

1 We would like to sincerely thank each of the reviewers for their time and for their insightful comments.

2 **Reviewer 1** You are right that with any version of PSRO, a potentially exponential number of pure strategies is needed
3 for an extensive form game. However, despite this limitation, PSRO is a very popular and promising algorithm, and
4 successful algorithms such as AlphaStar are based on it. You are correct that we do not actually find an approximate
5 Nash equilibrium in Barrage Stratego. It is our belief that Pipeline PSRO with enough compute would lead to expert-
6 level play in Stratego and Barrage Stratego, although it might still be subtly exploitable due to the reason you mention.
7 Also, as you mentioned, this limitation is due to PSRO itself, and not our improvement of PSRO, which performs better
8 than PSRO, DCH, and Rectified PSRO. We will add further analysis on the exploitability of the algorithm on Barrage.
9 We will include both of your suggestions: training a policy from scratch against the final meta-Nash, and analyzing
10 the final performance of each best response during training. We have analyzed the performance of the best response
11 over time against the meta-Nash in Barrage Stratego and we see that the performance goes down over time, providing
12 evidence that the meta-Nash is becoming less exploitable during training.

13 **Reviewer 2** We did not have sufficient computing power to compare with Deep CFR or NFSP on Barrage Stratego
14 or Leduc. However, we would expect both of these algorithms to outperform any PSRO variant (including Pipeline
15 PSRO) on Leduc poker based on the results reported in these papers. As for Barrage Stratego or similar large games,
16 it is somewhat of an open question whether NFSP or Deep CFR would outperform PSRO-based algorithms. One
17 of the downsides with NFSP is the need to store a large replay buffer of all past experience. Since we are fairly
18 limited with storage, the amount of storage needed to store enough experience to get good results on Barrage Stratego
19 could be too large. Similarly, it is unclear if Deep CFR would be able to get good results on Barrage Stratego. As
20 mentioned in the Deep CFR paper (and in private conversations with the author), the large branching factor could
21 be a problem, but we are not sure if it would or not. After we submitted this paper, DREAM was posted to arxiv:
22 <https://arxiv.org/pdf/2006.10410.pdf> which only samples a single action at each decision point. We think that this
23 method would be more promising for Barrage Stratego than NFSP or Deep CFR. We would be very interested to see
24 results of NFSP, Deep CFR, and DREAM on Barrage Stratego, but these experiments were outside the scope of this
25 work.

26 The theory behind PSRO works in both normal form and extensive form games, but you are correct that in extensive-form
27 games, PSRO could require an exponential number of pure strategies. We will edit the paper to clearly describe this
28 drawback of PSRO. Despite not performing as well as NFSP and Deep CFR on Leduc, PSRO is still a very promising
29 approach for large games. AlphaStar, which was inspired by PSRO, achieved expert-level performance on StarCraft,
30 and it is not clear that a method like NFSP or DREAM would have achieved a similar level. For these reasons, work on
31 PSRO is a promising and important research direction. Existing approaches to parallelizing PSRO are unstable (DCH)
32 or not guaranteed to converge (Rectified PSRO). Among PSRO variants, Pipeline PSRO is clearly state-of-the-art based
33 on theoretical guarantees and empirical performance. Our work is an important contribution to getting PSRO to work at
34 scale.

35 To summarize: PSRO does not perform as well as Deep CFR and NFSP on Leduc poker and could require an exponential
36 number of pure strategies in extensive form games. Despite these drawbacks, PSRO remains a popular and promising
37 algorithm, and it is an open question whether PSRO would perform better than these algorithms on large games.
38 Existing approaches to parallelizing PSRO are unstable or are not guaranteed to converge. Pipeline PSRO is the first
39 approach to parallelize PSRO while maintaining PSRO's convergence guarantees.

40 Other points: We have been in contact with the author of Probe in order to compare with it. During these discussions,
41 we asked if Probe played Barrage, to which we were explicitly told "no". The software description at the link provided
42 by the reviewer appears to conflict with this, however all download links that we could find to this software (including
43 the reviewer's) are dead. We are in the process of resolving this with the author of Probe, as a Barrage comparison
44 should be done if it is possible. All other bots that we found such as Master of the Flag were not able to play Barrage.
45 We will rephrase the abstract to describe P2SRO as "the first scalable PSRO-based method". We will add discussion
46 about how in extensive form games, a potentially exponential number of pure strategies are required to guarantee
47 convergence to a Nash equilibrium. We have added normal form experiments comparing to fictitious play, which gives
48 somewhat of a comparison to NFSP on normal form games and we find that PSRO outperforms fictitious play.

49 **Reviewer 3** You are correct that a limitation to PSRO is that it requires a long best response. Our approach is able to
50 significantly shorten this by training each best response in a hierarchy, but it still has this drawback.

51 **Reviewer 4** You are right that Pipeline PSRO is a simple change to PSRO. We think that this simplicity is a strength
52 because it is very easy to replicate and understand. Furthermore, we show drastically improved performance compared
53 with DCH and Rectified PSRO, which are similar to our algorithm. We will add a citation to lines 80-83 and will clear
54 up the language. We will add more reasoning on why we chose Barrage Stratego. The main reason is that it is a larger
55 game and has many more turns per game than heads-up Texas hold 'em, which only has four turns. We will fix the typo.