

An improved multiscale method for life-cycle production optimization

Diego F. B. Oliveira¹ · Albert C. Reynolds² · Jan Dirk Jansen³

Received: 31 July 2014 / Accepted: 11 September 2015 / Published online: 1 October 2015
© Springer International Publishing Switzerland 2015

Abstract This work proposes a new hierarchical multiscale optimization technique to improve the performance of life-cycle production optimization. This approach represents a combination of two previous multiscale approaches presented in the literature and is theoretically motivated by the concept of refinement indicators. The new algorithm is applied to two example problems, and its performance is compared with other life-cycle production optimization algorithms that have been proposed in the literature including the two hierarchical multiscale optimization methods on which it is based. In a separate paper (Oliveira and Reynolds 2015), the proposed algorithm is successfully applied to a field case.

Keywords Life-cycle production optimization · Multiscale optimization · Adjoint-gradient · Waterflooding optimization

✉ Diego F. B. Oliveira
diego.oliveira@petrobras.com.br

Albert C. Reynolds
reynolds@utulsa.edu

Jan Dirk Jansen
j.d.jansen@tudelft.nl

¹ Petrobras Research Center, Av. Horácio de Macêdo, 950, Cd. Universitária, Rio de Janeiro, RJ 21941-915, Brazil

² Department of Petroleum Engineering, University of Tulsa, 800 South Tucker Drive, Tulsa, OK 74104, USA

³ Department of Geoscience and Engineering, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, The Netherlands

1 Introduction

The life-cycle production optimization is applied in a general framework called closed-loop reservoir management (CLRM) to estimate optimal operational well settings based on the updated knowledge of the reservoir [2, 4, 7, 10, 12, 14, 16–20, 23–25]. The optimization aims to maximize an objective function for the remaining life of the reservoir, which is typically defined as the hydrocarbon recovery or the life-cycle net-present-value (NPV) of production.

In the standard optimization procedure, we divide the reservoir lifetime into a fixed number of control steps and perform a single optimization. These control steps are usually uniformly distributed in time, and all wells usually are assigned the same number of control steps. The choice of the number and length for the control steps is ad hoc and may affect the optimization result. Although widely applied in the oil industry, the standard procedure of defining a fixed set of control steps is not a straightforward task, since we do not know a priori the best control parametrization for each problem. A choice of few controls may not provide enough degrees of freedom to estimate the optimum solution, whereas an excessively high number of controls may lead to badly ill-conditioned problems, which are difficult to solve.

Multiscale techniques are a possible optimization strategy to overcome the inherent difficulty to determine a priori the ideal set of control steps. In multiscale optimization approaches, the control variable parametrization is progressively modified as the optimization proceeds, seeking for an appropriate parametrization which leads to the best estimate of the optimum well controls. For the life-cycle optimization problem, multiscale optimization has already been considered in the literature [1, 11, 14, 22].

We previously proposed a hierarchical multiscale approach [14] for adaptively selecting the lengths of control steps during the optimization, which we refer to as Hi-MO. With Hi-MO, we can merge and split the well controls as the optimization proceeds, but the splitting procedure is simple and straightforward, where we simply refine the controls eligible for splitting into a predetermined number of new controls evenly distributed. We propose an extension of Hi-MO where we keep the same merging step, but we propose a refinement procedure based on the refinement indicator concept following the ideas presented by Lien et al. [11].

This paper is organized as follows: first, we state the waterflooding optimization problem which we apply to evaluate the performance of our new hierarchical technique. Thereafter, we present details of the refinement-indicator based splitting procedure. Finally, we present the results for two numerical examples, a channelized reservoir and the Brugge field. Then, we summarize the results and present the conclusions of this work.

2 Waterflooding optimization problem

Our focus in this work is on waterflooding optimization processes, where we aim to maximize the net present value (NPV) over the remaining life of the reservoir which is given by

$$J(\mathbf{u}) = \sum_{n=1}^{N_t} \left\{ \frac{\Delta t_n}{(1+b)^{\frac{t_n}{\tau_r}}} \left[\sum_{j=1}^P (r_o \cdot \overline{q_{o,j}^n} - r_w \cdot \overline{q_w^n, j}) - \sum_{j=1}^I (r_{wi} \cdot \overline{q_{wi}^n, j}) \right] \right\} \quad (1)$$

where r_o (\$/STB) is the oil revenue; r_w (\$/STB) is the cost of disposal of produced water; r_{wi} (\$/STB) is the water injection cost, and b is the discount rate for a particular reference time τ_r . In the results presented here, we use $\tau_r = 365$ days. The time at the end of the n^{th} time step is denoted by t_n ; Δt_n is the size of the n^{th} time step; and N_t is the total number of time steps. P and I , respectively, are the number of producing and injecting wells; $\overline{q_{o,j}^n}$ (STB/D) and $\overline{q_w^n, j}$ (STB/D), respectively, denote the average oil rate and the average water rate at the j^{th} producing well over the n^{th} time step, whereas $\overline{q_{wi}^n, j}$ (STB/D) is the average water injection rate at the j^{th} injector well for the n^{th} time step.

We let \mathbf{u} denote the column vector of all well controls and let $N_{\mathbf{u}}$ denote the total number of controls. Thus, \mathbf{u} is a $N_{\mathbf{u}}$ -dimensional column vector. To evaluate the NPV function as defined in Eq. 1 for a given \mathbf{u} , it is necessary to perform a reservoir simulation run and compute the NPV from the outcomes of the a simulation run. Here, we simply apply the

gradient-based steepest ascent method [13], thus the reservoir simulator used must be able to compute the gradient of the objective function with respect to the vector of control variables ($\nabla_{\mathbf{u}} J(\mathbf{u})$) and the adjoint method is the only computationally efficient method to obtain an accurate gradient of the objective function. In this work, we use the commercial simulator Eclipse 300 [21] to obtain both the objective function and its gradient. However, Eclipse 300 does not output gradients for nonlinear state-dependent constraints, thus, only linear constraints are considered when applying the steepest ascent algorithm to solve the optimization problem defined below by Eqs. 2a, 2b, and 2c.

The dynamic optimization of a waterflooding process over the reservoir life-cycle subject to bound and linear constraints can be expressed as

$$\text{maximize}_{\mathbf{u} \in \mathbb{R}^{N_{\mathbf{u}}}} J(\mathbf{u}), \quad (2a)$$

$$\text{subject to } u_i^{\text{low}} \leq u_i \leq u_i^{\text{up}}, \quad i = 1, 2, \dots, N_{\mathbf{u}} \quad (2b)$$

$$\mathbf{A}_{\mathcal{I}} \mathbf{u} - \mathbf{c} \leq 0 \quad (2c)$$

where $J(\mathbf{u})$ is the objective function given by Eq. 1. The bound constraints are represented by Eq. 2b, where the terms u_i^{low} and u_i^{up} denote, respectively, the lower and upper limits for the i^{th} control variable. Equation 2c denotes the set of linear inequality constraints, where $\mathbf{c} = [c_1, c_2, \dots, c_{n_i}]^T$ is a n_i -dimensional column vector of with constant entries which represent the inequality constraint values, n_i denotes the number of inequality constraints, and $\mathbf{A}_{\mathcal{I}}$ is the $n_i \times N_{\mathbf{u}}$ Jacobian matrix of the inequality constraint functions. Since we do not have gradients of nonlinear state-dependent constraints, here we consider only linear constraints as in Eq. 2c. In this work, we enforce bound constraints by using a logarithm transformation [9, 23], and the general inequality linear constraints (Eqs. 2c) are handled with the augmented Lagrangian method [6, 13]. We denote the objective function to be maximized as \mathcal{L}_a , which refers to the augmented Lagrangian function [6, 13]. Because the log-transformation is applied to handle bound constraints, if the optimization problem does not contain constraints in the form of Eq. 2c, then we simply have that $\mathcal{L}_a(\mathbf{u}) \equiv J(\mathbf{u})$.

2.1 Well control vector representation

For the life-cycle production optimization problem, the optimization control variables are well operational settings (typically either well fluid rates or flowing bottom-hole pressures) at different times throughout the reservoir lifetime. Letting \mathbf{u}^w denote the vector of well controls at all control steps for well w , the vector \mathbf{u} can be written as

$$\mathbf{u} = \left[(\mathbf{u}^1)^T, (\mathbf{u}^2)^T, \dots, (\mathbf{u}^{N_w})^T \right]^T, \quad (3)$$

where N_w is the total number of wells. Denoting the number of control steps for a particular well w by N_c^w , \mathbf{u}^w , $w = 1, 2, \dots, N_w$, is a N_c^w -dimensional column vector and $\mathbf{N}_u = \sum_{w=1}^{N_w} N_c^w$.

A control step c for a particular well w is defined by an interval of the form

$$I_c^w = [t_{c-1}^w, t_c^w), \tag{4}$$

for $c = 1, 2, \dots, N_c^w$ and with $t_0^w = 0$, where the control variables are specified constant along each interval defined in Eq. 4.

Considering the numerical aspects of reservoir simulators which select time-steps adaptively, it is useful to specify the finest time-scale that will be used to define controls steps. To do so, we define the total number of fine-scale steps, N_s , and compute the fine-scale time interval

$$\Delta \hat{t} = \frac{L}{N_s}, \tag{5}$$

where L (days) is the remaining life of the reservoir. N_s must be a non-zero natural number. For convenience, we should select N_s to be some power of two and require that any coarser control step considered be the union of fine scale steps.

Given $\Delta \hat{t}$, the finest time-scale for control intervals (steps) is defined by the times $\hat{t}_n = \hat{t}_{n-1} + \Delta \hat{t}$, where for simplicity, we assume $\hat{t}_0 = 0$. The fine-scale time intervals are given by $\hat{I}_n = [\hat{t}_{n-1}, \hat{t}_n)$, for $n = 1, 2, \dots, N_s$. Note that no two distinct fine-scale time intervals intersect and the fine-scale steps (intervals) can be defined the same for all wells so that we can also express any control step for any well w in terms of an union of fine-scale time intervals. Therefore, we can define the set of fine-scale intervals that pertains to control step c of well w as

$$\mathcal{J}_c^w = \left\{ j \mid \hat{I}_j \subset I_c^w \right\}. \tag{6}$$

We denote the number of elements of \mathcal{J}_c^w by n_c^w , i.e., the coarse control step interval I_c^w contains n_c^w contiguous fine-scale intervals. Thus, given a specific u_c^w for I_c^w , we can define, for each fine-scale control interval contained in I_c^w , a corresponding fine-scale control variable, \hat{u}_j^w for $j = 1, 2, \dots, N_s$, such that

$$\hat{u}_j^w = u_c^w \quad \forall j \in \mathcal{J}_c^w. \tag{7}$$

By the adjoint method, we can compute all partial derivatives $\frac{\partial \mathcal{L}_a}{\partial \hat{u}_j}$, for $j = 1, 2, \dots, N_s \times N_w$, where N_s is the total number of fine-scale steps and N_w is the total number of wells. However, we need to compute $\frac{\partial \mathcal{L}_a}{\partial u_i}$, for $i = 1, 2, \dots, N_u$. We can write \mathcal{L}_a in terms of $\hat{\mathbf{u}}$, which allows us to apply the chain rule to obtain

$$d\mathcal{L}_a = \sum_{k=1}^{N_u} \sum_{j=0}^{n_k-1} \frac{\partial \mathcal{L}_a}{\partial \hat{u}_{j_k+j}} d\hat{u}_{j_k+j}, \tag{8}$$

where j_k is the index of the first fine-scale variable correspondent to the k^{th} control and n_k is the number of fine-scale control intervals contained in the interval which represents the control step for the k^{th} control variable. From Eq. 8, it follows that

$$\frac{\partial \mathcal{L}_a}{\partial u_i} = \sum_{k=1}^{N_u} \sum_{j=0}^{n_k-1} \frac{\partial \mathcal{L}_a}{\partial \hat{u}_{j_k+j}} \frac{\partial \hat{u}_{j_k+j}}{\partial u_i}. \tag{9}$$

