Reply to reviews of COMG-D-13-00011, "Adjoint formulation and constraint handling for gradient-based optimization of compositional reservoir flow" by D. Kourounis, L.J. Durlofsky, J.D. Jansen and K. Aziz

We thank the reviewers for the time they spent reviewing this paper and for their useful and constructive comments. We have modified the paper based on their suggestions, and we believe the revised version addresses their concerns. Below are our detailed responses, with the reviewer comments in italics and our responses in normal font.

REVIEWER 1

Manuscript COMG-D-13-00011 introduces a computational approach for optimization problems in an oil-gas compositional reservoir model. Based on a gradient formulation, the proposed optimization technique uses automatic differentiation to implement the adjoint of the reservoir model. One critical issue of the optimization problem with the compositional model is the incorporation of nonlinear constraints. For example, under controlled bottom hole pressure, the challenge is to incorporate the specification of the maximum gas in the production/injection wells. The paper proposes two approaches for handling the aforementioned constraints. In the "formal" approach the nonlinear constraints are apparently incorporated within the optimization routine. The second "heuristic" approach is an ad-hoc method that uses the forward model to enforce the constraints. The manuscripts provides experiments to show the capabilities of the proposed approaches for incorporating the constraints into the solution to the optimization problem. From the numerical results the authors conclude that the "heuristic" approach provides a less computationally costly method for handling the constrains. In addition, the authors discuss the differences and similarities between discrete and continuous adjoint formulations.

I believe that the manuscript addresses interesting computational aspects of solving practical problems with large-scale complex reservoir models like the oil-water compositional model under consideration. I also believe that the proposed approach based on automatic differentiation is worth of a thorough review. In my opinion the manuscript is reasonably well written and satisfies the standards of quality for its publication in Computational Geosciences. However, I consider that the scientific contribution of the paper is not sufficiently clear/emphasized. The numerical experiments are interesting but they should only reflect the main point of the proposed work which seems to be to efficiently incorporate nonlinear constraints in such complicated models. Below I describe some concerns that I believe should be addressed so that the manuscript can be considered for publication.

If I understood correctly, the "formal" approach is the proper mathematical way to handle the constraints. If this is correct, then it should be made more clear in the manuscript and the corresponding mathematical treatment should be included. Although in Section 3 the authors present both the continuous and discrete versions of the optimization problem, they do not include the nonlinear constraints(which seems to be the critical point!). Indeed, it is stated "In general, a number of linear and nonlinear constraints may need to be included in the optimal control problem. We postpone the discussion of their treatment until Section 4". However, in Section 4 it does not be-

come clear the precise optimization problem that is solved with the formal approach. Given the fact that incorporating the nonlinear constraints seems one of the central aspects of the paper, a more thorough discussion of the precise optimization problem should be provided (regardless that references [27] and [15] are included). More concretely, I believe Section 3 should be rewritten including the nonlinear constraints and thoroughly discussing the approximations/implementations performed. In other words, equations (4.1)-(4.4) should be omitted to the formulation of Section 3. I presume that Section 3 does not include the nonlinear constraints so that it is easier to see the differences between the discrete and the continuous adjoint. I personally believe that the continuous vs discrete adjoint has been sufficiently discussed in other references, and for the sake of this publications it suffices to say that the discrete adjoint (as usual) is the preferred one. I am not suggesting to remove it but just to include the nonlinear constraints in the formulation.

If my understanding of the paper is correct and the formal approach is the proper mathematical technique to incorporate the nonlinear constraints, then the results of experiment 3 are misleading. I understand that a "formal" (or more rigorous) approach may be computationally unfeasible and inaccurate unless fully resolved. However, the more mathematical way should always be the gold standard against which to compare approximate techniques. Therefore, I consider that the heuristic approach should also be compared in the case where the formal result gives the most accurate solution, even if that means to use more "control steps" or fully resolve the problem (and/or consider a smaller experiment if computations are overwhelming). Once the fully resolved exact case is obtained (highlighting the potential high computational cost involved), a less resolved example (like example 3) may be used to expose the loss of accuracy in more practical (less resolved) problems for which the heuristic approach outperforms the formal one in terms of accuracy and efficiency (properly quantified in terms of forward simulations).

In summary, if the "formal approach" is the proper mathematical way to handle the constraints, the manuscript should provide the full mathematical derivation that includes the how these constraints are handled within the optimizer. I believe this is essential to transfer (and reproduce) the scientific contribution of this paper (without necessarily having to learn the SNOPT optimizer). The numerical experiments should also include the case where the formal approach provides the most accurate solution against which to compare the heuristic approach.

