlight the importance of optimizing both \( f \) and \( Q \).

Table 6: Simulation and optimization parameters for 40 \times 30 \times 9 channelized model

| Grid dimensions | \( 40 \times 30 \times 9 \) |
|-----------------|--------------------------|
| Grid cell dimensions | 100 ft \times 100 ft \times 15 ft |
| Initial pressure \( p_i \), at datum | 6000 psi at 6000 ft |
| \( \mu_o \) | 1.0, 3.0 or 5.0 cp |
| \( \mu_w \) (constant) | 1.0 cp |
| \( \rho_o \) and \( \rho_w \) | 53.10 and 64.79 lbm/ft³ |
| \( B_o \) and \( B_w \) at \( p_i \) | 1.075 RB/STB |
| Oil compressibility at \( p_i \) | \( 10^{-5} \) psi⁻¹ |
| Water compressibility at \( p_i \) | \( 10^{-5} \) psi⁻¹ |
| Rock compressibility at \( p_i \) | \( 10^{-9} \) psi⁻¹ |
| Field porosity | 0.25 |
| Simulation period | 14,600 days |
| \( p_o \), \( c_{pw} \) and \( c_{iw} \) | $60, $5 and $5/STB |
| \( c_w \) | $20 MM |
| Discount rate | 0% |

Table 7: NPVs for base cases (\( M = 3 \)) for 3D reservoir model with a single control period

| Well scenario | NPV (BN $) |
|---------------|------------|
| \( M = 3 \) | \( M = 3 \) |
| \( 1^1 \) | 3.71 | 1.5 | 2.0 |
| \( 2^2 \) | 3.69 | 5.14 | 6.61 |

\( 1^1 Q_{1,M=3}^* = 1.44 \text{ PVI} \)
\( 2^2 Q_{2,M=3}^* = 1.84 \text{ PVI} \)

Comparison of ST and MADS results in Table 8 demonstrates that ST again provides results quite close to those achieved using a formal optimization method. The largest relative difference between ST and (median)
MADS results is for well scenario 1, $M = 5$, where ST underperforms MADS by 4.7%. For well scenario 1, $M = 1$, ST actually achieves a larger NPV than any of the MADS runs. We also see that there is very little variation between the MADS results for all cases.

Table 9 presents elapsed time, under full parallelization, for ST and MADS optimizations. Individual simulation runs require $\sim 105–115$ seconds for these cases.
Figure 11: Cumulative field oil production (a), water production (b), and water injection (c) for base cases ($Q^*$, 1 PVI, 1.5 PVI and 2 PVI) and ST (single control period, well scenario 2, $M = 3$, 3D example).

Table 8: Optimal NPVs for ST and MADS for 3D reservoir model with a single control period

| Well scenario | Mobility ratio | NPV (BN $) | ST | med  | min  | max  |
|---------------|----------------|------------|-----|------|------|------|
|               |                | MADS       |     |      |      |      |
| 1             | 1              | 17.64      | 17.44 | 17.35 | 17.53 |
|               | 3              | 5.45       | 5.52  | 5.48  | 5.60  |
|               | 5              | 3.85       | 4.04  | 4.04  | 4.04  |
| 2             | 1              | 22.70      | 22.88 | 22.88 | 22.88 |
|               | 3              | 7.32       | 7.56  | 7.56  | 7.57  |
|               | 5              | 4.65       | 4.80  | 4.80  | 4.80  |

We thus see that ST elapsed times continue to be less than the time required for two full-physics runs. MADS requires six iterations for these cases, and speedup using ST of about $5 \times$ relative to MADS is observed. It is again apparent that ST is able to achieve reasonable results along with substantial speedup.

