We thank all the reviewers for their valuable feedback. Regarding implementation and practicality, this work primarily aimed to address outstanding theoretical questions in the robust statistics and continuous optimization literatures on several fronts. The question of practicality of more standard filtering-based methods (which our first PCA algorithm is an example of), in terms of dimension-dependent bounds on the number of iterations, led us to the development of our second robust PCA algorithm (Algorithm 3). As each of its key subroutines is based entirely on matrix-vector multiplications, and it runs in nearly-linear time for covariance matrices with a mild spectral gap, our hope is that it can indeed find use in practice. We agree it is interesting and important future work to evaluate our methods empirically, but believe it is out of the scope of the current paper as its focus was establishing various theoretical primitives.

Reviewer 1. Regarding the relationship between the two problems studied: we give two algorithms for the robust PCA problem; one is developed in Section 3 (Theorem 1), and the other in Section 5 (Theorem 2). The algorithm in Section 5 uses our Schatten packing procedure in Section 4 as a crucial subroutine (see Line 2 of Algorithm 3, and Proposition 2); its development appears to us to be necessary compared to standard operator norm packing procedures, as one needs to reason about the amount of adversarial perturbation along multiple directions. We will make this more clear in the revision. The power method takes $\tilde{O}(d^2)$ with high probability as it can be implemented with logarithmically many matrix-vector multiplies. We agree with your remaining comments and thank you for your suggestions.

Reviewer 2. Thank you for your reading. We believe our work addresses an important outstanding question in the robust statistics literature by proposing an efficient algorithm, develops a useful optimization primitive, and makes progress towards the broader goal of covariance estimation under weaker assumptions than is known. We hope you are inclined to raise your opinion of our paper.

Reviewer 3. Since our algorithm is conceptually simple and its implementation only requires (parallelizable) matrix-vector multiplies, we are optimistic it (or variants of it) can find use in practice, but acknowledge your concerns. We agree that the algorithm and analysis structure of Section 3 more heavily uses known techniques in the literature (i.e. filtering via downweighting), but will clarify this more directly in the text. On the other hand, our use of a Schatten packing subroutine in Algorithm 3 (as well as using its guarantees in the analysis) is novel, and we hope it finds additional use in robust statistics tasks where operator norm conditions do not suffice, and also as an independently interesting optimization primitive. We appreciate that you suggest this as well; thank you for your detailed comments.

Reviewer 4. We will add additional expositional text discussing the utility of a Schatten packing subroutine for the robust PCA task; at a high level, an operator norm packing algorithm cannot detect an adversary which raises a different eigendirection to have a quadratic form equal to the top, which can fool a PCA algorithm. Regarding space complexity: none of our algorithms require more than storing a constant number of additional vectors, i.e. $O(n)$ additive memory overhead, at any point; we will make a note of this. This space complexity overhead is standard for approximate SDP algorithms. The $\circ$ notation means entrywise product; we will clarify this. We agree with your other comments and will revise appropriately; thank you for your careful reading.