We know that $\frac{\partial \hat{u}_{j_k+j}}{\partial u_i} = 0$, for $j = 0, 1, \dots, n_k - 1$ and $k \neq i$. Moreover, from Eq. 7, we see that for $k = i$, $\frac{\partial \hat{u}_{j_k+j}}{\partial u_i} = 1$ for $j = 0, 1, \dots, n_k - 1$. It follows that

$$\frac{\partial \mathcal{L}_a}{\partial u_i} = \sum_{j=0}^{n_i-1} \frac{\partial \mathcal{L}_a}{\partial \hat{u}_{j_i+j}}. \tag{10}$$

3 Refinement-indicator based hierarchical multiscale optimization

Oliveira and Reynolds [14] presented the hierarchical multiscale method (Hi-MO) for adaptively selecting the number and the lengths of control steps as the overall optimization proceeds and we begin by reviewing their procedure. After maximizing \mathcal{L}_a for a specific level of parametrization, they then select the new set of control steps for the next level. The reparametrization step consists of a merging procedure and a splitting procedure. The coarsening procedure is based on three criteria (see [14, 16] for details), which are checked well-by-well as follows:

Control-Variation criterion For $i = 2, 3, \dots, N_c^w$, verify whether

$$\frac{|u_i^w - u_{i,\text{ref}}^w|}{u_i^{\text{up},w} - u_i^{\text{low},w}} < \xi_u, \tag{11}$$

where ξ_u is the control-variation tolerance for merging well controls; u_i^w is the i^{th} control at well w , and $u_{i,\text{ref}}^w$ is a reference control value corresponding to the average among all consecutive controls previous to the i^{th} control which were merged, i.e.,

$$u_{i,\text{ref}}^w = \frac{1}{n_{\Omega}^{i-1}} \sum_{j \in \Omega_{i-1}^w} u_j^w, \tag{12}$$

where Ω_{i-1}^w is the set of control steps which were merged with the control step $i-1$ and n_{Ω}^{i-1} is the number of elements of Ω_{i-1}^w .

Gradient-Variation criterion If after optimization at the previous level of parametrization, two temporarily consecutive controls are approximately equal (verified with Eq. 11)

and the derivatives of \mathcal{L}_a with respect to these two controls are also approximately equal, these controls should be merged. Specifically, to verify if a control step is eligible for merging according to the gradient-variation criterion, we check whether

$$\frac{|g_i^w - g_{i,\text{ref}}^w|}{\max\{g_{\text{max}}^w, 1 \times 10^{-2}\}} < \xi_g, \quad (13)$$

where ξ_g is the gradient-variation tolerance for merging well controls, $g_{\text{max}}^w = \max\{|g_j^w|\}_{j=1}^{N_c^w}$ is the maximum magnitude of the gradient components for well w , g_i^w is the component of the gradient vector ($\nabla\mathcal{L}_a$) related to the i^{th} control at well w ; $g_{i,\text{ref}}^w$ is a reference gradient defined by

$$g_{i,\text{ref}}^w = \frac{1}{n_{\Omega}^{i-1}} \sum_{j \in \Omega_{i-1}^w} g_j^w. \quad (14)$$

Gradient-Magnitude criterion The third criterion requires that we should not merge controls which correspond to gradient components of large magnitude, so, specifically, we require that

$$\frac{|g_i^w|}{\max\{g_{\text{max}}^w, 1 \times 10^{-2}\}} < \rho_g, \quad (15)$$

where ρ_g is the gradient-magnitude tolerance for merging well controls in order to merge the control interval for u_i^w with the set of control steps which were merged with the control step $i - 1$ and $\rho_g \equiv (1 - \xi_g)$.

A particular control step is merged with the preceding control step only if all three criteria hold. Controls that were not merged should be split in a predefined number of evenly distributed new controls, which is the uniform splitting procedure for Hi-MO (see details in [14, 16]). This splitting procedure is quite simple and straightforward but is heuristic. The purpose of this work is to present a procedure for splitting control steps which has a more theoretical foundation. In the multiscale algorithm presented here, the merging criteria is identical to the one outlined above [14, 16] but the splitting criteria have a more fundamental basis.

3.1 Theoretical motivation of the refinement procedure

Based on the ideas discussed originally by Chavent and Bissell [5] and Ben Ameer et al. [3], and applied to the optimal well control problem by Lien et al. [11], we consider using so-called *Refinement Indicators* (RI) to determine the best way to split a particular control step at each refinement level of our hierarchical multiscale algorithm. The refinement-indicator concept is based on the equivalence of two optimization solutions for two different parameterizations of control variables, or, for the problem of interest here, two different levels of refinement.

Without loss of generality, consider for simplicity a problem with one single control variable \tilde{u} , which may be split into two new control variables, $\tilde{u}^{(1)}$ and $\tilde{u}^{(2)}$. The one-dimensional optimization problem based on \tilde{u} is given by

$$\text{maximize}_{\tilde{u} \in \mathbb{R}} \mathcal{L}_a(\tilde{u}), \quad (16)$$

which has a solution denoted by \tilde{u}^* . Similarly, the two-dimensional optimization problem based $\tilde{u}^{(1)}$ and $\tilde{u}^{(2)}$ can be stated as

$$\text{maximize}_{\tilde{u}^{(1)}, \tilde{u}^{(2)} \in \mathbb{R}} \mathcal{L}_a(\tilde{u}^{(1)}, \tilde{u}^{(2)}), \quad (17)$$

whose solution is denoted by $\tilde{u}^{*,(1)}$ and $\tilde{u}^{*,(2)}$. We may think of \tilde{u}^* as the optimal control for a time interval $[0, L]$ when only one control step is used for the entire time interval and $[\tilde{u}^{*,(1)}, \tilde{u}^{*,(2)}]^T$ as the optimal vector of controls when $[0, L]$ is split into two control steps $[0, t_1]$ and $[t_1, L]$ where the optimal vector depends on the choice of t_1 . Let $\delta\tilde{u}^* = \tilde{u}^{*,(1)} - \tilde{u}^{*,(2)}$ be the difference (perturbation) between the values of the two optimum controls for problem defined by Eq. 17. Note that if $\delta\tilde{u}^* = 0$, the solutions of the problems specified by Eqs. 16 and 17 are the same. On the other hand, if we know a priori the optimum perturbation $\delta\tilde{u}^*$, that will maximize $\mathcal{L}_a(\tilde{u}^{(1)}, \tilde{u}^{(2)})$ over all t_1 such that $0 < t_1 < L$, then the solution of problem in Eq. 17 is the same as the solution of the following constrained problem

$$\text{maximize}_{\tilde{u}^{(1)}, \tilde{u}^{(2)} \in \mathbb{R}} \mathcal{L}_a(\tilde{u}^{(1)}, \tilde{u}^{(2)}), \quad (18a)$$

$$\text{subject to } \tilde{u}^{(1)} - \tilde{u}^{(2)} = \delta\tilde{u} \quad (18b)$$

where $\delta\tilde{u} = \delta\tilde{u}^*$.

However, we do not know a priori the optimal perturbation $\delta\tilde{u}^*$, so we instead consider the generic equality constraint in Eq. 18b in terms of an unknown perturbation, $\delta\tilde{u}$, and investigate the sensitivity of the optimum value of the objective function \mathcal{L}_a to this perturbation. Chavent and Bissell [5], and other authors who followed their ideas [3, 11], derived the sensitivity of the optimum value of the objective function \mathcal{L}_a to the perturbation $\delta\tilde{u}$ and Lien et al. [11] showed that the Lagrange multipliers associated with the constrained problem in Eq. 18 represent the sensitivity of \mathcal{L}_a to $\delta\tilde{u}$. For the simple problem considered here, the Lagrange multiplier at the optimal $(\tilde{u}^{(1)}, \tilde{u}^{(2)})$ associated with the equality constraint in Eq. 18b is given by

$$\lambda = \frac{\partial\mathcal{L}_a}{\partial\tilde{u}^{(1)}} = -\frac{\partial\mathcal{L}_a}{\partial\tilde{u}^{(2)}}, \quad (19)$$

where λ must be nonnegative to satisfy the optimality conditions. It follows that

$$2\lambda = \frac{\partial\mathcal{L}_a}{\partial\tilde{u}^{(1)}} - \frac{\partial\mathcal{L}_a}{\partial\tilde{u}^{(2)}} = \left| \frac{\partial\mathcal{L}_a}{\partial\tilde{u}^{(1)}} - \frac{\partial\mathcal{L}_a}{\partial\tilde{u}^{(2)}} \right|, \quad (20)$$

where the last equality of Eq. 20 follows from the fact that $\lambda \geq 0$ by standard optimality conditions [13]. Higher values of λ correspond to higher values of the sensitivity of $\mathcal{L}_a(\tilde{u}^{(1)}, \tilde{u}^{(2)})$ to $\tilde{u}^{(1)} - \tilde{u}^{(2)} = \delta\tilde{u}$. This suggests that given a set of choices for the splitting of a given \tilde{u} into two controls, $(\tilde{u}^{(1)}$ and $\tilde{u}^{(2)})$, i.e., given a set of choices for t_1 , a good choice is to split the control step into the two steps that correspond to the highest sensitivity. More specifically, given an optimal control \tilde{u}_i for a control step $[\tilde{t}_{i-1}, \tilde{t}_i)$, the basic idea of the splitting procedure is to consider the consequences of splitting this control interval two control steps, $[\tilde{t}_{i-1}, \tilde{t}_s)$ and $[\tilde{t}_s, \tilde{t}_i)$ for a sequence of increasing values of \tilde{t}_s . The values considered for \tilde{t}_s would be chosen from the set of endpoints of fine scale time intervals that are contained in $(\tilde{t}_{i-1}, \tilde{t}_i)$. For each \tilde{t}_s , we let $\tilde{u}^{(1),s}$ correspond to the control variable on $[\tilde{t}_{i-1}, \tilde{t}_s)$ and let $\tilde{u}^{(2),s}$ correspond to the control variable on the control interval $[\tilde{t}_s, \tilde{t}_i)$. For each s , we evaluate the refinement indicator which is defined by

$$\pi_s \equiv \pi(\tilde{t}_s) = \left| \frac{\partial \mathcal{L}_a}{\partial \tilde{u}^{(1),s}} - \frac{\partial \mathcal{L}_a}{\partial \tilde{u}^{(2),s}} \right|, \tag{21}$$

where the derivatives required in Eq. 21 are evaluated at $(\tilde{u}^{(1),s}, \tilde{u}^{(2),s}) = (\tilde{u}_i, \tilde{u}_i)$. Then we choose the value of \tilde{t}_s that gives the highest value of π_s .

Consider a choice of \tilde{t}_1 that divides $[\tilde{t}_{i-1}, \tilde{t}_i)$ into two nonempty subintervals and let $\tilde{u}^{(1)}$ and $\tilde{u}^{(2)}$ denote the corresponding partition of \tilde{u}_i such that \tilde{u}_i is the optimal based on the non-partitioned control interval. It follows from Eqs. 10, 19 and the optimality of $\tilde{u}^{(i)}$ that

$$\frac{\partial \mathcal{L}_a}{\partial \tilde{u}}(\tilde{u}_i) = \frac{\partial \mathcal{L}_a}{\partial \tilde{u}^{(1)}} + \frac{\partial \mathcal{L}_a}{\partial \tilde{u}^{(2)}} = \frac{\partial \mathcal{L}_a}{\partial \tilde{u}^{(1)}} - \frac{\partial \mathcal{L}_a}{\partial \tilde{u}^{(1)}} = 0. \tag{22}$$

so if we try to perform an iteration of the steepest ascent method based on \tilde{u}_i , the optimization will terminate because $\nabla_{\tilde{u}} \mathcal{L}_a = 0$. However, Eq. 22 does not imply that $\nabla_{\tilde{u}^{(1)}, \tilde{u}^{(2)}} \mathcal{L}_a$ evaluated at $(\tilde{u}^{(1)}, \tilde{u}^{(2)}) = (\tilde{u}_i, \tilde{u}_i)$ is equal to zero. Thus, if one splits \tilde{u} into $\tilde{u}^{(1)}$ and $\tilde{u}^{(2)}$, it may be possible to obtain at greater value of \mathcal{L}_a based on the refined control steps. As discussed in more detail immediately below, our choice is to choose \tilde{t}_1 to maximize the refinement indicator (Eq. 21) as this choice maximizes the sensitivity of \mathcal{L}_a to the new control variables.

