Reviewer 1: Thank you for the review! We agree with all your suggestions. We'll incorporate them. Thanks: thorough!

Reviewer 2: Thank you for the review!

Novelty: To our knowledge, this is the first approach that gives exploitability guarantees while only sampling a small portion of an extensive-form imperfect-information game. This is a highly novel and very important direction (we discuss importance further in the application paragraph below). These kinds of approaches have not been applied to extensive-form imperfect-information games before nor tested experimentally. Also, this approach is very different than the game abstraction literature (discussed in the introduction of the paper).

Runtime guarantee: Obviously, since the algorithm expands at least one node per iteration, the total number of iterations it takes is at most the total number of nodes in the full (underlying) game. You are correct, though, in saying that there is no general runtime guarantee for infinite games, nor one that is a (polynomial) function of the size of the smallest certificate. As we prove in the paper, these are impossible to achieve in the general case. However, we provide experimental evidence that our algorithm finds small certificates in a variety of games. We also want to emphasize that our run time is a function of the size of the certificate, not the size of the entire tree. This is key, and is in sharp contrast to all prior approaches, including CFR, MCCFR, EGT, etc.

Application of small certificates: This we justify in the paper. Despite general hardness, our algorithm allows us to find small certificates in practice of size much smaller than the size of the full game. The natural application (which we discuss in the paper) is one where the agent only has blackbox access to a large game. With the techniques in this paper, strategies with exploitability guarantees can be computed for such games. This is not possible with prior techniques in very large or even infinite games, as we explain in the paper. We demonstrate that even some infinite games have small certificates. It would be impossible to run most known algorithms on such an infinite game, because most of them require either expanding the whole game tree beforehand, or at least (in the case of, *e.g.*, outcome-sampling MCCFR) being able to bound the size of the game tree and conducting more samples than the number of leaves in the tree.

Exploitability: It is well known that exploitability in zero-sum games is bounded by the Nash gap (ε throughout the paper), which we analyze extensively. We will remind the reader about this in the final version.

Typo: Yep, thanks! We'll fix that in the final version.

Reviewer 3: Thank you for the review!

"Text merging" in Prop 6.4: Typo: Prop 6.4 should have the two definitions flipped. We'll correct this in the final version.

Limit Leduc: We'll make this clarification in the final version.

Perfect info in Sec 6: Correct. And the dependence on imperfect information in Thm 6.5 is unavoidable: in perfect-info games, the smallest certificate can be computed from a game tree in linear time via an alpha-beta-like algorithm.

Schmid et al. paper: Thank you for the suggestion. We will try to obtain this paper (it is not currently available via aaai.org or any other site that Google can find) and will read it and cite it in the final version.

Reviewer 4: Thank you for the review!

(3.1) Section 4.2 doesn't assume perfect information. A normal-form game can be converted to a small *imperfect*-information game, as is described in that section. That is all that matters.

(3.2) Thank you for the suggestion. We will incorporate this into the final version.

(3.3) We include this section because it is sometimes unreasonable to assume direct access to the nature action distribution: in many black-box settings, we can only sample the nature distribution.

(3.4) As discussed in the response to Reviewer 3, the hardness in Thm 6.5 comes from the imperfect information, and is somewhat mitigated by the observation that we usually don't care about finding the *smallest* certificate, as long as we can efficiently find one of reasonable size. The hardness in Thm 6.6 is more fundamental: it comes from the fact that we're not assuming access to any reasonable heuristic of where to explore; thus, we may explore the optimal path of play last in the worst case, resulting in a large certificate. We'll include these summary sentences in the final version.

(3.5) Yes. 1) Upper and lower bounds on the value of a node: It is natural to assume trivial bounds (e.g., $[0, 1]$ or $(-\infty, \infty)$) on the utility of all nodes if we don't know any better. Often, we can do better. In our experiments, this was clear in goofspiel. More generally, often rewards are "incremental" (e.g., in a war game if one loses an asset, the value of the asset can be subtracted from the maximum payoff right then); in these cases, the bounds on deeper nodes are often much tighter than the trivial bound. 2) Player action list at player nodes: This seems reasonable. It is impossible to learn to play a game if we don't even know what we're allowed to do. 3) Nature samples at nature nodes: This seems reasonable. This is the minimum amount of access at nature nodes (those that are expanded) necessary for solving.