We believe both of the reviewer's key suggestions – adding more detail on formal nonlinear constraint handling, and demonstrating that the formal treatment does indeed outperform the heuristic approach in a (tractable) case with more control steps – are good ones, and we have modified the paper accordingly. Specifically, we have added a new section (Section 4, p. 6-7) on "Gradient-based optimization and related software" that describes the SQP approach in some detail, along with discussion of the treatment of nonlinear constraints. Additional description on the treatment of nonlinear constraints appears in Section 5.1 (Constraint handling in the optimizer). Our discussion in new Section 5.1 expands on the discussion in the original paper (which appeared in Section 4.1 in the earlier version). In total, we believe that the new text in Sections 4 and 5 provides a substantive description of the formal treatment of nonlinear

constraints.

Regarding the reviewer's second key point, along the lines of his/her suggestion, we have added a new example case (Example 1, starting on p. 10) that clearly shows the formal constraint handling procedure outperforming the heuristic approach with increasing number of control periods. This is indeed as would be expected if the problem remains sufficiently "tractable." More specifically, for this case, which entails a simple reservoir description (see Figure 6.1), we see that with 8 control steps the heuristic approach outperforms the formal approach by about 2%. With 64 control steps, however, the formal approach now outperforms the best heuristic solution (by 4.1%). For this example, the formal approach provides a solution that exceeds the feasible reference case by 11.6%, which is a substantial improvement.

We have also added new discussion regarding the relative performance of the two approaches. On p. 11, left column, 4th paragraph, we now state: "Theoretically, if the optimization problem remains sufficiently 'tractable', as we increase the number of control variables the formal approach should (eventually) outperform the heuristic approach. However, if the optimization problem becomes significantly more difficult with increasing numbers of control variables (which may be related to the constraint lumping procedure), or if a large number of local optima associated with relatively poor objective function values appear, then the formal approach will not necessarily outperform the heuristic treatment." For Example 1, the optimization problem does remain sufficiently tractable, while for old Example 3 (which is now Example 4) it does not, which is presumably why the heuristic procedure outperforms the formal approach for that case. We have also added a paragraph discussing this issue in Section 7 (Concluding remarks; see paragraph starting "In the simplest case ..." on p. 19).

In summary, we believe we have fully addressed the two important issues raised by Reviewer 1.

REVIEWER 2

The authors present an adjoint-based approach for optimal oil recovery using a compositional model. They study aspects of obtaining gradients (discrete/continuous adjoints, and automatic differentiation), discuss the handling of constraints and present three case studies. The main contributions of the manuscript are a discussion and implementation of adjoint-based gradients for a rather complicated flow model, and the application of the resulting algorithm for realistic optimization of oil recovery. For the optimization itself the publicly available SQP implementation SNOPT is used.
The paper is well written and most issues/challenges are clearly explained. I've a few comments and suggestions below, most of them are concerned with connecting the research better to the optimization/adjoint methods literature. However, the paper is an interesting contribution and I will support–after the authors incorporate/reply to the comments below–the publication of the manuscript.

Comments

1. *The discussion on continuous vs. discrete adjoints is restricted to the time discretization, i.e., an issue at the level of ordinary differential equations. In the applied maths literature, this is well studied for general (Runge Kutta) time step-*

ping methods [5] [6]. In [7], the same discrepancy with respect to the continuous and discrete final time condition as in the manuscript is reported. Note that a similar gradient discrepancy between discretization/optimization vs. optimization/discretization can also occur with respect to the spatial discretization, see e.g., the discussion on a shape optimization problem in [8]. I enjoyed reading your presentation and explanation of this discrepancy, but please put your findings in a somewhat larger perspective.

We thank the reviewer for providing us with these references. We have updated our discussion in Section 3.4 (Continuous versus discrete adjoint formulation; discussion appears on p. 5, right column, last paragraph) to now include references [5], [6] and [7] (from the reviewer's list), and we also now note that a similar problem may occur with respect to the spatial discretization, citing reference [8].

2. *Please introduce the abbreviation BHP for readers not familiar with reservoir modeling*

   Thank you for catching this oversight. We now state, on p. 7, left column, 2nd paragraph: "bottom-hole pressure, or BHP, is the wellbore pressure at a specified depth within the reservoir."