Similar to the above discussion, we now consider the i^{th} control variable of well w , u_i^w , which is defined on the control step (interval) $I_i^w = [t_{i-1}^w, t_i^w)$; see Eq. 4. Consider that we cut I_i^w at a time $t_{i,s}^w \in (t_{i-1}^w, t_i^w)$ to form the two new potential control intervals given by

$$I_i^{w,(1)} = [t_{i-1}^w, t_{i,s}^w), \tag{23}$$

and

$$I_i^{w,(2)} = [t_{i,s}^w, t_i^w), \tag{24}$$

which define, respectively, two new control variables $u_i^{w,(1)}$ and $u_i^{w,(2)}$. Then, based on Eq. 21, the refinement indicator associated with the cut at $t_{i,s}^w$ is given by

$$\pi_{i,s}^w = \left| \frac{\partial \mathcal{L}_a}{\partial u_i^{w,(1)}} - \frac{\partial \mathcal{L}_a}{\partial u_i^{w,(2)}} \right|, \tag{25}$$

where the derivatives presented in Eq. 25 are evaluated at $(u_i^{w,(1)}, u_i^{w,(2)}) = (u_i^w, u_i^w)$. Note that the derivatives in Eq. 25 are easily computed from the derivatives of the fine-scale control variables as defined in Eq. 7, so computing the refinement indicators, $\pi_{i,s}^w$, is virtually cost free; thus, we can compare a large number of potential cuts in u_i^w without increasing substantially the computational cost.

3.2 RI-based splitting procedure

As discussed previously, the multiscale method that we propose below incorporates refinement indicator (RI) values to split control intervals but in all other details follows the steps of the Hi-MO algorithm of [14, 16]. The new algorithm is referred to as Refinement Indicator-based Hierarchical multiscale method (RHi-MO). Similar to Hi-MO, the RI-based splitting procedure is applied only for the control steps that were not obtained by merging two or more control intervals in the merging procedure. We check the control step refinement on a well-by-well basis, so for each control interval eligible for splitting, I_i^w , which corresponds to the i^{th} control step at well w , we find the time-cut, $t_{i,s}^w$, which provides the highest RI ($\pi_{i,s}^w$) among all potential cuts considered for I_i^w , denoting by $n_{\text{pot},i}^w$ the number of potential cuts of the i^{th} control interval of well w . Lien et al. [11] proposed setting $n_{\text{pot}} = 6$ for all controls. However, since the computational cost required to evaluate the refinement indicators is negligible, we propose a procedure to specify the number of potential cuts for each control interval by solving

$$2^{n_{\text{pot},i}^w} = \frac{\Delta t_i^w}{\Delta \hat{t}} \tag{26}$$

for $n_{\text{pot},i}^w$, where $\Delta t_i^w = t_i^w - t_{i-1}^w$ and $\Delta \hat{t}$ is defined in Eq. 5. Then, we define the potential time-cuts as

$$t_{i,s}^w = t_{i-1}^w + s \left(\frac{\Delta t_i^w}{n_{\text{pot},i}^w} \right), \tag{27}$$

for $s = 1, \dots, n_{\text{pot},i}^w$, and compute the RI ($\pi_{i,s}^w$) associated with each time-cut $t_{i,s}^w$ using Eq. 25. To prevent suddenly splitting a control step into an excessively small new control interval, we add a safeguard that any new control interval obtained by splitting cannot be smaller than a quarter of the original control step. Thus, we set $\pi_{i,s}^w = 0$ for any cut $t_{i,s}^w$

such that $t_{i,s}^w - t_{i-1}^w < \Delta t_i^w/4$ or $t_i^w - t_{i,s}^w < \Delta t_i^w/4$. This safeguard allows the hierarchical refinement to be developed more homogeneously, i.e., we do not implement any abrupt reduction in length of any control interval. We compute the highest RI as $\pi_i^w = \max\{\pi_{i,s}^w\}_{s=1}^{n_{\text{pot},i}^w}$. For those controls not eligible for refinement, i.e., those which were obtained by merging two or more control steps, we simply set $\pi_i^w = 0$.

Once we have computed the best cut for each control step of well w , we propose three options on how to implement the best local split. The splitting procedure is performed according to one option which defines the number of control steps per well to be refined. Denoting the highest refinement indicator value among all control steps of well w as $\pi_{\text{max}}^w = \max\{\pi_i^w\}_{i=1}^{N_c^w}$, the three options available are as follows:

1. Split ONE: Only one single control step per well is split. The control step selected for splitting corresponds to the one with the highest refinement indicator at that particular well (π_{max}^w);
2. Split FEW: The i^{th} control steps at a particular well w will be split if and only if $\pi_i^w \geq \rho_g \pi_{\text{max}}^w$ where $\rho_g \in (0, 1)$ and $i = 1, 2, \dots, N_c^w$;
3. Split ALL: All control steps that were not obtained by merging two or more control intervals in the merging step are split.

We set the scalar ρ_g required when the splitting option is chosen as FEW equal to the gradient-magnitude tolerance for merging well controls (see Eq. 15). Note that if $\rho_g = 0$, then all control steps are selected for splitting, which is equivalent to selecting the option ALL. On the other hand, if $\rho_g = 1$, then only the control steps which have a refinement indicator equal to maximum indicator are selected for splitting, which is virtually always equivalent to selecting the option ONE. After defining the control steps at any refinement level, the solution of the resulting optimization problem is solved using the augmented Lagrangian method where at each inner iteration, the augmented Lagrange function is maximized by the steepest ascent algorithm; see [8, 13] for theoretical convergence results.

In terms of the initial set of control variables, similar to the method of Lien et al., we always initialize RHi-MO with one single control step per well. However, we perform one RI-based splitting procedure before the first optimization step is conducted followed by checking whether new control steps should be remerged. This initialization may provide a better initial set of control steps without introducing any additional cost since all required gradients can be obtained from the fine-scale gradients already computed for the initial guess.

For the sake of comparison, we also consider the other multiscale methods defined in [14], namely, the Hi-MO

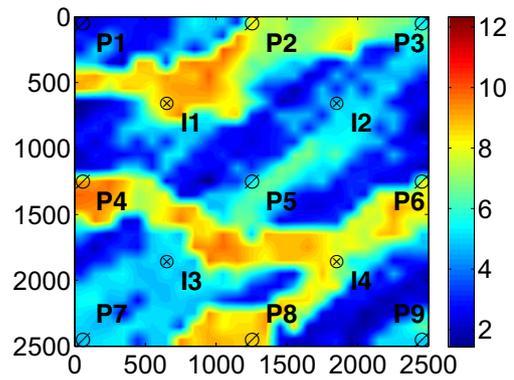


Fig. 1 $\ln k_h$, three channel case

of Oliveira and Reynolds [14], the successive splitting multiscale optimization (SS-MO) [11, 22], and the method proposed by Lein et al. [11] based on refinement indicators (RI-MO); see [14] for details.

We specify $\xi_u = 0.01$, $\xi_g = 0.10$, and $\rho_g = 0.90$ as the reparameterization tolerances for the examples presented in this paper. The maximum number of allowable reservoir simulations is set equal to 150 for the problems with only bound constraints, whereas we set 1000 as the maximum allowed number of simulation runs for the constrained problems solved with the augmented Lagrangian method. In terms of the tolerances for the change in the objective function and for the change in the control variables, we use $\varepsilon_f^0 = 0.1$ and $\varepsilon_u^0 = 0.1$, where ε_f^0 and ε_u^0 , respectively, are the initial convergence tolerances for the change in the objective function and for the change in the control variables. We also use $\varepsilon_f^* = 1 \times 10^{-4}$ and $\varepsilon_u^* = 1 \times 10^{-3}$, where ε_f^* and ε_u^* , respectively, are the final convergence tolerances for the change in the objective function and for the change in the control variables.

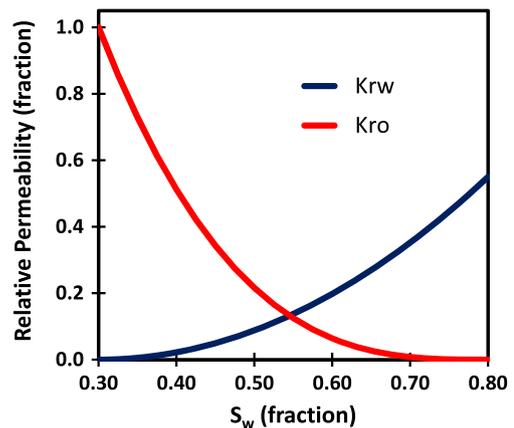


Fig. 2 Relative permeability curves, three channel case

Table 1 Reservoir properties, three channel case

Reservoir grid	$25 \times 25 \times 1$
Depth	4800 ft
Initial pressure	4000 psi
Porosity (ϕ)	0.20
Compressibility (c_t)	$6.9 \times 10^{-5} \text{ psi}^{-1}$
Initial saturation (S_{wi})	0.20
Viscosity (μ)	2.2 cP

4 Examples

In this section, we apply the refining and coarsening procedures to two examples and the performance of RHi-MO is compared to the performance of the other methods. In the examples considered, the objective is to optimize the overall NPV for a reservoir under waterflooding. The first example pertains to a two-dimensional reservoir with a geological heterogeneity represented by a channelized depositional environment. The second example considers a large field-scale reservoir with a complex geological description. The reservoir contains several wells under control, both producers and injectors. The specified reservoir life is long enough to observe significant water production in many production wells. In all examples, we apply two possible constraint scenarios. First, we consider only bound constraints (Eq. 2b). In the second scenario, in addition to the bound constraints, we also consider linear inequality constraints (Eq. 2c).

4.1 Example 1: three-channel case

The three-channel model is based on a uniform grid with $25 \times 25 \times 1$ grid blocks, $\Delta x = \Delta y = 100$ ft, and the thickness of the reservoir is 20 ft. The permeability distribution, which is shown in Fig. 1 in terms of $\ln k_h$, represents a channelized reservoir, which contains nine producers and

four injectors arranged in four five-spot patterns; the reservoir is produced under a waterflooding strategy with relative permeability curves presented in Fig. 2, and properties summarized in Table 1.

The total liquid rates of each producing well and the water injection rate of each injector are controlled for the waterflooding optimization problem during the anticipated total project lifetime, which is 1800 days. The upper bounds for the total liquid rate at the producers is 200 STB/D and for the water injection rate at the injecting wells is 500 STB/D. For both producers and injectors, the lower bound is 0 STB/D. We define 80 STB/D as the initial guesses for the total liquid producing rates and 200 STB/D as the initial guess at the injectors. In this example, we optimize the NPV, considering the oil revenue is \$50.0 /STB; the water production cost is \$5.0 /STB; the cost of injecting water is \$2.0 /STB; and an annual discount rate of 10.0 %.

Only bound constraints We optimize the three-channel case using the steepest ascent method based on the adjoint-gradient. The multiscale techniques SS-MO, Hi-MO, RI-MO, and RHi-MO are compared against the base parametrization case. For the base case, there are 10 fixed control steps of 180 days duration for each well; i.e., there are 130 control steps (and 130 control variables). We also consider a fine parametrization such that each well contains 128 fixed control steps which results in 1664 control variables. We start SS-MO with two control steps per well, each of duration equal to 900 days, and set $n_{split} = 2$. For RI-MO, we apply one initial control step per well as in Lien et al. [11]. For Hi-MO, we consider the cases where either $n_{split} = 2$ or $n_{split} = 4$ which are applied in two possibilities for the initial set of control steps, 2 and 10 initial control steps per well. In terms of RHi-MO application, we vary the number of control steps actually split by using the options ONE, FEW, and ALL.