Table 9: Elapsed time (fully parallelized) for ST and MADS for a single control period for 3D reservoir model

| Well scenario | Mobility ratio | Elapsed time (s) |
|---------------|----------------|------------------|
|               |                | ST | MADS$^1$       |
| 1             | 1              | 151| 741             |
|               | 3              | 146| 736             |
|               | 5              | 147| 738             |
| 2             | 1              | 132| 741             |
|               | 3              | 136| 735             |
|               | 5              | 137| 734             |

$^1$Median values

4.2.2. Multiple control periods

In our final example, we consider well scenario 1 (three producers and three injectors) with $M = 3$, and optimize a multiple control period strategy. Specifically, the well rates are now determined four times during the simulation time frame, which increases $n_{opt}$ to 20 (16 well-rate ratios plus 4 field rates). Two approaches can be used to assess ST performance relative to MADS for this case. First, we can simply apply the ST well rates over the entire time frame. Second, the single-period ST solution can be used as an initial guess for MADS (we designate this case ST & MADS).

The NPV results and corresponding elapsed times obtained with ST (standard single control period solution), ST & MADS, and MADS (we give the median, minimum and maximum NPVs for the five standalone MADS runs) are presented in Table 10. We see that both methods provide generally similar NPV results, and there does not appear to be substantial benefit from using multiple control periods in this case. We do, however, observe that the NPV resulting from the ST & MADS approach provides an objective function value that is 3% higher than the median solution from standalone MADS. The corresponding optimized injection and production rates for the four control intervals obtained by ST & MADS are presented in Figs. 12 and 13.

We do see some shifts from control period to control period, though these shifts are relatively slight. From the timings in Table 10, we see that ST & MADS is slightly faster than MADS (elapsed time of 971 seconds compared to median MADS time of 1355 seconds). The improvement in NPV with iteration is shown in Fig. 14 which displays the evolution of NPV for ST, ST & MADS, and standalone MADS (median case). ST & MADS leads to speedup relative to standalone MADS because it requires 9 iterations, while standalone MADS performs 13 iterations (both meth-
Table 10: Optimal NPVs and elapsed time (fully parallelized) for ST and MADS for 3D reservoir model with multiple control periods (well scenario 1, $M=3$).

| Result | ST and MADS | ST MADS | med | min | max |
|--------|-------------|---------|-----|-----|-----|
| NPV (BN $) | 5.45 | 5.63 | 5.46 | 5.36 | 5.55 |
| Elapsed time (s) | 136 | 971 | 1355 | 1228 | 1377 |

A key advantage of ST & MADS, however, is that standalone MADS must be run multiple times to avoid poor local optima, while ST & MADS is run only once. This is possible because the initial guess provided by standard ST is expected to lead to a promising (local) optimum, as is indeed observed for this example. This behavior, if achieved consistently over a range of test cases, would represent a real advantage for ST & MADS, as multiple optimization runs would not be required.

5. Concluding remarks

In this work, we introduced a new two-step surrogate treatment (ST) for well control optimization of oil fields under waterflood. The method provides injection and production rates for all wells for a single control period. Our ST proceeds by first determining optimal well-rate ratios, which are expected to lead to more uniform reservoir sweep. In the second step, overall field rate is optimized such that the objective function of the actual problem (NPV in the cases considered here) is maximized or minimized. In the first step we use simplified physics, specifically the assumption of unit-mobility-ratio displacement and exact voidage replace-
ment. This step is accelerated through use of velocity response matrices, which can be constructed efficiently in a preprocessing step. Full-physics simulations are used in the second ST step, but this minimization can be performed quickly since it involves only a single optimization variable. Near-global optima are obtained in both steps, though this does not provide the global optimum to the actual problem due to the approximations introduced in the ST. The overall two-step procedure requires an elapsed time corresponding to the runtime for 1–2 full-order simulations, assuming full parallelization in the second ST step.

