We thank all the reviewers for their incisive comments and pointing out typos, all of which will be fixed in the final draft. In summary, our paper presents a new randomized algorithm for fair $k-$center/supplier clustering, together with theoretical analysis and experimental evidence. Below we address the concerns and questions from each reviewer.

R1. R1 questions the variance in $\mathcal{E}$, and we agree that such concern is on point. Indeed, we observe different variance in additive violations $\mathcal{E}$ on different datasets; nevertheless, $\mathcal{E} \leq 2$ is always obtained within 5 runs irrespective of the dataset. Even in the most extreme case across all experiments and datasets, the maximum additive violation observed was 8. In fact, the variance in $\mathcal{E}$ arises from fractional LP variables, but empirically almost all LP variables were integral, leading to a consistently small variance in $\mathcal{E}$. For a strictly fair comparison, we also note that Bera et al. method can be used in our algorithm to **guarantee** $\mathcal{E} \leq 4\Delta + 3$ using iterative rounding. However, we observe that the expectation guarantee was already sufficient to get few fairness violations. We will add error bars to demonstrate the robustness of $\mathcal{E}$ in the final draft. At issue is also our significantly faster runtime despite the use of binary search (BS), this is primarily attributed to the introduced joiner-signature concept. Our number of variables/constraints of the LP is very small compared to the number of points, leading to an extremely fast feasibility check in BS. Regarding δ in experiment 3, δ is a parameter defined by Bera et al. to control α_i and β_i, where $\beta_i = r_i(1-\delta)$ and $\alpha_i = r_i/(1-\delta)$, and r_i is the ratio of size of group i to the total number of points. Also, ϵ is used in the stopping condition $r - l < \epsilon$ during BS. The additional experiment requested is shown in Figure 2, comparing runtime with the number of protected groups on the full version of bank dataset, which has 7 protected groups in total. The results for the remaining datasets will be in the final draft. Lastly, we confirm that 30 minutes is the actual TLE, also after submission we privatized the GitHub link, which is identical to the supplementary code submitted, to maintain anonymity.

R2. The recommendation of enhancing the notation of the LP will be addressed in the final draft. R2 also suggests reducing overlapping groups to non-overlapping distinct groups. While the preprocessing indeed leads to distinct groups, a fair cluster for the new distinct groups does not necessarily lead to a fair cluster to the original groups. We would be happy to rerun Ahmadian et al. and Bercea et al. algorithms on the datasets with overlapping groups if there is a way. R2 also rightfully mentions that Bercea et al algorithm was not included in the experiments. The main reasons were that the algorithm used $\Theta(N^2)$ variables and constraints in the LP, which was likely to TLE on large datasets (Just like the $\Theta(N^2)$ Algorithm 2 of Ahmadian et al. did), and did not have any practical variant, such as Algorithm 4 in Ahmadian et al. The algorithm also did not have any publicly available implementation that we were aware of and did not support overlapping groups. Nonetheless, we agree and will definitely include the algorithm in the discussion, and will try to implement a practical variant of the Bercea et al. algorithm (Instead of setting $L = C$, set L as $2k$ centers returned by the greedy $k - center$) and report its performance in the final camera-ready version.

R3. R3 questions the $5-$approx ratio for the $k-$supplier variant of the algorithm ($F \neq C$)). Consider an optimal cluster $B(o_i^*, \lambda_\alpha^*)$ and the greedy centers S. While it is true that: $\forall x \in B(o_i^*, \lambda_\alpha^*) \cap C, \exists f \in S$ s.t $d(f, x) \leq 3\lambda_\alpha^*$, this is not sufficient to prove that the LP is feasible. To prove the LP admits a solution, it is sufficient to prove that there is a greedy center $f \in S$ such that $B(o_i^*, \lambda_\alpha^*) \cap C \subset B(f, 5\lambda_\alpha^*)$ (See Figure 1). $4\lambda_\alpha^*$ is not sufficient to guarantee this for all points. In addition, the introduced ideas of joiners and signatures used to parametrize the LP are entirely novel in this context, leading to a natural and more practical algorithm with an intuitive theoretical analysis. Furthermore, our algorithm surpassed the current state of the art both theoretically and practically by several orders of magnitude. R3 also questions if the method can be used for $k - \{means, median\}$. While it is entirely possible that this method can be adapted for those objectives (for example, define a joiner for k-median as area which has “close” nearest-neighbors at the expense of the approximation ratio), the joiner-signature method was specifically designed to parametrize fair $k - \{center, supplier\}$. Finally, line 97 should be "3-approx Algorithm using linear programming and rounding".

R4. R4 has two on point questions regarding the proof of Theorem 1, which we refer to as $Q1, Q2$. For $Q1$, this is indeed as pointed out incorrect, it should be that every point remaining has *some* greedy center within distance $3\lambda_\alpha^*$, however this doesn't affect the correctness of Theorem 1 (using the explanation shown above for R3) and will be fixed in the final draft. For $Q2$, we indeed implicitly used Claim 4 of Bera et al. (Neurips 19) to justify why combining a fair group leads to a fair cluster. The claim will be explicitly cited in the final draft.

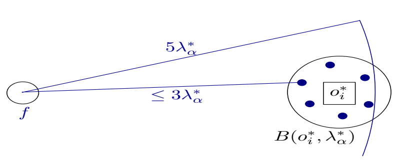

Figure 1: Clarification of Theorem 1. Blue points are points in C

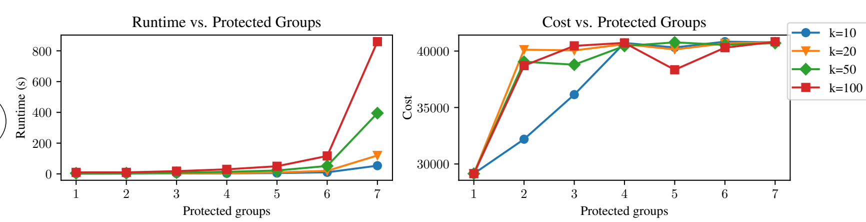

Figure 2: $\epsilon = 0.5, \delta = 0.2$ on bank dataset. Groups 1-6 are either binary (2 categories) or tertiary (3 categories). Group 7 has 11 categories.