Table 2 Summary; only bound constraints, three channel case

	NPV, $\times 10^7$ \$	# Sim.	Final N_u
Base parametrization: 10 fixed control steps	6.6924	33	130
Fine parametrization: 128 fixed control steps	6.6158	44	1664
SS-MO	6.7087	59	832
RI-MO (1 init. control step)	6.6560	39	34
Hi-MO (2 init. control steps; $n_{split} = 2$)	6.7122	36	105
Hi-MO (2 init. control steps; $n_{split} = 4$)	6.7368	49	197
Hi-MO (10 init. control steps; $n_{split} = 2$)	6.6766	55	211
Hi-MO (10 init. control steps; $n_{split} = 4$)	6.7389	43	217
RHi-MO (split ALL)	6.9037	81	252
RHi-MO (split ONE)	6.9262	36	56
RHi-MO (split FEW)	6.9264	33	60

Fig. 3 Optimal control rates for producing wells; only bound constraints for base parametrization (10 fixed control steps), fine parametrization (128 fixed control steps), SS-MO, RI-MO, Hi-MO with 10 initial control steps ($n_{\text{split}} = 4$), and RHi-MO with option split FEW, three-channel case

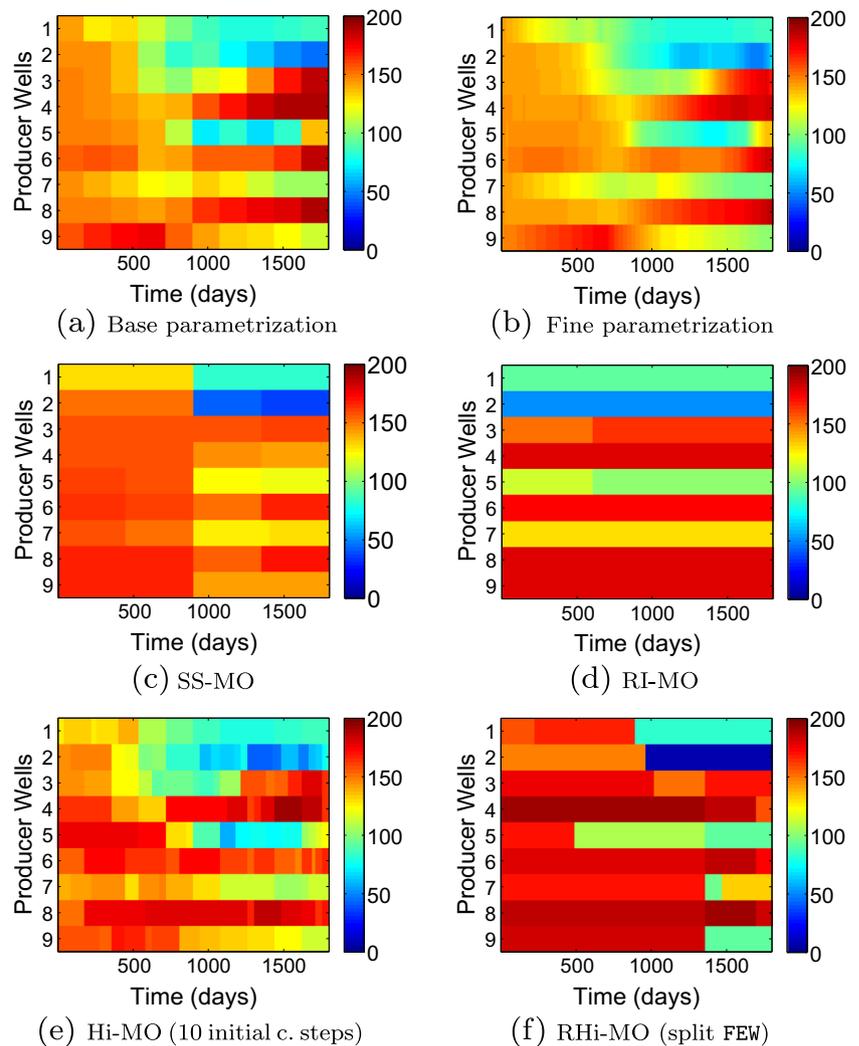
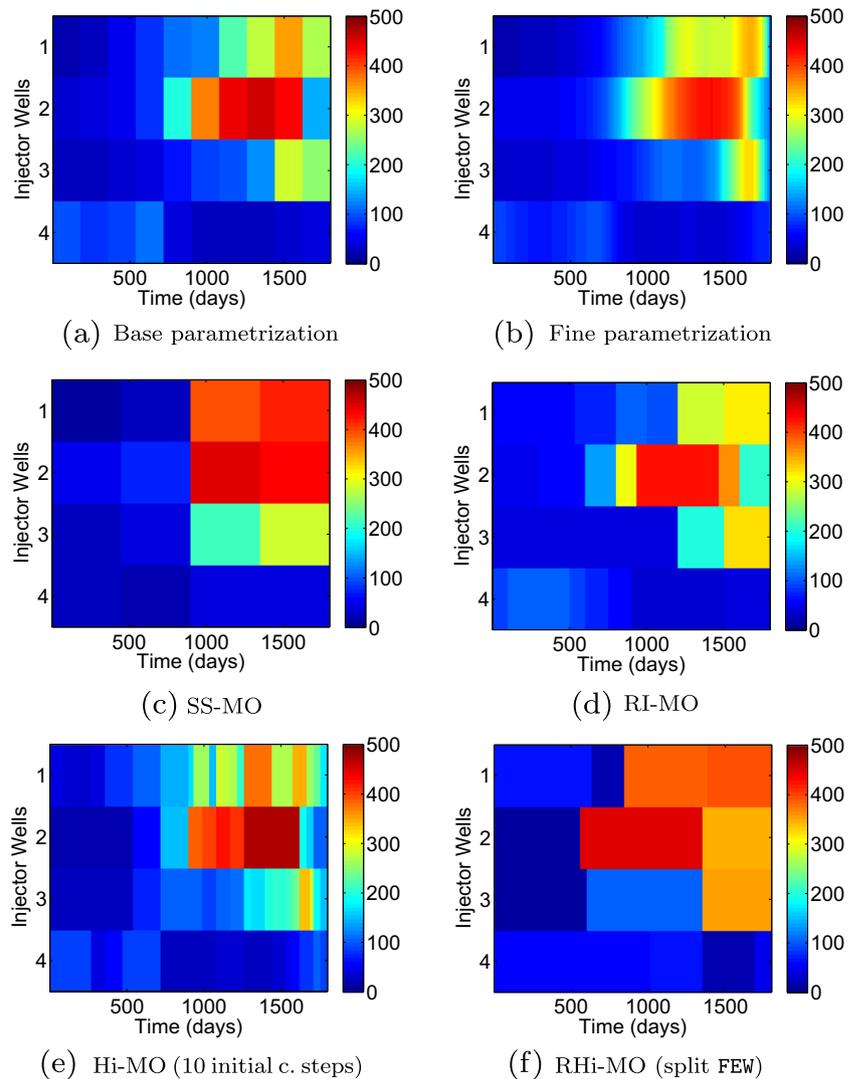


Table 2 presents the summary of all results obtained for the three-channel case. The results are compared in terms of the final NPV obtained by each approach, the number of reservoir simulation runs required to reach that final NPV, and the number of control variables used by each technique. We note from the results presented in Table 2 that, except for RI-MO (method of Lien et al.) and for Hi-MO (10 initial control steps; $n_{\text{split}} = 2$), all multiscale techniques outperform the base parametrization. The results of SS-MO and the majority of the Hi-MO applications are superior to the base parametrization, however, by a small margin of 0.25 % for SS-MO and at most 0.7 % for Hi-MO. For this example, all Hi-MO schemes lead to results better than SS-MO in terms of the final NPV and the number of simulation runs required, except for the case with 10 initial control steps and $n_{\text{split}} = 2$, but the difference on the final NPV is small. On the other hand, the results obtained with the three different versions of RHi-MO give fairly close optimal values of NPV

and the final NPVs for these methods are more than 3 % higher than the NPV obtained with the base parametrization or any other method. Additionally, RHi-MO is particularly efficient in terms of its computational cost for this example, since the best result obtained with RHi-MO requires the same number of simulation runs as the base parametrization but leads to a final NPV that is 3.5 % higher.

We also note from the results presented in Table 2 that we do not obtain better results by introducing an extremely refined set of control steps as we have for the fine parametrization case. In fact, the fine parametrization leads to a final NPV 1.1 % lower than the base parametrization and requires more simulation runs. This fact is not unexpected since a certain level of ill-conditioning may exist for large scale optimization problem, which can affect the overall optimization performance in terms of both the final objective function and the number of simulation runs required.

Fig. 4 Optimal control rates for injecting wells; only bound constraints, for base parametrization (10 fixed control steps), fine parametrization (128 fixed control steps), SS-MO, RI-MO, Hi-MO with 10 initial control steps ($n_{split} = 4$), and RHi-MO with option split FEW, three-channel case



In terms of the final number of control variables, RHi-MO with splitting option ONE and FEW utilizes far fewer control variables than all other methods except RI-MO which yields a much lower NPV than the RHi-MO method. Note that SS-MO leads to an extremely large number of controls but results in essentially the same NPV as is obtained with the base parameterization.

Although the final number of control variables is not a performance criterion per se, it is important to highlight that optimization approaches which require fewer well controls are preferable from the operational point of view as long as they still ensure a reasonably high final NPV. Strategies which result in a small number of well controls correspond to strategies which require few changes in the well operational settings in the field production operation.

Optimum well controls obtained by the methods considered are presented in Figs. 3 and 4, which depict the controls obtained with the base parametrization (10 fixed

controls steps), the fine parametrization (128 fixed control steps), SS-MO, RI-MO, Hi-MO with 10 initial control steps ($n_{split} = 4$), and RHi-MO with option split FEW. In Figs. 3 and 4, the well controls are presented in a color map, where the horizontal axis represents the time (days) and the vertical axis indicates the wells. The producer well controls are displayed in Fig. 3, whereas Fig. 4 shows the controls for the injecting wells.

