We thank the reviewers for their thoughtful feedback. We believe that a few misreadings of our work made some of evaluations overly harsh and would ask reviewers to reconsider our paper in the light of clarifications provided below.

R1: We organize the reviewer's questions (Q) and provide answers below. Point #6 clarifies questions in "Correctness".

1. Are graphs necessary? (Q1-2, Q4) The departing point of our work is the realization that an imitating policy is generally underdetermined by the observational data alone. For concreteness, consider models M_1, M_2, unknown to researchers, where in M_1, $X \leftarrow U$, $Y \leftarrow X$; in M_2, $X \leftarrow U$, $Y \leftarrow X \oplus U$; in $M_i, i = 1, 2$, $P(U = 0) = P(U = 1) = 0.5$. We assume that Y, U are unobserved; Y is the reward. In both M_1 and M_2, the observational distribution $P(X = 0) = P(X = 1) = 0.5$. In M_1, $P(y)$ is imitable with policy $\pi(x) = P(x)$; while in M_2, $P(Y = 0|\text{do}(\pi)) = 0.5$ for any $\pi(x)$, which is far from $P(Y = 0) = 1$. This example shows that when unobserved confounders (UCs) are present, imitation learning from observations alone is generally impossible.

2. Learning causal diagrams (Q1-2, Q4) A common approach to address the challenges of UCs is to explore functional relationships among variables, represented as a causal diagram. For instance, M_1, M_2 in the previous example would be distinguished using their corresponding causal diagrams $\mathcal{G}_1, \mathcal{G}_2$ (since bi-directed arrow $X \leftrightarrow Y$ exists in $\mathcal{G}_2$ but not $\mathcal{G}_1$). One could construct causal diagrams with the assistance of domain experts; see examples in (Bottou et al., 2013). In addition, efficient methods for learning causal diagrams from data have been studied under the rubrics of causal discovery (Spirtes, Glymour, and Scheines, 2000), so that dependence on domain experts is minimized. Our methods could be combined with causal discovery algorithms to learn an imitating policy after the graph is obtained. Finally, to address the challenges of incorrect causal diagrams, one could apply causal discovery methods to detect "model misspecifications", i.e., incompatibilities between the diagram and the collected data.

3. Examples, benchmarks (Q3, Q6) Since most RL benchmarks do not explicitly model the presence of UCs, we study the highD dataset which contains trajectories of human driving. We demonstrate how challenges of UCs could arise when applying imitation learning with highD, some of which have been acknowledged in (Etesami and Geiger, 2020). Overall, we agree that imitation learning from observational data contaminated with UCs in sequential decision-making settings is an important and pervasive challenge. This paper takes the first step towards a solution by studying the feasibility of learning an imitating policy from observational data when the causal diagram is obtained.

4. GAIL (Q3, Q6) In all experiments (Sec. 4 and Appendix D), the learner *bc* is able to learn the nominal expert's policy $P(x|pa(\Pi))$, but still diverges significantly in the performance $P(y)$. Since GAIL is not concerned with UCs, it converges to the nominal policy $P(x|pa(\Pi))$, which means that its performance coincides with *bc*. Having said that, our methods could certainly be combined with GAIL to ensure both the causal robustness and the scalability with high-dimensional data, which we'll acknowledge in the paper. In particular, once an i-backdoor/i-strument is found, one could then apply GAIL to obtain a policy that imitate the expert's performance.

5. Related work (Q5) We appreciate the suggested references, but they are somewhat orthogonal to our problem. As we mentioned in Footnote 1, (de Haan, Jayaraman, and Levine, 2019) assume that "input to the expert policy is fully observed", i.e., there is no unobserved confounding. (Arjovsky et al., 2019; Zhang et al., 2020) attempt to find a set of variables so that adjustment on them leads to invariant parameters across multiple environments. Non-parametric methods for such problem have been studied under the rubrics of transportability (Pearl and Bareinboim, 2011).

6. Other comments (1) U is an independent exogenous noise affecting only L, which "are not explicitly shown" in a causal diagram (Line 110-112) (2) By "any policy", we mean a policy taking observed variables (i.e., Z) as input. (3) A causal diagram is imitable where an imitating policy exists regardless of parameters of the observational distribution (e.g., $P(x, w, y)$). Fig. 3(b) is not imitable due to the counterexample in Line 167. However, one could still find an imitating policy in Fig. 3(b) for some *specific* $P(x, w, y)$ (called p-imitable); Line 255-256 shows such an example.

R2: (1) A causal diagram containing latent reward Y generalize the traditional settings of imitation learning. For instance, inverse reinforcement learning requires the parametric form of the reward function, and all input must be observed. (2) While it is not complete, Alg. 1 explores settings where existing imitation methods are not applicable, i.e., when unobserved confounding is present. (2) (Etesami and Geiger, 2020) assumes that reward Y is observed, and the goal is to ensure $P(y) = P(y|\text{do}(\pi))$. This paper assumes Y is unobserved and develop non-trivial methods to find its surrogates (Def. 6). Also, (Etesami and Geiger, 2020) consider a canonical causal diagram and explore parametric assumptions (e.g., linearity); while we focus on non-parametric causal identification methods in an arbitrary graph.

R3: We appreciate your feedback and suggestion. A possible solution to scalability is discussed in #4 to R1. Thanks.

R4: We respectfully disagree with the statement that "practical applications" of our methods "are limited." As we discussed in #2 (to R1), there are causal discovery efficient algorithms to construct causal diagrams from the data; methods could be applied after a graph is obtained. Nevertheless, UCs are still present in the environment despite the hardness of finding its causal diagram, as demonstrated in #1. This paper explicitly acknowledges the existence of UCs and provides methods to circumvent its possibly nefarious effects. As for a discussion on GAIL, please refer to #4.