We thank the reviewers for their effort, valuable insights, and comments!

Presentation We will, as suggested by the reviewers, improve the presentation of the paper as we will:

1. Include a concise summary of the related work section in the main body.
2. Further clarify the description of the reduction from fair online batch classification to online batch classification.
3. Attempt to re-arrange and improve the presentation of the result sections (3,4) of the paper.
4. Add a description of the CONTEXT-FTPL algorithm (and also expand a bit about the transductive setting, separator sets), and add a discussion on the motivation behind using it in our setting. We will make sure that no prior knowledge related to CONTEXT-FTPL on the reader's side is assumed or required.
5. Attempt to be more specific when referring to prior work (exact theorems, bounds) throughout the paper.

How results would be different when the auditor still returns the set of all pairs of individuals with fairness violations (R3): Indeed an interesting question. We have thought about it quite a bit while writing the paper - it is still not clear to us how that can be leveraged to improve the overall guarantee. One implication of such a strong requirement could potentially be a faster fairness convergence rate. However, for this to be done, we have to penalize rounds diferentially according to the amount of violations, not just according to the existence of one or more violations, as we suggest by the reduction approach in our paper. This is definitely an interesting avenue for future work, although in terms of practicality, requiring a human auditor to point out all violating pairs might be prohibitive if the number of individuals per round is large.

Dependence of the bounds on $\log(|H|)$, comparison with the result of Gillen et al. [9], which depends on the dimension of the instance space, d (R1): It is important to stress that Gillen et al. [9] operate under a strong set of additional assumptions, in the form of: a) Linear rewards with sub-gaussian noise, b) A metric assumption which must be a Mahalanobis distance function, and c) The assumption that all fairness violations must be detected on every round. It is these assumptions that, in turn, allow them to achieve bounds that depend on the dimension of the instance space, partly due to the fact that algorithms for this problem absent fairness constraints that achieve such a dependence are well-known. They indeed utilize a form of the LinUCB algorithm. Our setting, however, removes all three mentioned assumptions, leaving us in the more difficult, non-parametric case - for which no algorithms with dependence on the dimension of the instances are known. Also, note that $\log(|H|)$ term in our bounds stems directly from the regret guarantee of CONTEXT-FTPL, while any other algorithm in the adversarial setting can be used as a blackbox for our problem; it's just that we don't know of any other algorithm that can achieve better guarantees than CONTEXT-FTPL in terms of the complexity of the hypothesis class without additional assumptions.

In addition, given that we operate in a non-parametric, adversarial setting, we cannot even hope for a mistake bound which depends on the VC-Dimension of H—there exist simple classes H with bounded VC dimensions (e.g., 1-dimensional thresholds) for which sub-linear regret bounds are not possible with adversarial contexts. As an interesting future direction, it would be interesting to see if, when operating in the stochastic arrivals setting (as in section 4), the fair online batch problem can be reduced to the (stochastic) online batch setting. Such a reduction would allow, for example, to incorporate efficient algorithms for the stochastic online batch setting, which would then replace the dependence on $\log(|H|)$ by the VC-dimension of H. One obvious hurdle stems from the fact that even though individual instances arrive stochastically in this setting, the auditor is still allowed to select an arbitrary violating pair on every round adaptively. Replacing the $\log(|H|)$ dependence is therefore non-trivial in our setting, and we consider it a challenging and intriguing question for future work.

The ϵ parameter (R1): This is a slack parameter, representing the sensitivity of the auditor. Due to the nature of the adversary that can choose very similar instances and charge a pair whose fairness violation is infinitesimally bigger than the allowed threshold α, linear fairness regret seems unavoidable without the slack. Thus, in our model, the auditor reports fairness violations of size at least $\alpha + \epsilon$. As shown in the regret guarantee Corollary 3.6 and Theorem 3.8, we characterize the trade-off between the slack allowed and the actual regret for fairness and accuracy.

Practically, to what extent the between-round individual fairness can be achieved in the proposed approaches (R3): We note that enforcing individual fairness across rounds is challenging with existing impossibility results from Gupta and Kamble [4]. Their results show that in the adversarial arrival setting, enforcing individual fairness across rounds would imply linear regret even when the fairness metric is known: linear regret is unavoidable if the learner has to treat even the future instances as similar as the past instances that were misclassified. However, in the stochastic arrivals setting, our fairness generalization result does imply it is possible to achieve approximate individual fairness across rounds.

The rough idea of the composition covering argument could also be discussed in the main body (R3): Given space constraints, we have made an attempt to convey the core idea in the "high-level proof idea" of lemma 4.6.