Although the well controls depicted in Figs. 3 and 4 vary as the algorithm used to generate them varies, there are some qualitative similarities. For instance, we note that in the pair PRO-02/INJ-01, which are connected through a highly permeable channel in the North part of the reservoir, we tend to produce PRO-02 at high rates at early time, decreasing the total liquid production at PRO-02 later; on the other hand, we inject at low rates at the beginning, when the reservoir still provides a good pressure support, then we increase the water injection rate at INJ-01. All methods somehow capture this trend, except for RI-MO whose solution for the

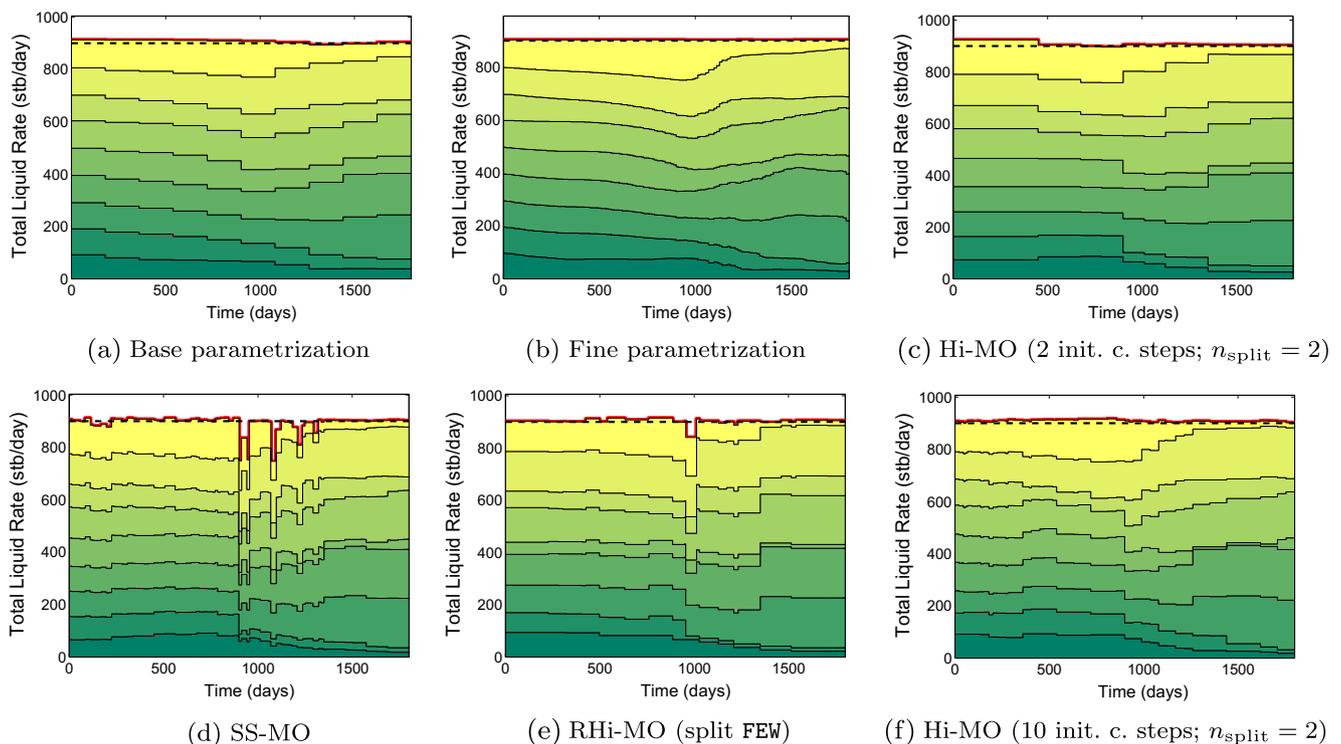
Table 3 Summary; bound plus inequality constraints, three channel case

	NPV, $\times 10^7$ \$	# Sim.	Final N_u
Base parametrization: 10 fixed control steps	6.0239	132	130
Fine parametrization: 128 fixed control steps	6.0570	92	1664
SS-MO	6.0687	375	1664
Hi-MO (2 init. control steps; $n_{\text{split}} = 2$)	6.1049	814	117
Hi-MO (10 init. control steps; $n_{\text{split}} = 2$)	6.1390	644	280
RHi-MO (split ALL)	6.0750	649	189
RHi-MO (split ONE)	6.0792	485	88
RHi-MO (split FEW)	6.1110	989	116

controls of well PRO-02 contains only one control step. It is important to note that the controls at each well exhibit significant temporal smoothness although we do not explicitly impose any correlation between the well controls.

Bound plus linear inequality constraints We reconsider here the three-channel model with the same bound constraints plus two linear inequality constraints. The first linear inequality constraint is that the total liquid production rate is less than or equal to 900 STB/D at all times. The second linear inequality constraint is that the total injection rate is less than or equal to 1000 STB/D at all times. The constrained optimization problem is solved using the augmented Lagrangian method applied with the adjoint-based steepest ascent method.

We summarize in Table 3 the results obtained for the constrained problem associated with the three-channel case. Table 3 presents the results for the base and fine parametrization, successive splitting, and the hierarchical optimization methods (Hi-MO and RHi-MO). The original method of Lien et al. [11] was not designed for constrained optimization problems, particularly within the augmented Lagrangian framework; so we do not apply RI-MO here or in any other constrained optimization problem in this paper. Similar to the problem with only bound constraints, the base parametrization case has 10 fixed control steps per well (130 controls at total), whereas in the fine parametrization we consider 128 control steps per well (1664 controls at total). As usual, SS-MO starts with two control steps per well and uses $n_{\text{split}} = 2$ at every level of parametrization. We consider two initialization strategies for Hi-MO, where in one

**Fig. 5** Optimum total liquid production rate; bound plus inequality constraints, three-channel case

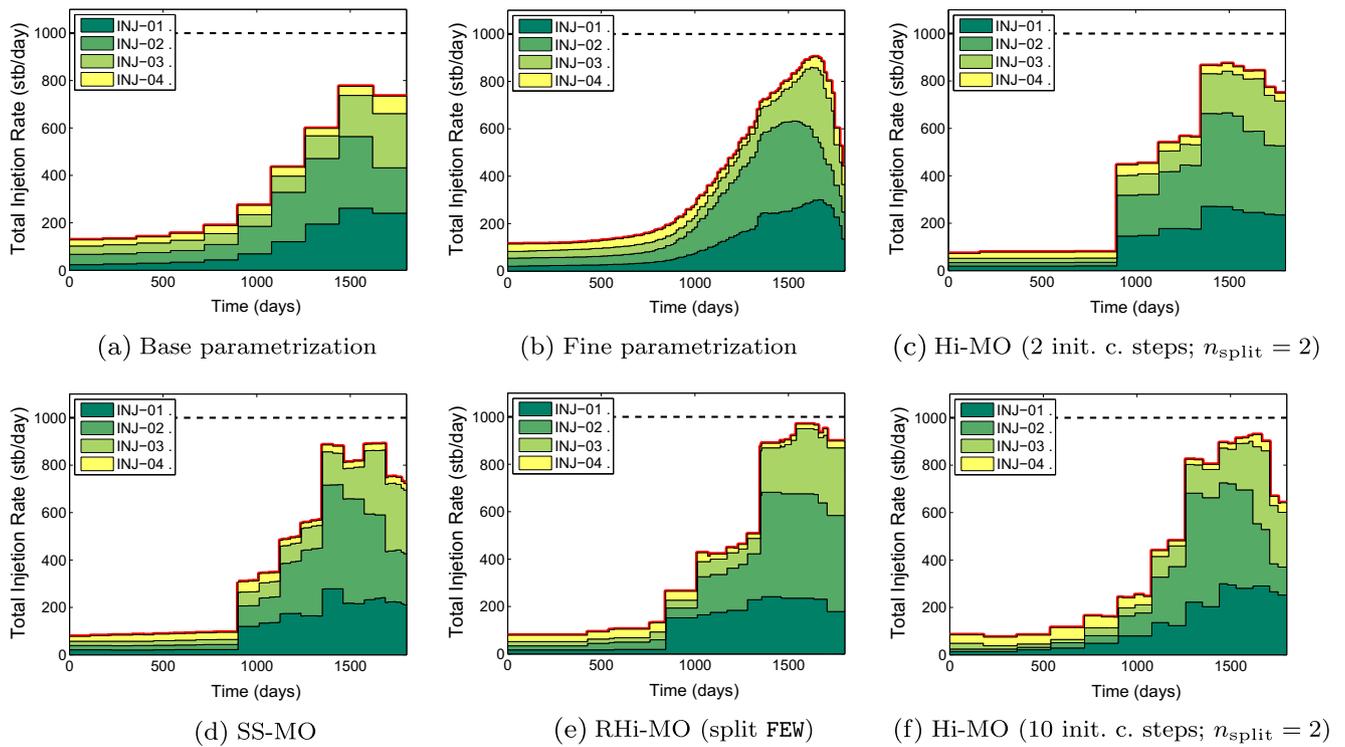


Fig. 6 Optimum total injection rate; bound plus inequality constraints, three-channel case

case we initialize with two control steps for each well and in another case we use 10 initial control steps per well. In both cases, we set $n_{split} = 2$. For RHi-MO, we again apply

the initialization strategy based on the refinement indicators and we consider again the splitting options ALL, ONE, and FEW.

Fig. 7 Depth (ft) of the top of the structure for the Brugge field example

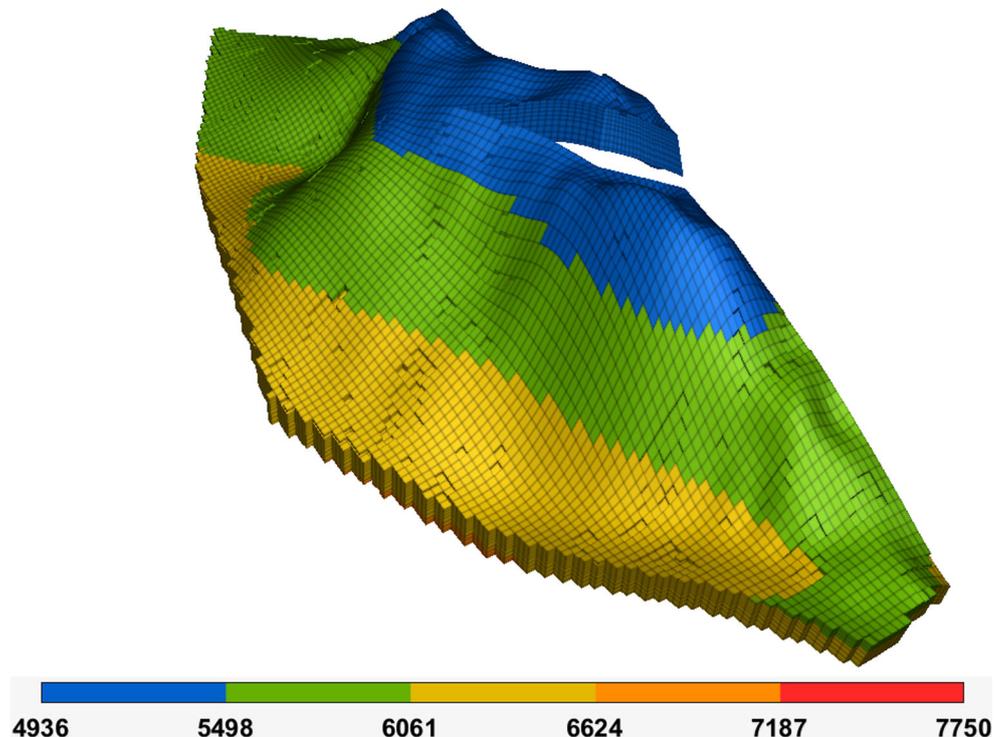


Table 4 Summary; only bound constraints, Brugge field

	NPV, $\times 10^9$ \$	# Sim.	Final N_u
Base parametrization: 40 fixed control steps	5.0730	71	3360
SS-MO	5.0922	55	1344
RI-MO (1 initial control step)	5.0481	87	246
Hi-MO (10 initial control steps; $n_{\text{split}} = 2$)	5.1228	150	1720
Hi-MO (10 initial control steps; $n_{\text{split}} = 4$)	5.1054	102	1926
RHi-MO (split ALL)	5.1298	150	753
RHi-MO (split ONE)	5.0725	58	256
RHi-MO (split FEW)	5.1005	56	501

From the results in Table 3, we note that the optimization with fine-scale parametrization performs just slightly better than the base parametrization, both in terms of the final NPV and the number of simulation runs required to obtain that final NPV. One might expect that the fine-scale parametrization problem is more difficult to solve due to the higher dimensionality of the problem, but for this constrained case, the solution of the fine-scale problem is particularly efficient in comparison to the base parametrization, leading to a higher final NPV with fewer reservoir simulation runs. However, the excessive number of control variables in the fine-scale parametrization would be more difficult to implement in field operations than the base-case controls. The results we observe from Table 3 implicitly suggest that ill-conditioning effects due to the high dimensionality of the problem have a low impact for this particular example. Because of this lack of ill-conditioning, only very small increases in NPV can be obtained with the hierarchical multiscale methods.

