We would first like to thank the reviewers for their insightful comments on our work. We appreciate that you like our paper and that you are helping us making it even stronger with your comments. We will happily incorporate all of your suggestions for improving our paper.

Opinion on Nash equilibrium. Reviewer #1 felt that we were critical of the notion of NE based on line 282. We definitely did not intend to give the impression that we are "unhappy" with the notion of Nash equilibrium: it is a natural and elegant solution concept, and we believe its role as a "gold standard" in game theory is well-deserved. Line 282 concerns the stability of **mixed** Nash equilibria: in contrast to *strict* equilibria, mixed equilibria are fragile because a small "trembling hand" imperfection in the agents' mixed strategies could lead to off-equilibrium behavior. They also effectively presuppose coordination between the agents as, in a mixed equilibrium, any agent could deviate to another mixed strategy with the same support without any penalty; however, doing so would destroy the equilibrium. These shortcomings of mixed equilibria are well documented in the game theory literature. By contrast, *strict* equilibria are uniquely robust in this regard (since they are preserved by small payoff perturbations), and we find the fact that this robustness is picked up by no-regret strategies (like FTRL) quite important and insightful from a learning standpoint. In a certain sense, this provides a crucial link between game-theoretic learning and the extensive equilibrium refinement literature in game theory.

Effects of averaging in general games. Reviewer #1 asked whether averaging can help in general games. There exist small, simple games for which even the time-average of regret-minimizing dynamics does not converge to Nash equilibrium (either in terms of strategies or payoffs) [1, 3]. In fact, [3] shows an example of a class of games with no pure NE such that for all but a zero-measure of initial conditions replicator converges to a limit cycle whose social welfare everywhere (and hence its time average) is optimal and can be arbitrarily better than the best Nash. Thus, time averaging cannot hope to restore any meaningful connection between regret minimization and NE in general games.

Discovering Nash equilibria. Reviewer #1 inquired about promising approaches to finding NE. Optimism can help resolve very special cases of games where the payoff field is monotone without being strictly/strongly monotone (e.g., in bilinear zero-sum games with an interior equilibrium). One can construct game instances where optimistic approaches do not necessarily converge to meaningful equilibria, e.g., [2]. Given that optimism is a much newer technique its failure modes are less well understood, but we suspect that similar to averaging even stronger negative results are possible. Given the PPAD completeness of the problem, there is little hope that a non-exhaustive method can help in the worst case.

Extensive form games. Reviewer #2 wanted to know which of our results carry over to extensive form games. That is an excellent and important question! A key component of our proof is that the FTRL dynamics are divergence free in the payoff space for normal form games. For sequential imperfect information game setting under perfect recall FTRL remains divergence free [5]. It should be straightforward to extend our results for these cases, modulo a heavy notational overhead. Although this direction lies outside our current scope, it should be clear that the insights of our paper are valuable for this line of work and we hope that they will trigger more theoretical/experimental investigations.

Lifting to $(N+1)$-player zero-sum games. Reviewer #3 wants to know if our results can be directly derived from the zero-sum case of [4] by lifting N-player games to $(N+1)$-player zero-sum games. [4] captures only *polymatrix* zero-sum games, i.e., games that can be described by a network of zero-sum *bimatrix* games. General N-player games cannot be lifted to $(N+1)$-player polymatrix zero-sum games. This distinction is important because otherwise all FTRL dynamics on games with an interior Nash equilibrium would have recurrent trajectories based on [4].

References

[1] C. Daskalakis, R. M. Frongillo, C. H. Papadimitriou, G. Pierrakos, and G. Valiant. On learning algorithms for nash equilibria. In *Symposium on Algorithmic Game Theory 2010*, 2010.

[2] C. Daskalakis and I. Panageas. The limit points of (optimistic) gradient descent in min-max optimization. In *NeurIPS 2018*, 2018.

[3] R. D. Kleinberg, K. Ligett, G. Piliouras, and É. Tardos. Beyond the nash equilibrium barrier. In *Innovations in Computer Science 2011*, 2011.

[4] P. Mertikopoulos, C. H. Papadimitriou, and G. Piliouras. Cycles in adversarial regularized learning. In *Symposium on Discrete Algorithms 2018*, 2018.

[5] J. Pérolat, R. Munos, J. Lespiau, S. Omidshafiei, M. Rowland, P. A. Ortega, N. Burch, T. W. Anthony, D. Balduzzi, B. D. Vylder, G. Piliouras, M. Lanctot, and K. Tuyls. From poincaré recurrence to convergence in imperfect information games: Finding equilibrium via regularization. *CoRR*, abs/2002.08456, 2020.