To all: Thank you for exceptionally thoughtful and helpful comments!

Reviewer 2:

- **Further motivation for pseudo-regret:** You are correct that we mostly assumed that the setting (with the pseudo-regret as the performance metric) was already motivated in previous work, and indeed there is room for including more motivation in our introduction—we will accommodate this in our final version. In short, one application domain where corruptions are natural is content/ads recommendation: the presence of malicious users affects the feedback signal received by the learning algorithm, but the objective one cares about is the performance of the system (measured via pseudo regret) on the true population of non-malicious users.
- **Performance of FTL:** This is an excellent comment: FTL is indeed a very natural algorithm in the pure stochastic setting, and it would be interesting to see how it performs in the mildly corrupted case. We will give this some thought for the final version, and at the very least include a comment about it as you suggested.
- **Plots for FTRL:** We have inspected in depth the issue you are pointing out to (we do agree that something appears wrong there), and it turns out that while there is no bug in the experiments, they do illustrate a rather non-intuitive behaviour: recall that there is a trivial upper bound of $O(\Delta T)$ on the pseudo-regret, which kicks in once C becomes very large; the latter bound actually increases with the gap Δ! This explains the artifact you mentioned, which indeed takes place only at high levels of corruption C. At the same time, there is of course no contradiction to our upper bounds. Many thanks for highlighting this—to avoid confusion, we will rework the plots in the more interesting regime where this artifact is negligible (or at the very least carefully discuss this confusing behavior).
- **Minor glitches in proofs:** Thank you for carefully inspecting the proofs and spotting those! We will of course make sure all are corrected for the final version.

Reviewer 3:

- **Corruption in the losses and not only in the feedback:** This is a fantastic point, on which we will remark in the final version: a similar analysis can give for the same MW algorithm an upper bound of order $\sqrt{C/\Delta}$ with respect to the corrupted losses, which is also tight for any value of C. (In this sense, MW enjoys a "best of all worlds" guarantee for any corruption level.)
- **Additive +Δ term on page 13, line 387:** Note that the summation is changed from $p_{t+1,i}$ to be over $p_{t,i}$; to include the last term of the original summation an additive Δ is required. We will elaborate more in the final version.
- **Where $p_{t,i}(\mu_i - \mu_{i^*}) \geq 0$ is being used:** In Eq. (10) we upper bound the summation over $t = t_0 + 1, \ldots, T$ by the summation over $t = 1, \ldots, T$; this holds due to the fact that each term of the summation is non-negative.

Reviewer 4:

- **Practical impact of the result:** Our primary focus in this paper was indeed theoretical, and we do not claim the results to have immediate practical consequences. However, we believe that the broader issue of statistical learning under adversarial corruptions is highly relevant to practice, and that understanding the basic and fundamental questions in this space is crucial before moving on to studying more complex settings.
- **Significance of the technical contribution:** It is true that parts of our development rely on existing techniques in online learning (and we tried to be super transparent about the relationships to those in our writing). Granted, the experts problem is an extremely well studied one and it is always possible to find similarities in the vast literature on the subject. That said, note that our arguments differ from those of [32,33] (for the analogous MAB setting) in a substantial way and rely on somewhat surprising properties of the classic Entropy regularization (e.g., the statement of Lemma 7 is entirely new and was quite illuminating to us). These are crucial for obtaining sharp regret bounds, which are logarithmic in N and independent of T (vs. linear in N, logarithmic in T in the MAB case).
- **Relation to other "beyond worst-case" analyses:** Our discussion of related work in this context focused on prior work within (online) learning. As you correctly remark, going beyond standard worst-case analysis is an active research agenda relevant to many other fields, some of which are surveyed in the pointer you provided. We will do our best to include some more broader context in the final version (but it is hard to do justice to the vast literature on that).

Reviewer 5:

- **Value of simulations and comparison between FTRL and OMD:** We partially agree with your view that these are secondary results, and the main contribution of the paper being the analysis of the (FTRL variant of) MW in the corrupted setting. On the other hand, we also think that the proven gap between FTRL and OMD in this setting is quite surprising given the literature on these meta-algorithms, and the fact that this gap grows with the corruption level is particularly insightful and directly related to the problem at hand. (See also the insightful comments made by Reviewer 2 on this aspect, who actually found this a notable strength of the paper.)