

Author's responses to Reviewers' comments:

Thank you very much for reading the manuscript carefully and providing useful suggestions. The following is our response to your suggestions and comments one by one.

Reviewer #2:

1. **Reviewer:** *We thank the authors for addressing all points raised in our report. In particular, we appreciate the new numerical experiments and the modifications explicitly stating the necessary assumptions. We kindly ask the authors to comment in the text on the results by SPGL1. It seems that the recovery errors of SPGL1 suggest that the solver computed critical points of poorer quality in the interior of the feasible set. Is that correct?*

Response: Thanks for the comment. There seems to be some confusions here (which is likely due to mis-communication). Below we will clarify on this.

First of all, the SPGL1 algorithm is not solving the same problem as $\text{FPA}_{\text{retract}}$, but a problem with the objective function replaced by the group Lasso $\sum_{j \in \mathcal{J}} \|x_j\|$. These problems may have different sets of critical points as their objective functions are different. Thus, instead of commenting on the quality of critical points, we have now added the following comments concerning solution residual at the end of Section 6.3.1: “the solutions returned by $\text{FPA}_{\text{retract}}$ and ESQM_{ls} tend to have smaller residuals (in absolute value) than those obtained by SPGL1”.

Furthermore, concerning the comment “in the interior of the feasible set”, while this comment is true for the test instances described in Section 6.1, it may not be true for those instances in Section 6.2. Indeed, the problem that SPGL1 solves in the latter case is (6.8), whose constraint set is contained *inside* the feasible set of the original problem. Thus, for this latter class of problems, it is natural to expect SPGL1 to return a solution in the interior of the feasible set of the original problem.

2. **Reviewer:** *P3, L12-14: You mention that “Thus, the computational effort solving each subproblem is similar in our method.” This claim is invalid for the Supporting Hyperplane Method (SHM) variant of [1] that requires solving a strongly convex quadratic program per iteration (assuming C is a polyhedron) instead of a convex program as in your method.*

Response: Thanks for the comment. We believe that there could be some misunderstanding. In this sentence, we did not mean to compare our methods with [1]. We just wanted to emphasize that each subproblem in our method bears similar computational cost. To avoid misunderstanding, we rephrased it as “Thus, each subproblem in our method bears a similar computational cost”.

Communicating Editor:

1. **Reviewer:** *I know of a bundle method having a low computational cost per iteration and yet being able to handle more general DC programs than those considered in the revised manuscript. This bundle method, published in <https://doi.org/10.1007/s10589-020-00241-8>, does not require differentiability of either objective or constraint function, does require a Slater point, and needs solving only a strongly convex QP program of limited size per iteration. In the interest of completeness, I ask the authors to consider this work in the context of their own contribution.*

Response: Thanks for suggesting this interesting and highly related reference. Like ESQM_{ls} , this bundle method is an infeasible method, while our $\text{FPA}_{\text{retract}}$ is a feasible algorithm. This reference is now cited in the introduction.