We thank the reviewers for their thoughtful comments and diverse perspectives. The suggestions are excellent and will lead to numerous improvements to the paper. We will answer some of the questions directly in the paper, and state others as interesting open problems for future work. See below for our detailed plan.

[R1, Optimal Mechanisms] The reviewer makes a good observation by pointing out that the solution to the linear program does not define full mechanism, but just the probabilities of a mechanism on the 2Δ-lattice. It may have been a slight abuse of terminology to refer to the solution of the linear program as a "mechanism", so we will add in some supporting text to clarify this point.

It is not clear how one could extend the linear programming approach to define a full mechanism, as that would result in an infinite number of variables and constraints. It would be very nice if we could characterize the optimal mechanism using expected error *integrated* over all quality score vectors within some fixed box (the continuous analog of our current result), or even precisely quantify/bound the optimality gap for permute-and-flip under this objective. We don't know the answer to these questions, but will highlight it as an open problem.

[R2, High probability error] We thank the reviewer for encouraging us to think about this question. We acknowledge that the expected additive error is not the only relevant metric, but do point out that the main theorems of MT07 (e.g., Thm 8) do deal with expected error as we've defined it, although agree that DR14 derives a high-probability additive error bound instead.

We feel that expected error is a natural metric to optimize, although recognize that it is still useful to have guarantees w.r.t. the high probability error metric as well. We are unsure in general which of our results (e.g. optimality) may have analogs for high-probability additive error. However, we can easily show that PF inherits the guarantees of the exponential mechanism w.r.t any high probability error metric, much in the same way that it inherits the bound on expected error.

A little more precisely, we can show that PF *stochastically dominates* EM, that is:

$$\Pr[q_{PF} \geq x] \geq \Pr[q_{EM} \geq x]$$

for all x, where q_{PF} and q_{EM} are shorthand notation for q_r where $r \sim M_{PF}(q)$ and $r \sim M_{EM}(q)$ respectively.

This follows directly from the Proof of Lemma 1, where we argue that $\Pr[\mathcal{M}_{PF} = r] - \Pr[\mathcal{M}_{EM} = r]$ monotonically increases with q_r.

Raising this question allowed us to derive a stronger statement about the performance of PF and simplify the existing proof of Theorem 5 at the same time, so thank you R2!

[R2, Universality of EM]

We will clarify the distinction of EM as a universal mechanism vs. EM for selection, as we use here. The reviewer raises an interesting question about non-standard scoring rules, which can hopefully be further explored and answered in future work.

[R2, Expected Runtime]

The reviewer rasises an interesting point on something we did not emphasize in this paper. We had previously thought a little about this idea, and concluded that for it to work, we need to know q_* a-priori or have a sublinear time algorithm to find it. In these special cases, it would be possible to realize computational savings. We will state this as an open problem in the paper.

[R3, Report Noisy Max]

This is a good point. Report noisy max is tricky to reason about exactly because evaluating the probability mass function requires computing an $(n-1)$-dimensional integral in general, as it is the probability that one noisy answer is larger than $n-1$ others.

In the preliminary experiments hinted at in the related work, we studied the expected error of report noisy max in the special case where quality score vectors take on the form $(c, ..., c, 0)$. We will add those details and this experiment into the supplementary material and the full version of this paper, but leave further comparisons to future work.

[R4, Error reduction of other mechanisms]

This is another excellent idea, and something that we feel would greatly improve experiments. It would require additional space, so we state it as an important open problem.