The ST was applied for both 2D and 3D optimization problems. For the 2D case, 40 well scenarios (corresponding to a wide range of well locations and types), generated by PSO during a field development optimization run, were used to compare ST and MADS. Relatively close correspondence between the two sets of optimization results, for \( M = 3 \), was observed. More detailed comparisons were presented for three of the well scenarios for systems with \( M = 1, 3 \) and 5. Multiple MADS and adjoint-gradient runs were performed to assure avoidance of poor local optima with these methods. Again, close correspondence between results from our two-step ST and the formal optimization procedures was observed. Similar observations applied for the 3D case, where two well scenarios were considered, again for \( M = 1, 3 \) and 5. In terms of timings, under full parallelization ST provided speedups of \( 5 \times \) or more relative to MADS for the cases considered. For a case involving multiple control periods, where the ST result provided the initial guess for MADS, a speedup of less than \( 2 \times \) was observed. However, for such cases standalone MADS should be run multiple times to avoid poor local optima, while the ST & MADS approach need be run only once.

There are a number of areas in which future work extending our two-step ST could be performed. More general well control optimization setups, involving, e.g., nonzero discount rates or the consideration of optimal economic project life \[48\], could be considered. Optimization under geological uncertainty is another important aspect that should be addressed. In this case realization selection, using procedures described in \[42\], could be applied. The two-step ST may be particularly useful within the context of field development optimization, where well number, locations, types and controls must be determined. Such MINLP problems are extremely demanding computationally, so the use of surrogate treatments is highly useful. The applicability of the two-step ST for this problem is suggested by the results shown in Fig. \[7\] and our results in a followup study, which is currently underway, indeed confirm this. Finally, the two-step ST could be extended to handle more complex production scenarios, such as primary production followed by waterflood, or water-alternating-gas processes.

Acknowledgements

We thank Nikhil Padhye for providing the Unified Optimization Framework (which facilitated testing and comparing the different optimization algorithms used in this study) and for helping us with its use. We are also grateful to Marco Antonio Cardoso for useful discussions and suggestions. We thank Petrobras and the industrial affiliates of the Stanford Smart Fields Consortium for partial funding of this work.

References

[1] P. Sarma, L. J. Durlofsky, K. Aziz, W. H. Chen, Efficient real-time reservoir management using adjoint-based optimal control and model updating, Computational Geosciences 10 (2006) 3–36.
[2] K. Zhang, G. Li, A. C. Reynolds, J. Yao, L. Zhang, Optimal well placement using an adjoint gradient, Journal of Petroleum Science and Engineering 73 (3) (2010) 220–226.
[3] J. D. Jansen, Adjoint-based optimization of multi-phase flow through porous media: A review, Computers & Fluids 46 (1) (2011) 40–51.
[4] L. F. Almeida, Y. J. Tupac, M. A. C. Pacheco, M. M. B. R. Velasco, J. G. L. Lazo, Evolutionary optimization of smart-wells control under technical uncertainties, in: Latin American and Caribbean Petroleum Engineering Conference, 2007.
[5] T. J. Harding, N. J. Radcliffe, P. R. King, Hydrocarbon production scheduling with genetic algorithms, SPE Journal 3 (1998) 99–107.
[6] T. Foroud, A. Baradaran, A. Seifi, A comparative evaluation of global search algorithms in black box optimization of oil production: A case study on Brugge field, Journal of Petroleum Science and Engineering 167 (2018) 131–151.
[7] V. Lushpoev, A. Margarit, Optimization of oil field development process based on existing forecast model, Journal of Applied Engineering Science 544 (2018) 391–397.
[8] A. A. Emerick, E. Silva, B. Messer, L. F. Almeida, D. Szwarccman, M. A. C. Pacheco, M. M. B. R. Velasco, Well placement optimization using a genetic algorithm with nonlinear constraints, in: SPE Reservoir Simulation Symposium, 2009.
[9] J. E. Onwunalu, L. J. Durlofsky, Application of a particle swarm optimization algorithm for determining optimum well location and type, Computational Geosciences 14 (1) (2010) 183–198.
[10] C. Audet, J. Dennis, Mesh adaptive direct search algorithms for constrained optimization, SIAM Journal on Optimization 17 (2006) 188–217.
[11] D. Echeverra Ciaurri, O. J. Isebor, L. J. Durlofsky, Application of derivative-free methodologies to generally constrained oil production optimization problems, International Journal of Mathematical Modelling and Numerical Optimisation 2 (2) (2011) 134 – 161.
