Response to all We thank the reviewers for their nice and thorough reviews. All reviewers mention variants of the following question: **“No optimistic OMD results for the Leduc Poker domain”**. We omitted OOMD from Figure 2 (bottom row) because we were not able to trigger the interesting property that OFTRL enjoys: simultaneously having a worse saddle-point gap and a better cumulative regret, as compared to CFR+. In fact, we found OOMD’s performance in Leduc3 to be significantly worse than OFTRL’s, both in terms of saddle-point gap and regret. Understanding these qualitative differences between the two algorithms is an interesting direction of research. We will include this discussion in the final version of the paper.

Response to Reviewer #1

- **“Section 3, line 187”** Yes, it was a typo. Thanks!
- **“Leduc Poker experiments shown in the bottom row, the OOMD and OFTL results are still for the last iterate?”** No, we used the “standard” average $\bar{\boldsymbol{x}} := \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{x}^t$ of all iterates $\boldsymbol{x}^1, \dots, \boldsymbol{x}^T$, in line with the theory (Line 153). Note that the choice of averaging strategy has no effect on the bottom right plot. We will include this discussion in the final version of the paper.
- **“Comparing the Cumulative Regret (and Conclusions)”** Even if our paper mainly uses regret minimizers as a way to compute a saddle point, our regret minimizers can be used in other contexts too. For example, any of our regret minimizers can be used as an online decision maker that plays against an opponent controlled by the environment (see for example (Farina et al., 2019a) cited in our paper, where they use this approach to find exploitative strategies in games). In those settings, regret is the meaningful quality metric of the regret minimizer.
- Fair point regarding both places where we say "prove" but really mean “give relatively conclusive experimental evidence.” We will update both.

Response to Reviewer #2

- **Re “...where is used the flexibility given by the choice of the $\boldsymbol{m}^t$.”** There are at least two cases in which being able to freely choose $\boldsymbol{m}^t$ (rather than setting it to the fixed choice $\boldsymbol{m}^t = \boldsymbol{\ell}^{t-1}$) helps:
  - Beyond saddle-point solving: even if our paper uses regret minimizers as a way to compute a saddle point, these regret minimizers can be used in other contexts too. For example, any of our regret minimizers can be used as an online decision maker that plays against an opponent controlled by the environment; in that case, if a statistical model of the opponent is available, the best prediction of the next loss might very well be different from $\boldsymbol{\ell}^{t-1}$.
  - In other saddle-point algorithms: for example, Farina et al.[1] show how to combine different optimistic regret minimizers and obtain a composite regret minimizer that they use to find a saddle point in two-player zero-sum extensive-form games. They found that being able to pick $\boldsymbol{m}^t$ to something other than $\boldsymbol{\ell}^{t-1}$ is beneficial in their construction (see top left column on page 7, as well as Equation 17 in their paper).
- **Re “...OOMD-type methods do not work in deep games (but those of the OFTRL type do)...”** OOMD and OFTRL both work in deep games, in the sense that they are guaranteed to converge to a saddle point at a rate of $O(T^{-1})$. But experimentally we did find worse performance for OOMD; see the response to all.
- **Re “...decomposition...advantage of such idea in the general case?”** We think the decomposition is interesting for at least two reasons. First, it is the first accelerated method that allows a “CFR-like” interpretation of the method, in the sense that updates are local. From a theoretical perspective we think this is interesting in its own right. Second, it enables a lot of practical experimentation. While this would not technically retain the theoretical rate, it would be interesting to find ways to incorporate ideas such as stepsizes that adapt at a local level, pruning, and other “local” ideas that have been popular in CFR variants.

Response to Reviewer #3

- **Re “... only OFTRL in Leduc?”** See response to all.
- **“not the Entropy DGF used in the empirical results?”** See lines 71-83 in our paper for the reason why we are more interested in the Euclidean DGF than entropy. We have experiments with the entropy DGF as well. We can add them to the appendix for future versions of the paper. They are consistent with the observations for other settings that we mention in lines 71-83; it is worse than Euclidean DGF.
- **Re “Is there no hope that a similar effect would occur in Leduc if run longer?”** It’s possible but we think it’s unlikely. If that were to be the case, we think it would happen at such high precision that it would not be interesting from a practical perspective.
- **Re “... relationship between minimizing regret and exploitability ...”** The connection between regret and saddle-point gap (or exploitability) is one-way: if the two regret minimizers (one per player) have regret R_1 and R_2, then the saddle point gap can be easily shown to be $\leq R_1 + R_2$. However, nothing prevents it from being *much* smaller than $R_1 + R_2$. What we empirically found is that for CFR$^+$ this bound is very loose. We are not sure as to why this is the case, and we agree that it is an interesting fact that deserves more investigation in the future.

[1] Farina, Kroer, Brown and Sandholm. Stable-Predictive Optimistic Counterfactual Regret Minimization. AAAI 2019.