Regarding the linear constraints imposed for this problem, the total liquid production rates obtained with the different methods are presented in Fig. 5, and the total injection rate is shown in Fig. 6. Figures 5 and 6 present the rates cumulatively, i.e., they also show the individual contribution of each well. We can clearly see from Figs. 5 and 6 that all inequality constraints are honored at all times within the specified tolerance. In all results, the constraint on the total liquid production rate remains active at almost all control steps, whereas the limit on the total injection rate is not reached at any time by any method. Finally, we observe that the optimum controls are in general quite smooth, but SS-MO and Hi-MO with 10 initial control steps provide the roughest set of well controls.

4.2 Example 2: Brugge field

The second example corresponds to the Brugge field, which is a reservoir model built by TNO [17] as a benchmark example to analyze different approaches to closed-loop reservoir management. According to the available data, 104 geological realizations were generated and provided to the

participants in the benchmark study. A full description of the Brugge field can be found in [17].

The reservoir simulation model was built based on the geological information provided and it consists of a 138×48 areal grid and nine layers. The reservoir contains only oil and water, and the initial reservoir pressure is 2465 psi (17×10^3 kPa). The Brugge reservoir contains 30 vertical wells, including 20 smart production wells and 10 smart water injection wells. Wells are equipped with inflow control valves (ICVs) to individually control production or injection from distinct geological zones. Most of the producing wells are equipped with three ICVs, which are completed in the top three zones, however, some producers have only one or two ICVs. All the injectors are equipped with three ICVs which are completed in the bottom three geological formations. Considering all wells, we have 84 ICVs installed in the field.

The reservoir lifetime is 30 years, which includes a history-matching step for the first 10 years. We focus exclusively on the production optimization step for the last 20 years of the reservoir life. Moreover, instead of optimizing well controls based on the unknown true model, we optimize the well controls for years 10 through 30, based on the reservoir properties of one particular realization provided by TNO. We have selected the particular realization FY-SF-KP-7-33, where the reservoir is subdivided

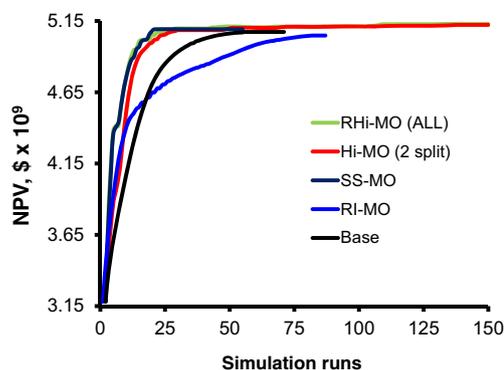
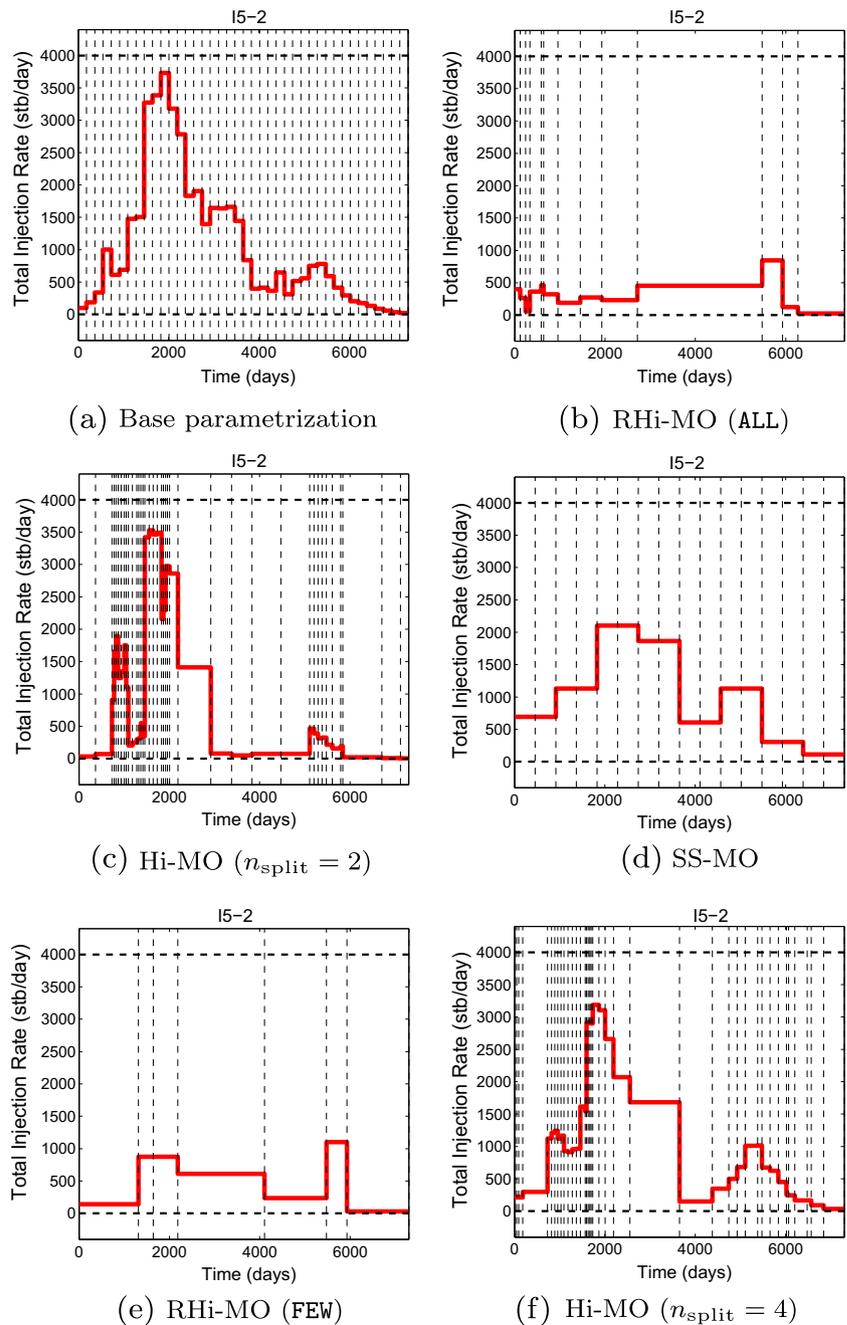


Fig. 8 Optimization performance in terms of NPV vs. the number of simulation runs required for bound constraints only, Brugge field

Fig. 9 Only bound constraints; optimum controls for well segment I5–2, Brugge field

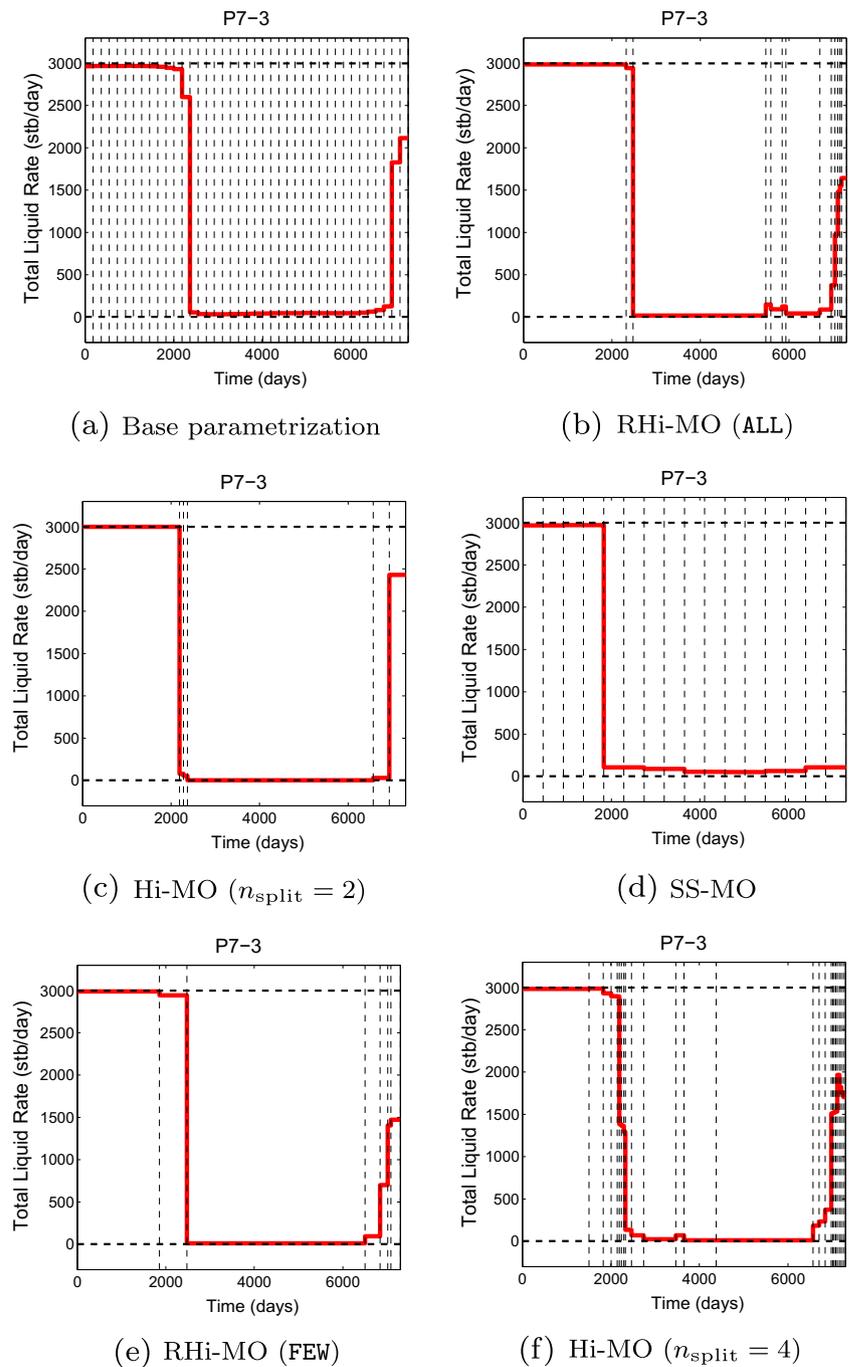


into facies classes with associated pro-permeable characteristics of the fluvial reservoir zone modeled as channel objects in a shale background. The permeability was generated stochastically with co-Kriging on porosity. Figure 7 shows the top structure map as well as the well locations.

