We thank the reviewers for the thoughtful review of our work. We would like to briefly clarify a few minor points and will happily incorporate the suggested improvements and changes to future drafts of this work.

There are four main contributions in the paper: 1) defining the multicriterial dimensionality reduction problem, and observing on real-world datasets the failure of standard techniques to adequately solve this problem, 2) designing and analyzing new algorithms to solve multicriteria dimensionality reduction 3) analyzing the complexity of solving the general multicriteria dimensionality reduction problem (for constant k and then for general k) 4) empirically evaluating our algorithms.

The theoretical results are original, rather than background material known prior to this work. We will happily separate the relevant background content into a preliminaries section to make this distinction clear. We agree with R5 that extremal solutions to mathematical programs have found use in many contexts, and our analysis relies on this as well as new insights to the specific problem we formulate.

Specifically responding to R3, we agree that our work doesn't measure the 'downstream' effects of different amounts of reconstruction error; doing so seems interesting but domain-specific. We will spend some time thinking about an appropriate setting for measuring this.

Regarding R4's question about MW: we implemented MW to the same datasets and tested its performance and runtime. MW, in fact, runs very efficiently on both datasets (within few seconds) and scales even up to random instances of original dimensions 1000. The empirical objective value of MW produces optimal solutions in many cases. Moreover, we have a theoretical understanding as to why: in these real datasets, we find no degeneracy in SDP feasible set, and hence the solution is necessarily unique and extreme. This implementation is publicly available at the same link in the paper.

We note however some limitations to MW in this setting. The framework of MW in Samadi et al does not apply to as many objectives as SDPs. MW depends on solving linear constraints, and so the objective that cannot be obtained from applying min/max to linear functions cannot be solved by MW. For example, the marginal loss objective is solvable by MW, but not Nash social welfare. Also, one must tune the parameter of MW and hope that it will converge faster than the theoretical bounds (Samadi et al mentioned 10-15 steps on real-world datasets but theoretically the runtime bound is much looser).