We would like to thank the reviewers for their comments and constructive feedback. We will implement all the minor comments by the reviewers. Below, we address the main issues raised and clarify some misunderstandings.

R2: *Limitation of intervention not changing parent set:* There are many settings in the empirical sciences where interventions do not change the parents. For instance, a drug injected into a biological system may change the baseline of certain proteins (the underlying causal mechanisms) but it neither entirely suppresses their expression nor create new connections. Also, the work of Yang et al. (2018) characterizes soft interventions in systems without latent variables. Mooij et al. (2013) discussed interventions of this nature in the context of equilibrium in cyclic causal models.

Usage of MAGs: The reviewer's observation only holds for hard interventions. Within the soft intervention setting, Augmented MAGs actually do not lose information with respect to the tests we consider. Fig. 2 is excellent to illustrate this: Since the intervention is soft, the inducing path between X and Y through Z is not eliminated upon X's intervention. Thus, we cannot rule out $X - Y$ adjacency. In general, soft interventions do *not* break inducing paths.

Comparison to nested Markov models: Nested Markov models consider constraints that are computable from the observational distributions, also known as the Verma constraints, as pointed out by the reviewer. We, on the other hand, focus on constraints imposed by a given set of interventional distributions that are not necessarily computable from observational data. Moreover, we provide a graphical characterization to the proposed interventional equivalence class while, to the best of our knowledge, there is no characterization for equivalence in nested Markov models.

Can't we employ rule 9 directly on FCI output?: Consider the diagram in Fig. 2a. By running FCI on observational data, we would get the following PAG $\{X \circ\!\!\rightarrow Y, X \circ\!\!-\!\!\circ Z, X \circ\!\!-\!\!\circ W, Z \circ\!\!\rightarrow Y, W \circ\!\!\rightarrow Y\}$. Our approach recovers more orientations even without Rule 9 as shown in Fig. 2c. Note that F-nodes are not variables, but they are instrumental to recover additional invariant graphical properties such as orientations through our newly proposed machinery.

Maximality of Augmented MAG: The augmented MAG is indeed the maximal ancestral graph constructed from Aug(D). The current definition does not reflect this, as correctly observed by the reviewer and we will update it.

R3: *Controlled experiment assumption:* We apologize for the confusion. Your observation that the assumption is required in the proof of Theorem 4 is correct. Without this assumption, the proposed characterization would still apply under a different F-node construction: Instead of an F-node for every symmetric difference, we would use an F-node for the variables whose mechanism differ across interventions. For instance, if $I = \{A, B, C\}$, $J = \{B, C\}$, $P_I(B|pa_B) \neq P_J(B|pa_B)$, $P_I(C|pa_C) = P_J(C|pa_C)$, then $F_{I \triangle J} \rightarrow \{A, B\}$. We will clarify this with an example.

R4: *Faithfulness issues of [9]:* Note that similar to Yang et al. (2018), we do not treat F-nodes as random variables, but simply devices that are instrumental in graphically representing the results of the invariance tests formalized in the paper. Therefore, faithfulness involving F-nodes does not arise in our approach. Still, a different type of faithfulness assumption is still required to map test results directly back to its graphical representation (Def. 7, pg. 7 lines 285-286).

Advantages over [9]: To the best of our knowledge, Theorem 2 in the paper is the first characterization of an interventional equivalence class under causal insufficiency (presence of latents). As for the algorithms, the FCI-JCI algorithm of [9] is indeed similar to ours, although the resultant graphs are different (they consider regime variables to be random variables connected via bidirected edges). One obvious difference due to our characterization is Rule 9 which can orient additional edges, see Fig. 2 in pg. 8.

Edges between F-nodes and non-intervened variables: The adjacency between F-nodes and nodes outside the corresponding intervention has an important role in our characterization. They signify the existence of an inducing path between the intervention set and the node outside the intervention. For instance, the adjacency of F_X and Y in Fig. 2b encodes the presence of an inducing path between X and Y through Z in Fig. 2a. Meanwhile, we still cannot rule out the presence of a directed edge from X to Y since such a causal diagram would entail the same invariance tests. These edges in the augmented MAG allows us to capture all the test results using a single graph.

Comments on Fig. 1: Our goal with Fig. 1 is to illustrate that, for a ground truth causal diagram with latent variables (Fig. 1a), and for a given set of interventions (p_x, p in Fig. 1b, p_x, p, p_{xz} in Fig. 1c), F-node separation statements can be used to "store", or represent, the results of do-calculus tests. Our goal is not only to learn the MAG, but mainly to establish a graphical equivalence relation that maps directly to the equivalence of the corresponding do-calculus tests.

Soundness of Rule 9: Consider an F-node F_k and $Y \notin S_k$. If there is an inducing path p between them, then every observed non-endpoint Z along the path must be a collider and an ancestor of either endpoint. The crucial point is that Z cannot be an ancestor of F_k since F-nodes are root nodes by construction. Then, Z must be an ancestor of Y. It follows that the subpath of p from Z to Y is an inducing path as well since every observed non-endpoint is a collider and ancestor of Y, i.e., one of the endpoints. Thus, Z and Y are adjacent in the augmented MAG and the edge is directed out of Z, since Z is an ancestor of Y as described. As for Fig. 2c, X could be an ancestor of Y through different ways in the causal diagram as suggested, but this implies that the edge is out of X by the ancestral property of MAGs.