We would like to thank the reviewers for their detailed and helpful feedback. We will address all of the typos and make another editorial pass to incorporate additional reviewer suggestions (such as, adding relevant references etc.) to improve the readability of the paper. We address their concerns below:

1. Comparing experiments with [41,51] and extending beyond Gaussian sketches (R1): Thanks for bringing it up. We emphasize that although we focus on a practice-oriented approach (as compared to recent TCS work), this paper focuses on theoretical aspects of using randomized preconditioning in the context of IPM (whereas the algorithms of [41,51] deal with second-order optimization methods and are of general-purpose), and our experiments are preliminary proof-of-concept showing that the linear solvers in IPM can be accelerated via our technique. For the same reason, we only considered Gaussian sketches in our experiments. We do not anticipate significant differences if other sketching matrices are used and we will extend our experiments to accommodate sparse embeddings as well.

2. Relation to prior works on sketching (R2, R3): Thanks for such insightful comments. We do believe that we highlighted prior works on sketching and preconditioning, especially the ones dealing with solving regression problems or linear systems, but, in light of this suggestion, we will update the 'prior works' section in the final version of the paper. As R3 further suggested, instead of just saying RLA is used, we will state earlier in the paper that the construction of our preconditioner depends on the known randomized linear system solving tools from subspace embeddings. In this context, we will also include the references pointed out by R2 and R3, as well as other relevant papers that we might have missed out.

3. Extending our experiments (R2, R3): In addition to R1's suggestions on extending our empirical evaluations, we can also include additional experiments comparing our algorithm with other standard preconditioning techniques applied to IPMs for solving linear programs, specifically, as suggested by R3, the inner-iteration preconditioning of [13]. As a supplement to our main numerical results and as suggested by R3, we will also include in our experiments how the convergence gets affected in practice if CG is replaced with SD (or Richardson) in Algorithm 1. We agree with R2 and R3 on the fact that additional experiments will further validate the efficiency of our proposed approach, but we do believe that our experimental results are already a strong proof-of-concept that our proposed theory works well in practice.

4. Outlining the novelty (R3): Thanks for the detailed feedback. In the final version, as R3 suggested, we will attempt to further emphasize the novelty, including the way we handled the error incurred by the iterative solver by proposing a fast, sketching-based solution to a linear invariant that needed to be exactly satisfied. We do note that the construction of our solution is original *and* computationally efficient. We will also revisit and perhaps improve the description of our contributions with respect to the prior works on inexact IPMs for LP.

5. Why is [14] closer to this paper? (R3): In our opinion, this is due to the fact that their objective is quite similar: the analysis of an approximate solver in each iteration. However, important differences do exist between their work and ours, as mentioned in our work.

6. SVD of $\mathbf{ADW}$ could cause large runtime issues (R3): In the analysis, we use an SVD to provide a cleaner and simpler theoretical analysis and since it does not increase the asymptotic complexity. We can replace the use of SVD with a Cholesky factorization and significantly improve the constants.

7. Reporting running times (R3, R4): We refrained from reporting running times to avoid direct comparisons with heavily optimized benchmark LP solvers; in industrial-grade solvers the "true" algorithmic efficiency is grossly confounded by built-in optimization strategies. For better comparisons with other standard preconditioners (please see item 3) as well as the direct solvers, we will include the running times of our experiments in the updated version.

8. Outlining the novelty (R4): As also recommended by R2 and R3 (please see item 2), we will update the "prior work" to further emphasize the line of research on sketching and preconditioning. In addition, as per the suggestions by R3 , we will further elaborate on the novelty of our paper with respect to handling the error due to the approximate solver (please see item 4).

9. Singular values of ADW (R4): Thanks for pointing out the typo. The matrix of the singular values of $\mathbf{ADW}$ is indeed $\mathbf{\Sigma}_{\mathbf{Q}}^{1/2}$, not $\mathbf{\Sigma}_{\mathbf{Q}}^{-1/2}$. We will fix it in the revised version.

10. Construction of W (R4): We will add a short note on how to construct $\mathbf{W}$.

11. Citing prior works on the oscillatory behavior of CG residual (R4): By "...the norms of the CG residuals may oscillate", we implied the non-monotonic decrease of the residual norms in CG. For example, in Theorem 8 of [5], if we have $\frac{\kappa(\mathbf{A})-1}{2} > 1$ *i.e.* if $\|r_k\| \leq \beta \|r_{k-1}\|$ for $\beta > 1$, it doesn't guarantee $\|r_k\| \leq \|r_{k-1}\|$. In practice, similar behavior of CG residuals also discussed in many papers including: *CG vs. MINRES: An Empirical Comparison* by Fong and Saunders (2012). We will add this reference in the updated version.