We would like to thank the reviewers for their time, constructive feedback, and critical suggestions that will help us improve the paper. As noted in the reviews, the final submission of the paper needs additional refinements, especially in the proof sections presented in the supplementary material. Before providing our individual reponses, we would like to bring two important points to the reviewers' attentions. First, the theoretical results provided in this paper have many potential applications. Due to space limitations, we confined ourselves to only a few applications in learning from quadratic measurements, i.e., sparse covariance estimation, low-rank matrix recovery, and PhaseLift in phase retrieval. Exploring all potential applications of our main result (Theorem 1) is beyond the scope of a single conference paper. As a result, we anticipate that our theoretical findings may be helpful to many researchers and, in particular, to the NeurIPS audience. Second, we must emphasize that, unlike most previous works, in this paper we do not assume the independence of the entries of the measurement vectors. Indeed, to the best of our knowledge, the precise phase transition for measurements vectors with non i.i.d. entries has been provided for the first time in this paper. Please find our individual responses to each of the reviews in their respective threads.

Reviewer1: This paper provides a generalized framework from which the precise performance of signal reconstruction using linear measurements with iid vectors can be characterized. The importance of the result is that it extends the performance analysis beyond the iid Gaussian setting (which has been studied extensively in the literature). We thank the reviewer for the detailed comments and suggested improvements. Nonetheless, we must point out that previous results for non-Gaussian random matrices (e.g., binary or Rademacher) either have assumed that the entries of the measurement vectors are independent (and demonstrate a performance identical to the iid Gaussian case), or only provide orderwise bounds on the phase transition (such as those that have been given for PhaseLift). Our main result (Theorem 1) precisely characterizes the performance without assuming independence of the entries (which is why we were able to resolve the long-standing problem of determining the phase transition of PhaseLift). However, to address the reviwer's point, in the final submission we will add a discussion to compare with previous results for non-Gaussian random matrices. The statement in line 51-55 is informal and was meant to provide the gist of the main contribution. To add clarity, we will edit the text and more formally illustrate the result.

We will add some references to specific parts in the corresponding papers (e.g. Section V.C in [29]) to address the connections of our corollaries with the previous results exploiting CGMT. The remark on the "Pseudo Gaussian Width" will be updated to avoid ambiguities. The expectation defined in (5), is a generalization of the Gaussian width that is known in the literature. More specifically, as noted in the paper, it reduces to the Gausssian width in the iid Gaussian setting, i.e., when $\mu = \mathbf{0}$, and $\mathbf{\Sigma} = \mathbf{I}_n$. In line 247, the matrix $\mathbf{M}$ is only used for generating a PSD covariance matrix through $\mathbf{\Sigma} = \mathbf{M}\mathbf{M}^T$, where $\mathbf{\Sigma}$ is used later on in our numerical simulations. Therefore, normalization is not needed. The reviewer is correct that the result in Corollary 3 is assymptotic, based on which we should be more careful in our statement in line 307 (we will have perfect recovery if $m/n > 3r$). It is worth noting that even though the theoretical results are asymptotic, we observed in our numerical simulations that when n and m are moderately large, the theory well matches the empirical results.

We appreciate the typographical comments from the reviewer. We also apologize for the broken references in the supplementary material. They will all be taken care of in the final submission of the paper.

Reviewer2: As correctly pointed out by the reviewer, this paper generalizes results in linear inverse problems to beyond iid Gaussian settings. The main theoretical result (in Theorem 1) demonstrates the equivalence of the phase transition in the Gaussian and non-Gaussian settings provided the first and second order statistics match. We provide detailed specialization of the result for sparse and low-rank matrix recovery, and PhaseLift in phase retrieval. We are glad to see the reviewer finds the results interesting and widely applicable. We also thank the reviewer for the suggested improvements. We will provide some clarification for the definition of "pseudo Gaussian width". We will also add a section to discuss potential further research directions and suggest some other cases where our theory can be applied.

Reviewer3: As stated by the reviewer, this paper extends earlier results on the performance of convex-relaxation-based methods to a broader class of linear measurements. We appreciate the efforts made by the reviewer to provide us with the insightful historical context. The comments are definitely valid and will be addressed in the final submission of the paper. We are planning to extend the introduction section by (a) adding a detailed discussion on the foundational works that analyzed the performance of convex relaxations for solving the linear inverse problem, and (b) a comparison with other universality results which have especially gained attention in the community in the past few years.

Theorem 1 was developed as an effort to resolve some of the current trending topics in phase retrieval, learning from quadratic measurements, interference analysis in MIMO communications, etc., for which the classical results could not provide a precise analysis. Regarding the comments on alternative non-convex methods, as the reviewer points out, the theoretical understanding is very limited. As such, it is not straightforward—and certainly beyond the scope of this paper—to compare our theoretical findings with the plethora of available non-convex methods. Nonetheless, we are currently investigating whether (under certain conditions) the phase transition of the optimization program (3) with non-convex objective $f(\cdot)$, would remain the same if we replace the rows of the measurement matrix, $\mathbf{A}$, with vectors drawn from Gaussian distribution with the same first and second moments. Finally, at this point we feel that the abstract of the paper sufficiently describes the convex limitation of our results. However, if the meta-reviewer feels that this restriction should also be reflected in the title, we will happily do so.