We thank the reviewers for their insightful feedback. We are encouraged that they found our approach to be interesting (R2, R4) and distinct from existing approaches (R1) with thorough and detailed experiments (R3) and reasonable baselines for a fair evaluation (R1). We are pleased that R1 recognizes the novelty and value of using *both* upper and lower bounds, *in contrast to existing approaches*, for resource allocation and for the provision of performance guarantees (regret) that current methods lack. Indeed, incorporating existing estimators of MI that are *biased in known directions* as bounds (rather than proxies for true MI) is the critical insight that directly leads to both algorithmic improvement *and* performance guarantees.

(R2, R4) **Baselines** R1 feels the paper does "a good job of including reasonable baselines" while R2 and R4 prefer comparison to additional MI bounds. We emphasize that our goal is not to identify the most accurate MI proxies, but to propose an approach which exploits available bounds to guarantee performance with minimal computation. While we consider specific bounds (Eqn. 4) **other bounds are easily substituted** including variational bounds suggested by R2 (Supplement Sec 3). R4 considers the comparisons to be "a variant of proposed methods", which we disagree with since the typical Bayesian optimal experimental design (BOED) approach uses our chosen bounds ('bed-lb', 'bed-ub') as proxies [5, 2, 3]. Additional comparisons suggested address a different, *continuous design*, problem (R2 [3], R3 [4]).

(R3) **Dimension Experiment or Discussion** R3 is concerned that the experiments are 1D designs. Design dimension is relevant only for continuous designs, whereas in discrete settings the number of distinct design elements is a better measure of complexity. For the Gaussian MRF we use a set size of 100 and for the Tracking experiment there are 6669 choices. We do not have experiments explicitly analyzing the impact of increasing set size, but expect our approach to yield greater savings (w.r.t. baseline) as allocating resources to promising designs is increasingly important.

(R2) **One needs to compute the expected information gain (EIG), requiring a nested estimator, given Eq(2), right?** Not exactly. The bounds (*e.g.* Eqn. 4) *are typically used as proxies for the true MI (Eqn. 2)*, but we explicitly treat them as bounds. We select the design with the highest performance guarantee (L104-110) which is afforded by two-sided bounds. We use nested estimators in our experiments because they are simple – and common in the BOED literature [1, 3]. One could use (non-nested) alternatives (Supp. Sec 3), but we illustrate the benefits of an approach incorporating two-sided bounds. We will also add references as suggested by R1, R3.

(R2) **Added cost of knapsack algorithm, scaling with variables** The algorithm only adds a small cost ($< .01\%$ of the total computational cost for GMRF experiment) because the marginal utility (MU) of each design doesn't require any computation over the samples. In general, cost of bound evaluation (quadratic in samples) will far outweigh that of knapsack. The knapsack cost is *linear* in the number of designs since the MUs depend on each lower and upper bound.

(R2) **Motivation for the cost function formulation (L125) unclear.** The computational cost arises from the particular bounds used in the BOEDIR framework. The nested bound evaluations in Eqn. (4) are quadratic in the total number of samples ($= |\mathcal{Y}| + N$ where N is the incremental update when refining), resulting in the cost function of L125. The costs take a different form for other bounds (Supp. Sec 3).

(R3, R4) **Estimating Costs** The costs for sampling can be directly estimated using any method for measuring code performance, including functions that measure wall time. The coefficients of the bounding function can be estimated by a quadratic fit to timing measurements at various sample sizes. Alternatively, one could learn these parameters online; measuring and adaptively estimating the computational cost adds little computation.

(R2) **Does Eq(4) need to be computed for each experimental design setup?** Yes, we bound EIG of each design (Eqn. 4) with an initial amount of computation. This may suffice to exclude some designs from further computational resources; over half of the designs in the tracking experiment do not receive additional evaluation.

(R1) **Why are approximations in Sec. 3.1 made?** One can exactly evaluate the change in performance guarantee under an assumed update to the lower/upper bounds. However, the result is sensitive to the update assumption due to the discontinuous $\max$ function, so we use a standard smooth approximation: LogSumExp.

(R4) **Selection of the refinement set should be described.** The refinement set, $\mathcal{R}$ in Alg. 2, consists of all designs with upper bound greater than the highest lowest bound. These are all designs that may feasibly be optimal.

(R1, R3, R4) **Presentation** In addition to comments above we will: clarify the definition of $g(I_a, I_a^*)$ (R1), give an example of cost parameter estimation (R3, R4), explicitly reference Alg. 1 (R3), discuss suboptimality gap of greedy (myopic) vs. non-myopic BOED (R3), and expand derivations of the bounds (Eqn. 4) in the supplement (R3). Additions to the main text are minor, but we will shift details of the GMRF experiment to the supplement (R1) for extra space.

[1] H. Cho, B. Berger, and J. Peng. Reconstructing causal biological networks through active learning. *PloS One*, 11(3):e0150611, 2016.
[2] A. Foster, M. Jankowiak, E. Bingham, P. Horsfall, Y. W. Teh, T. Rainforth, and N. Goodman. Variational bayesian optimal experimental design. In *NeurIPS*. 2019.
[3] A. Foster, M. Jankowiak, M. O'Meara, Y. W. Teh, and T. Rainforth. A unified stochastic gradient approach to designing bayesian-optimal experiments. In *AISTATS*, volume 108. PMLR, 2020.
[4] S. Jiang, H. Chai, J. Gonzalez, and R. Garnett. Binoculars for efficient, nonmyopic sequential experimental design. *arXiv*, pages arXiv–1909, 2019.
[5] J. Pacheco and J. Fisher. Variational information planning for sequential decision making. In K. Chaudhuri and M. Sugiyama, editors, *PMLR*, volume 89 of *PMLR*. PMLR, 2019.