The control variables are the liquid production rate at each individual segment for the producers and the water injection rate at each individual segment for the injectors. During optimization, the liquid production rate for each production segment is constrained to the interval [0, 3000]

STB/D with a lower bound on bottom-hole pressure (BHP) equal to 725 psi; the injection rate for each injector segment is constrained to the interval [0, 4000] STB/D with an upper bound on the injection pressure equal to 2611 psi. These pressure constraints, which represent nonlinear constraints, are not taken into account in the example in this paper. The oil price is \$80.0/STB and both the water production and the injection costs are \$5.0/STB. The annual discount rate is 0.1. For the base parametrization, there are 40 fixed control steps, each one half-year long, and 84 control variables for

Fig. 10 Only bound constraints; optimum controls for well segment P7-3, Brugge field



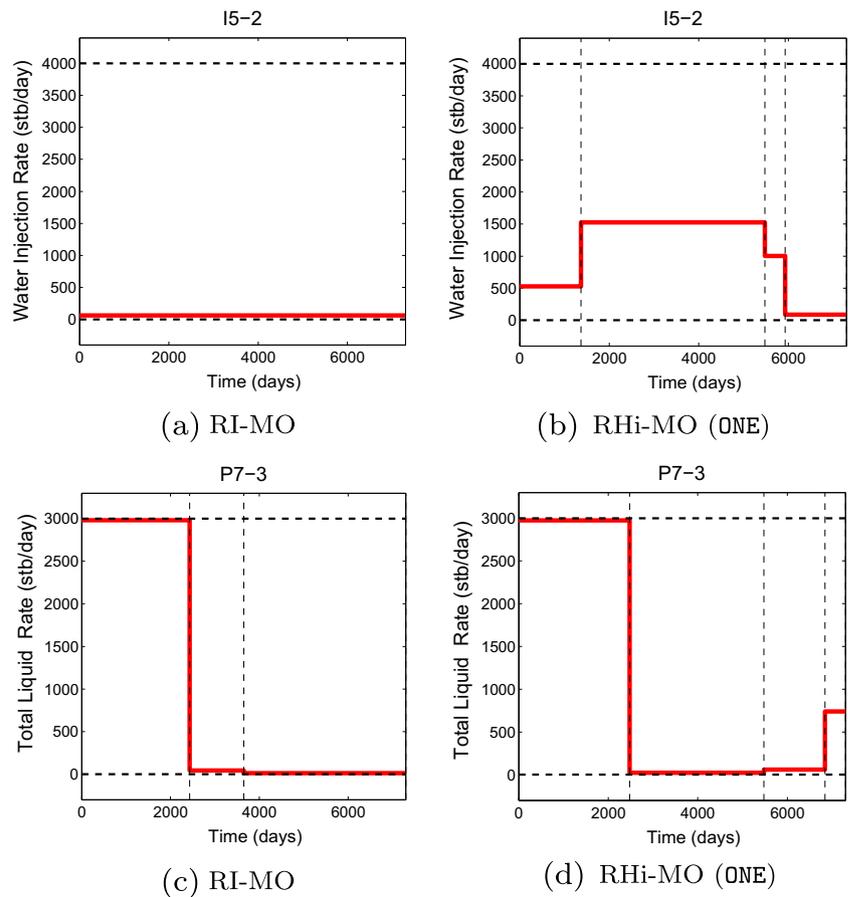
each control step, so the total number of control variables is 3360. The initial guess is 1333.3 STB/D for each injecting segment and 700 STB/D for each producing segment. Again, optimization is done in the log-transform domain.

Only bound constraints We optimize the Brugge example using the base parametrization (40 fixed control steps for each well segment), SS-MO, RI-MO, Hi-MO, and RHi-MO. Since the control variables here are the production or

injection rates at each smart completion, the bound constraints refer to the rates at each well segment. Here, we consider the steepest ascent method with adjoint-gradient algorithm in our hierarchical multiscale framework.

Table 4 summarizes the adjoint-based steepest ascent results obtained for the Brugge reservoir model when the only constraints are bound constraints. The hierarchical optimization method Hi-MO starts with 10 equally distributed control steps for each well segment and uses either $n_{\text{split}} = 2$ or $n_{\text{split}} = 4$, whereas for RHi-MO we once more

Fig. 11 Only bound constraints; optimum well controls for RI-MO [11] and RHi-MO (ONE), Brugge field



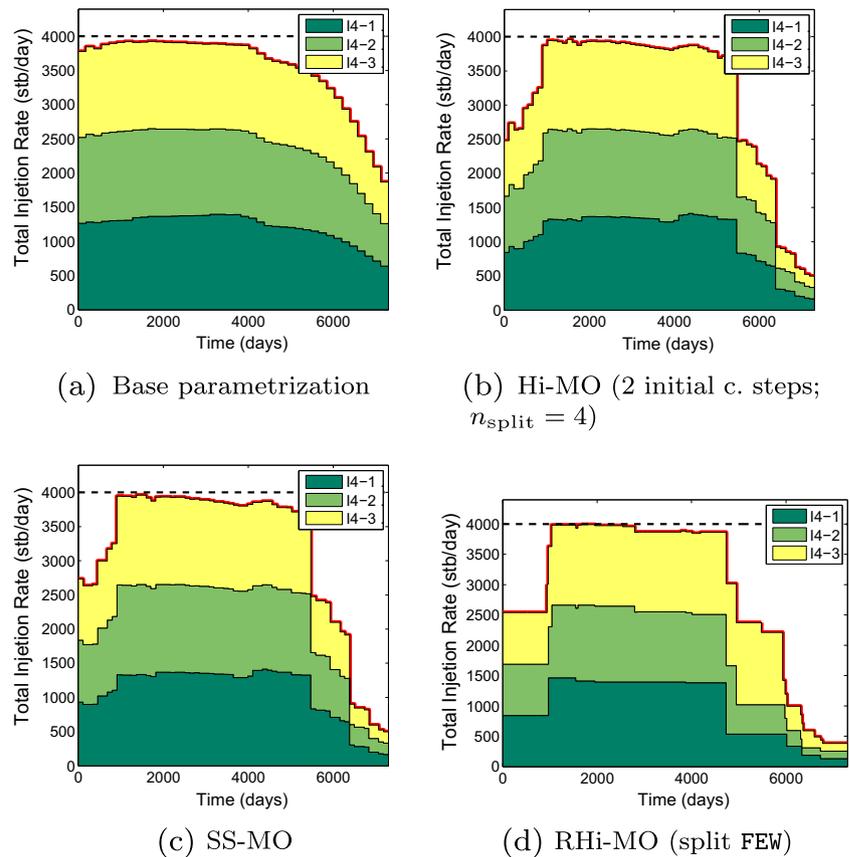
focus on the splitting options ALL, ONE, and FEW. SS-MO starts with two uniform control steps at each well segment, and for RI-MO, we consider only the case where we initialize with one single control step per well as in Lien et al. [11]. We note from the results in Table 4 that again the multiscale approaches achieve higher final values of NPV than we obtain with the base parametrization, except for RHi-MO (ONE), which reaches basically the same NPV as the base case, and RI-MO of Lien et al., whose final NPV is 0.5 % lower than we obtain with the base parametrization. Later, we present a comparative discussion on the results obtained with RHi-MO (ONE) and RI-MO.

Both results obtained with Hi-MO lead to a NPV higher than SS-MO; 0.2 % higher for $n_{split} = 4$ and 0.6 % higher for $n_{split} = 2$, however, the number of simulation runs required in both cases are larger than is required for SS-MO. In terms of the results for RHi-MO with splitting options ALL and FEW, we achieve a final NPV similar to the results obtained with Hi-MO and also higher than the SS-MO result. The best NPV is attained with RHi-MO using option ALL, a result that is 0.74 % higher than SS-MO and 1.1 % higher than the base parametrization. However, RHi-MO (ALL) requires almost three times as many simulation runs as SS-MO. On the other hand, the result for

Table 5 Summary, bound plus inequality constraints in augmented Lagrangian method, Brugge field

	NPV, $\times 10^9$ \$	# Sim.	Final N_u
Base parametrization: 40 fixed control steps	4.0749	46	3360
SS-MO	4.4587	508	5376
Hi-MO (2 initial control steps; $n_{split} = 2$)	4.4766	936	918
Hi-MO (2 initial control steps; $n_{split} = 4$)	4.5118	842	1727
Hi-MO (10 initial control steps; $n_{split} = 4$)	4.5091	840	1349
RHi-MO (split ONE)	4.4756	883	599
RHi-MO (split FEW)	4.5258	992	752

Fig. 12 Well I4: optimum controls with bound plus inequality constraints for base case, SS-MO, Hi-MO and RHi-MO, Brugge field



RHi-MO with the splitting option FEW also gives a reasonably good result in terms of the NPV, 0.2 % higher than the result with SS-MO and requires almost the identical number of simulation runs as SS-MO.

The performance of the parametrization strategies is compared in Fig. 8, which presents a plot of NPV versus the number of simulation runs required to obtain that NPV. Figure 8 contains the results for the base parametrization, SS-MO, RI-MO, and the best results obtained with Hi-MO and RHi-MO. We see that all SS-MO, Hi-MO, and RHi-MO perform similarly at early iterations and then they asymptotically converge to their final NPV. We can also see that the convergence rate becomes very slow at later iterations. Hi-MO and RHi-MO terminate due to reaching the maximum allowable number of simulation runs, which in this case is 150, whereas SS-MO terminates at its fourth level of parametrization (with 16 control steps per well segment) because the algorithm fails to find an optimum different from the optimum at the end of the third level. We can also note from the results depicted in Fig. 8 that RI-MO performs poorly for this example.

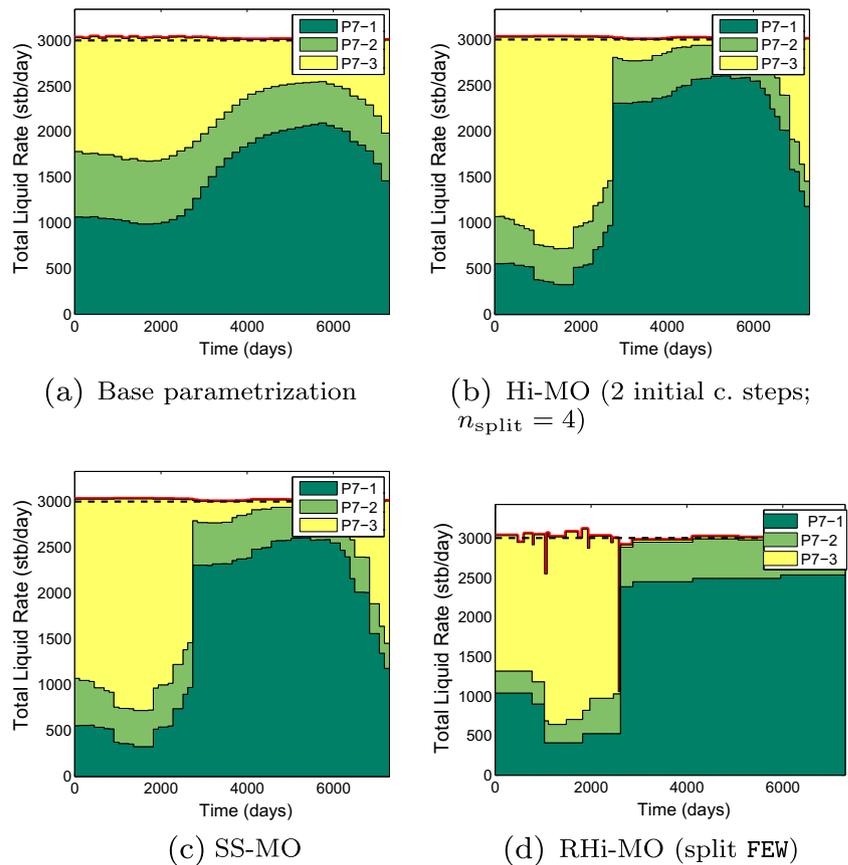
In terms of the optimal solution, Figs. 9 and 10, respectively, display the optimal well controls obtained for well segments I5-2 and P7-3. Although the methods yield distinctly different well controls, the qualitative behavior of the variation in the controls with time obtained with the

base parametrization, SS-MO, Hi-MO, and RHi-MO are vaguely similar. In general, the well controls vary smoothly, however, we can observe abrupt changes in the controls for some cases, loosely resembling bang-bang responses. We also observe for the Brugge example that the RHi-MO approaches result in far fewer control variables than the other techniques; few well control steps represents a potential advantage in terms of field implementation as fewer controls steps means we avoid excessive changes in the well settings.

As we note from the results presented in Table 4, RI-MO and RHi-MO with splitting option ONE perform poorly in comparison to the other multiscale methods. Even in comparison to the base parametrization, the results for RI-MO and RHi-MO (ONE) are not satisfactory. Apparently for this large-scale problem, the strategy of not performing many refinements is not efficient. RI-MO requires that we refine one single control variable per cycle/iteration and we note that this is not sufficient to obtain a final NPV better than the base case, although RI-MO leads to far fewer control variables than the other multiscale techniques. In fact, the optimization with RI-MO leads to a NPV 0.5 % lower than we obtain with the base parametrization and 1.6 % lower than the best NPV result.