3. *I was not able to find a discussion or reference describing the methods and implementation of the reservoir modeling code...please add a pointer and/or a short description of the methods used there.*

   This is discussed in Section 2 (p. 2-3), and several key references [7, 12, 39, 43] are provided there. The solution of the linear adjoint system is discussed in Section 3.5; the key reference given there is [19].

4. *Page 3, middle of column 2: I would not claim that AD is a recent development. This is quite a classical field by now, see for instance [3], [4] and references in these books.*

   In the revised version, in this discussion (p. 3, right column) we no longer refer to AD as a recent development.

5. *Page 4, equation (3.8): What is f_n here? Should this be f_A? Perturbations of the n-th cost term make limited sense since the adjoint variable couples the entire time history, i.e., $\boldsymbol{\lambda}_n$ depends on $\boldsymbol{\lambda}_k$ with k>n, and thus also on all $\vec{x}_n$. Same concern in equation (3.16).*

   Equation (3.8) has been rewritten and the notation problem has been clarified. Text has been added (p. 4, right column, last paragraph in Section 3.2) explaining gradients at time steps and at control steps, and how the latter (which are forwarded to the optimizer) are obtained from the former. Equation (3.16) in the original version has been removed to avoid unnecessary repetition, since it is the same as that for the discrete adjoint formulation (a pointer to equation (3.8) has been added instead).

6. *Is (4.1) not simply a bound constraint that holds for each time step? What is the advantage of formulating it as (4.2) and thus artificially increasing the nonlinearity?*

   Equation (4.1) in the original version (equation (5.1) in the revised version) represents a nonlinear constraint because the control variables are BHPs, and rates and BHPs are related nonlinearly (i.e., a nonlinear simulation is required to compute rates from BHPs). This is now explained in Section 5.1.2 (p. 8, left column, 1st paragraph), where we state: "Many constraints appear as simple bound constraints (e.g., BHP limits in a problem where BHPs are the control variables), but in other cases the constraints are nonlinear since a (nonlinear) simulation is required to evaluate them. Examples of this are the specification of maximum gas injection or production rates (either for an individual well or for a group of wells) in a general compositional problem where the control variables are BHPs."

7. *Page 14, first column, line 7: "terminating boundary condition" seems unusual. What about calling this the "final time condition"?*

   We now use "final time condition" instead of "terminating boundary condition."

8. *Discussion on page 6/7: Why is the number of control steps limited? Is the performance of SNOPT that depending on the size of the problem? Under certain assumptions, SQP performs similarly from the problem discretization ("mesh independence"), since it approximates the underlying continuous problem. Would it make sense to start with a coarser control discretization, solve the problem and then refine the control space (see also the next point).*

   It certainly would make sense to start with a coarser control discretization and then sequentially refine the control space. We believe, however, that the detailed exploration of such a procedure is beyond the scope of this paper, though this is a topic we intend to pursue in future work. This type of approach has been shown to be effective in production optimization problems by other investigators, and it would clearly be effective in new Example 1 for the heuristic procedure, where the best optimum using 8 control periods exceeds that using 64 control periods. At the end of our discussion of Example 1 (p. 11, right column, 2nd paragraph), we have therefore added the sentence: "These results suggest that it may be useful to explore the use of a sequence of optimizations, with increasing numbers of control periods, for production optimization problems." We added a similar sentence (noting this as a consideration for future work) in the Concluding remarks section (p. 19).

9. *A significant limiting factor in the optimization seems to be the existence of multiple local minima, in which the optimizer gets trapped. Thus, depending on the initialization, very different (local) solutions are found. Have you tried a continuation with respect to the number of control parameters, e.g., start with two unknowns per well, solve the optimization problem, use the obtained solution as initialization for an optimization problem with 4 unknowns per well, etc.? Often, the cost has less local minima if the control space is coarse and the solution for*

the coarse space is a good initialization for the next finer control parameterization that help avoid local minima. A similar continuation could also include upscaled models of the permeability. Similar ideas have been successful for other optimization problems where multiple local minima are a problem, e.g., in parameter estimation problems involving wave equations [1] [2]. Have you tried to use the heuristic solution as initialization for the optimizer?
I consider 9) only as suggestion; the present paper is sufficiently interesting to be published without implementing such a continuation procedure, but I am confident that such a procedure will help significantly.

Consistent with our response to Comment 8 above (regarding the use of an increasing number of control periods), we believe these to be excellent suggestions. In addition to our response above, please note that in current research we are in fact investigating the use of upscaling for (related) optimization problems, and preliminary results are very promising. But our work on this topic is still at an early stage, so we prefer not to discuss this approach in this paper.