We thank all reviewers for their valuable and helpful comments. They are addressed in detail in our response below.

Reviewer # 1:

"... it is unclear why the disparity is only measured in one direction (over-emphasis of higher relevance item) with the direction based on relevance ..." We will expand the current explanation (Line 137-144 in Section 2.2) in the final version, further elaborating why we designed the metric to be one-sided. Most importantly, imposing the two-sided proportionality constraint in Singh & Joachims (2018) and Biega et al. (2018) results in an infeasible problem when a small portion of groups/items contribute most of the relevance for the query. Consider the extreme case where only one item is relevant. After putting this item at rank one, we have to put some non-relevant item in position two. This item is now overexposed and violates the two-sided metric, but not the one-sided metric. In this way, the one-sided metric together with utility maximization allows non-relevant items to get higher exposure when this is unavoidable.

"Ideally, fairness critera would be defined in a way that is cognisant of historical/natural directions of bias ..." We appreciate your idea of selecting the direction of the one-sided metric based on historical disadvantage. We will further clarify the generality of our algorithm in the final version — specifically that such constraints, driven by specific application concerns, can easily be trained with the proposed policy-gradient algorithm as well. We focused on the metrics in Equations (3) and (4) for their broad applicability. For example, when ensuring fair exposure of products (items) by different sellers (groups) in a marketplace, there might not be a no notion of historical disadvantage.

"... figure 3 (right) seems to suggest that the post-processing method has a configuration where increasing the NDCG score also decreases the disparity?" There is indeed a tiny increase in NDCG in the second and third point. Note that if the post-processing method had access to the ground-truth relevances, this increase wouldn't happen and both NDCG and fairness would change monotonically with λ. However, the post-processing method has to work with regression estimates of item relevances, that are inaccurate and possibly biased. This leads to the erratic behavior of the post-processing method, and the method actually decreases fairness for much of the range of λ making it impractical, while our policy-gradient method allows a meaningful selection of λ that succeeds in driving unfairness to zero.

"the kinds of experiments presented are all over the place– the yahoo dataset and the german credit dataset show entirely different types of experiments, which makes it difficult to assess". Thank you for the feedback. To clarify, the experiments were designed with three very specific goals in mind:
(Section 4.1) How does our method perform relative to traditional LTR approaches when not using fairness constraints? To be comparable to the state of the art, we used the standard Yahoo! LTR challenge dataset without any modifications. This experiment verifies that our policy-gradient approach achieves reasonable ranking performance.
(Section 4.2) Does the fair ranking model learn to ignore the biased feature? We chose a synthetic dataset that can be visualized to provide direct evidence for our claim that the method can learn to discount biased features. This is a key property that sets it apart from post-processing methods where the learning step is indifferent to fairness, making it impossible for post-processing methods to recover from a biased regression estimates.
(Section 4.3) Can our method effectively enforce both individual and group fairness constraints on real datasets? To evaluate the ability to enforce individual fairness constraints, we again use the Yahoo! LTR dataset for consistency with Section 4.1 and easy reproducability. Unfortunately, the Yahoo! LTR dataset does not provide suitable attributes to define groups for evaluating group fairness. We therefore adapted the German Credit dataset to evaluate the ability to enforce group fairness.
For the final version, we will clarify the layout of the experiments section currently presented in lines 238-243 and re-align the figures with text description.

Reviewer # 3

"Ironically, the motivating example ... is slightly biased itself." Thank you for the thoughtful feedback. It seems that our point about system-endogenous amplification (through disparate exposure) of exogenous biases (e.g. 1% of the employers may be biased against women) is not coming through. We will think about further clarifying, or we will simply replace male/female with neutral group labels.

"... NDCG is typically used as NDCG@k." You are right that NDCG@k is just another set of position weights. Specifically, note that the gradient doesn't vanish even for small k for probabilistic policies, although it may be necessary to use a larger sample of rankings for gradient estimation.

"... minimize w.r.t. λ for a chosen δ." To achieve this behavior, one could simply implement a wrapper that searches for λ given δ. While more efficient solutions might exist, such search is common practice in ML (e.g. regularization parameter).

"... limitation of allowing only differential ML models..." Yes, our method works only for differentiable models, but this is a pretty large class including deep neural networks, SVMs, Lasso, conditional random fields, matrix factorization, etc.