The optimum well controls for segments I5-2 and P7-3 obtained with RI-MO and RHi-MO (ONE) are presented

Fig. 13 Well P7: optimum controls with bound plus inequality constraints for base case, SS-MO, Hi-MO and RHi-MO, Brugge field



in comparison in Fig. 11. It is observed that for RI-MO the controls for some well segments end up without any refinement (see, e.g., the solution for well segment I-5-2 in Fig. 11a). In fact, the controls obtained with RI-MO for 45 well segments end up with no refinement whatsoever. The procedure we implemented for RHi-MO, where we split one control step per well, rather than simply refine one single control variable per cycle as in RI-MO, improves the result of RI-MO both in terms of the optimum NPV and the computational cost (see Table 4).

Bound plus linear inequality constraints The Brugge example is reconsidered here; however, we add an inequality constraint that the total liquid production rate is less than or equal to 3000 STB/D at each production well at all times and that, at each water injection well, the total injection rate is less than or equal to 4000 STB/D for all times. Thus, when a particular well has multiple completions, we require that the sum of the segment liquid rates at each producer is less than or equal to 3000 STB/D and the sum of the segment water injection rate at each injector is less than or equal to 4000 STB/D. The bound constraints are that each production segment rate is in the interval [0, 3000] STB/D and each injection segment rate is in the interval [0, 4000] STB/D. As in previous examples, we use the augmented

Lagrangian method to solve the constrained optimization problem using the steepest ascent algorithm with the adjoint-gradient.

Table 5 summarizes the results for base parametrization, successive splitting, and hierarchical optimization (Hi-MO and RHi-MO) considering the inequality constrained problem. Here, we use three strategies for Hi-MO. In the first Hi-MO strategy, we initialize the multiscale optimization using only two control steps for each well and set $n_{split} = 2$, reproducing the simplicity of SS-MO. The other two Hi-MO strategies we set $n_{split} = 4$ and vary the initial set of control steps, using 2 or 10 initial control intervals per well segment. We applied RHi-MO with the splitting option FEW, because results from the previous examples show that, in terms of the three objectives of (i) maximizing final NPV, (ii) minimizing the number of simulation runs required, and (iii) minimizing the final number of control variables, it is a highly competitive method. However, we also consider RHi-MO with splitting option ONE for this example.

Again, all multiscale strategies outperform the base parametrization in terms of the optimal NPV. We observe an increase of 10.7 % in the final NPV over the base parametrization when we use either Hi-MO strategy, whereas RHi-MO (split FEW) leads to a NPV 11.2 % higher than the base case. SS-MO also improves the optimal

NPV by 9.4 % in comparison to the base parametrization. The hierarchical technique based on refinement indicators obtains the highest optimum NPV, which is \$16 million more than the best Hi-MO strategy and \$67 million more than SS-MO. Although the computational cost required by all multiscale methods is considerable, an impressive increase in the final NPV is obtained compared to the NPV of the base case. Although the optimum NPV estimated with Hi-MO using two initial control steps per well and $n_{\text{split}} = 2$ is still higher than the result obtained with SS-MO, we note that this Hi-MO strategy leads to a result inferior to the other hierarchical approaches we applied. Note that, for this example, when we initialize Hi-MO ($n_{\text{split}} = 4$) with only two control steps per well, the number of simulation runs required is about the same as when we use 10 initial control steps for each well ($n_{\text{split}} = 4$), and both estimates for the optimum NPV are quite close. In terms of the RI-based strategies, we see that using the splitting option ONE is not as good as using the splitting option FEW in terms of the final NPV obtained. RHi-MO (FEW) requires about 20 % more simulation runs than both Hi-MO approaches, but it also obtains a final NPV slightly higher (about 0.3 % higher). SS-MO utilizes fewer simulation runs than any other multiscale technique in this example. We also observe from the results in Table 5 that all hierarchical strategies (Hi-MO and RHi-MO) require far fewer control variables than the successive splitting approach and the base parametrization. In particular, RHi-MO (split FEW) obtained the highest final NPV requiring the smallest number of control variables.

Figures 12 and 13 display the optimal total injection rate for well I4 and the optimal total liquid rates for well P7 obtained with the base parametrization, SS-MO, Hi-MO with two initial control steps ($n_{\text{split}} = 4$) and RHi-MO (FEW). Rates are presented cumulatively in Figs. 12 to 13 which also show the individual contribution of each well segment. We note that, for these three wells, the solutions we obtain with the different methods are similar. In all cases, the constraints are honored within 1 % tolerance and the controls are reasonably smooth, which is preferable for field implementation.

5 Remarks and conclusions

We proposed a new hierarchical multiscale optimization strategy, RHi-MO, for adaptively selecting the number of control steps and their lengths as the overall optimization procedure progresses. Like the Hi-MO method of Oliveira and Reynolds [14], RHi-MO can merge control intervals when it is appropriate, but for the splitting step RHi-MO uses refinement indicators which potentially result in a better refinement because splitting based on refinement

indicators has a theoretical motivation. The new hierarchical multiscale optimization method was successfully applied to linearly constrained and unconstrained problems. Based on the applications considered in this work, the following conclusions are warranted:

1. The proposed multiscale methodology, RHi-MO, always outperforms the base parametrization strategy based on a fixed number of control steps which is set a priori and used throughout the optimization;
2. We always obtain a higher NPV with Hi-MO and RHi-MO than SS-MO. However, occasionally the NPVs obtained with Hi-MO and RHi-MO will be insignificantly higher than the SS-MO's final NPV;
3. Typically, the higher NPV obtained with Hi-MO will come at a higher computational cost than the SS-MO result; RHi-MO may also require more simulation runs than SS-MO in some cases, although this happens less often than it happens for Hi-MO;
4. RHi-MO consistently outperforms the method of Lien et al. [11] (RI-MO) in terms of both the final NPV and the number of simulation runs required for the examples presented here as well as an additional example shown in [16], even though both methods use refinement indicators to drive the splitting procedure. In fact, since RHi-MO can be viewed as a combination of RI-MO and Hi-MO, one should expect that RHi-MO will generally outperform RI-MO and Hi-MO;
5. In general, RHi-MO performs better than Hi-MO, both in terms of the final NPV and the number of simulation runs required;
6. The results obtained with the new proposed multiscale technique, RHi-MO, usually lead to fewer control variables than other approaches. This aspect can be relevant for field implementation, because a control strategy with few interventions at the wells would be preferable due to operational reasons.

Acknowledgments The authors are gratefully thankful for the support of the member companies of The University of Tulsa Petroleum Reservoir Exploitation Projects (TUPREP). The first author also acknowledges the support provided by Petrobras S/A.

References

1. Asadollahi, M., Nævdal, G.: Selection of decision variables for large-scale production optimization problems applied to Brugge field. In: Proceedings of the SPE Russian Oil & Gas Technical Conference and Exhibition. SPE 136378 (2010)
2. Asheim, H.: Maximization of water sweep efficiency by controlling production and injection rates. In: Proceedings of the SPE European Petroleum Conference. SPE 18365 (1988)
3. Ben Ameer, H., Chavent, G., Jaffré, J.: Refinement and coarsening indicators for adaptive parametrization: application to the

- estimation of hydraulic transmissivities. *Inverse Prob.* **18**, 775–794 (2002)
4. Brouwer, D.R., Nævdal, G., Jansen, J.D., Vefring, E.H., van Kruijsdijk, C.P.J.W.: Improved reservoir management through optimal control and continuous model updating. In: *Proceedings of the SPE Annual Technical Conference and Exhibition*, Houston, Texas, 26–29 September. SPE 90149 (2004)
 5. Chavent, G.M., Bissell, R.: Indicators for the refinement of parametrization. In: Tanaka, M., Dulikravich, G.S. (eds.) *Inverse Problems in Engineering Mechanics*. Elsevier, Amsterdam (1998)
 6. Chen, C.: Adjoint-gradient-based production optimization with the augmented Lagrangian Method, PhD. thesis, University of Tulsa, Tulsa, Oklahoma, USA (2011)
 7. Chen, Y., Oliver, D.S., Zhang, D.: Efficient ensemble-based closed-loop production optimization. *SPE J.* **14**(4), 634–645 (2009)
 8. Conn, A., Gould, N., Toint, P.: A globally convergent augmented langrangian algorithm for optimization with general constraints and simple bounds. *SIAM J. Numer. Anal.* **28**(2), 545–572 (1991)
 9. Gao, G., Reynolds, A.C.: An improved implementation of the LBFGS algorithm for automatic history matching. *SPE J.* **11**(1), 5–17 (2006)
 10. Jansen, J., Brouwer, D., Naevdal, G., van Kruijsdijk, C.: Closed-loop reservoir management. *First Break* **23**, 43–48 (2005)
 11. Lien, M., Brouwer, D.R., Mannseth, T., Jansen, J.D.: Multiscale regularization of flooding optimization for smart field management. *SPE J.* **13**(2), 195–204 (2008)
 12. Lorentzen, R.J., Berg, A.M., Nævdal, G., Vefring, E.H.: A new approach for dynamic optimization of waterflooding problems. In: *Proceedings of the SPE Intelligent Energy Conference and Exhibition*. SPE 99690 (2006)
 13. Nocedal, J., Wright, S.J.: *Numerical Optimization*. Springer, New York (2006)
 14. Oliveira, D.F., Reynolds, A.: An adaptive hierarchical multiscale algorithm for estimation of optimal well. *SPE J.* **19**(5), 909–930 (2014). SPE-163645-PA
 15. Oliveira, D.F., Reynolds, A.C.: Hierarchical multiscale methods for life-cycle production optimization: a field case study. *SPE J. Preprint*. SPE-173273-PA (2015)
 16. Oliveira, D.F.B.: A new hierarchical multiscale optimization method: gradient and non-gradient approaches for waterflooding optimization, Ph.D. thesis, The University of Tulsa, Tulsa, Oklahoma (2014)
 17. Peters, L., Arts, R., Brouwer, G., Geel, C., Cullick, S., Lorentzen, R., Chen, Y., Dunlop, K., Vossepoel, F., Xu, R., Sarma, P., Alhuthali, A., Reynolds, A.: Results of the Brugge benchmark study for flooding optimisation and history matching. *SPE Reserv. Eval. Eng.* **13**(3), 391–405 (2010)
 18. Sarma, P., Chen, W., Durlafsky, L., Aziz, K.: Production optimization with adjoint models under nonlinear control-state path inequality constraints. *SPE Reserv. Eval. Eng.* **11**(2), 326–339 (2008)
 19. Sarma, P., Durlafsky, L., Aziz, K.: Implementation of adjoint solution for optimal control of smart wells. In: *Proceedings of the SPE Reservoir Simulation Symposium*. SPE 92864 (2005)
 20. Sarma, P., Durlafsky, L., Aziz, K., Chen, W.: Efficient real-time reservoir management using adjoint-based optimal control and model updating. *Comput. Geosci.* **10**, 3–36 (2006)
 21. Schlumberger Ltd.: *ECLIPSE Reservoir Simulation Software: Reference Manual*, Schlumberger Software, London, UK www.slb.com, version 2011.2 edn., 2011
 22. Shuai, Y., White, C.D., Zhang, H., Sun, T.: Using multiscale regularization to obtain realistic optimal control strategies. In: *Proceedings of the SPE Reservoir Simulation Symposium*, The Woodlands, Texas, USA, 21–23 February. SPE 142043 (2011)
 23. Wang, C., Li, G., Reynolds, A.C.: Production optimization in closed-loop reservoir management. *SPE J.* **14**(3), 506–523 (2009)
 24. Zakirov, I.S., Aanonsen, S.I., Zakirov, E.S., Palatnik, B.M.: Optimizing reservoir performance by automatic allocation of well rates. In: *Proceedings of the 5th European Conference on the Mathematical Oil Recovery - Leoben, Austria*, 3–5 September (1996)
 25. Zhao, H., Chen, C., Do, S., Oliveira, D., Li, G., Reynolds, A.: Maximization of a dynamic quadratic interpolation model for production optimization. *SPE J.* **18**(6), 1012–1025 (2013)