

1 We would like to thank all three reviewers for their careful review and constructive feedback. We have addressed points
2 made by reviewers 1 and 3 below. We added a reference to (Basse and Airoidi, 2015); other references can be found in
3 the main paper.

4 **Reviewer 1**

5 “*In the introduction the connection to optimal experimental design is alluded to.*” Thank you for this feedback. If
6 accepted, we would be happy to detail the connection with the optimal experimental design literature. Because our
7 cluster randomized design is chosen so as to optimize a statistical criterion (the empirical exposure variance) within
8 the class of cluster randomized designs, our approach fits squarely within the optimal experimental design literature.
9 Canonical work by Raudenbush (1997) explores optimal cluster randomized designs under a different optimization
10 objective than ours since no violations of SUTVA are considered in their paper. More recently, Basse and Airoidi (2015)
11 consider a non-bipartite setting where interference is present but explores optimal allocation strategies within a given
12 graph, rather than finding an appropriate clustering. Like many traditional work on optimal designs, our objective is
13 tied to the Fisher information matrix of an inference problem. In fact, conditionally on a random assignment and under
14 a linear exposure model with Gaussian noise, maximizing the Fisher information matrix is equivalent to maximizing the
15 empirical exposure variance, as detailed in the proof of Proposition 2.

16 “*It is not clear what the quality of the approximation is for the heuristic presented [...]*” We agree that the quality of the
17 approximation for the proposed heuristic could have been made clearer. As a measure of quality, we can compute how
18 well our heuristic approximates an upper-bound of the objective. When treating 10% of all diversion units, the maximal
19 exposure variance we can hope for is that each outcome unit individually gets exposure 1 with probability .1 and 0
20 otherwise. The value of this optimistic upper-bound is $\sqrt{.1 \times .9} = .3$. The objective value of .23 that we achieved with
21 our algorithm and reported in Fig. 1.a. is therefore a 76% approximation of the upper-bound.

22 “*What is the space complexity of the proposed heuristic?*” Thank you for bringing this omission to our attention, which
23 we would be happy to clarify in a final version of the paper if accepted. The folding stage has worst-case $O(M^2)$ space
24 complexity if the resulting folded graph is complete, where M is the number of diversion units. In practice, we found
25 that most folded graph were sparse and that filtering low-weight edges in certain graphs did not substantially affect the
26 quality of the final solution, thus reducing the practical space complexity. The correlation clustering heuristic stage uses
27 space linear in the size of the folded graph since each of the maintained data structures has size linear in the number of
28 diversion units (cf. l. 264), plus the adjacency list of the folded graph.

29 “*Why are other correlation clustering heuristics not compared against in the experiment section?*” During the devel-
30 opment phase, we explored several variations of the local search heuristic before settling on the one presented in the
31 paper. We chose local search because it has been shown to perform well in practice, while being very scalable unlike
32 semi-definite-programming-based solutions (Elsner and Schudy, 2009). If accepted, we would be happy to include
33 comparisons to other attempted local search heuristics.

34 “*There are no results given for the variance of the estimators / designs [...]*” Thank you for this feedback. By reporting
35 the mean-squared error of the estimators / designs in Figure 1.b., we include both bias and variance under one metric:
36 $MSE = bias^2 + variance$. If accepted, we would be happy to plot bias, variance, and coverage in separate figures. We
37 would also be happy to plot the empirical relation between each of the three objective values for each algorithm: the
38 maximizing-agreement correlation-clustering objective, the empirical exposure variance, and final estimator variance.

39 **Reviewer 3**

40 “[...] *a more thorough explanation of novelty may be required.*” We have evidence of significant improvements using
41 this methodology over previous baselines in a real-world setting. Unfortunately, that dataset cannot be made public. We
42 repeated the experiment on a publicly-available Amazon dataset for which we saw clear gains over prior art.

43 “*There is no conclusion section.*” Due to space constraints, we did not include a conclusion section but we would be
44 happy to include one that recaps the main contributions of the paper and suggests future work: understanding how
45 estimation can be improved with generalized propensity scores and finding a similar algorithm to cluster the outcome
46 units are two directions that come to mind.

47 **References**

48 Guillaume W Basse and Edoardo M Airoidi. 2015. Optimal design of experiments in the presence of network-correlated
49 outcomes. *ArXiv e-prints* (2015).