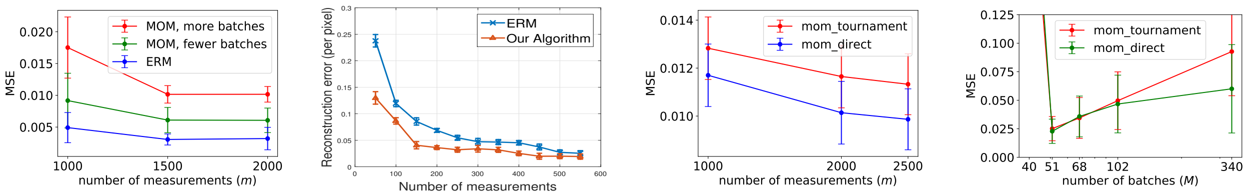

(a) Response to R3 | (b) R4, Weakness 1 | (c) R4, Weakness 2 | (d) R4, Weakness 4

Figure 1: (a) MSE vs number of measurements when the measurements are *sub-Gaussian without corruptions*. (b) Plot adds confidence intervals for MSE in the heavy tailed setting *without corruptions*(Figure 1(b) in the original paper). (c) Aggregate statistics for MOM in the presence of corruptions. (d) MSE vs number of batches for MOM on 1000 heavy-tailed measurements and 20 corruptions. All error bars indicate 95% confidence intervals. Plots (a),(c),(d) use a PGGAN on CelebA-HQ while (b) uses a DCGAN on MNIST.

We thank the reviewers for their positive comments and useful suggestions. We are delighted that the reviewers found: our problem well formulated & relevant; our theoretical results meaningful & novel; our experiments impressive & practical. As noted (R2, R4), robust compressed sensing is an important and practically relevant problem. The proposed Median-of-Means (MOM) algorithm advances the understanding of robust CS using generative models, in the setting where measurements and measurement matrix are heavy-tailed and contain adversarial corruptions. We will modify the notation and definition according to suggestions by R1, R2, R3, R4. We now address individual reviewer concerns.

@R3: Thank you for the positive feedback, we genuinely appreciate it. **"Are there cases where ERM outperforms MOM?"** Yes – it is known that ERM is optimal when the measurements are subgaussian, with no adversarial corruptions. We took this as a primitive and did not evaluate it in our submission. In Figure 1a, we plot MSE vs number of measurements under *sub-Gaussian measurements without corruptions*. We believe the gap between ERM and MOM is because the proposed objectives are harder to optimize, and can be reduced via fancier optimization routines that use negative momentum. We will include this in future versions and conduct more extensive experiments to see if MOM objectives are fundamentally harder to optimize. Note that as we reduce number of batches, MOM approaches ERM.

@R4: Thanks for the positive feedback and thought provoking questions. **Weakness 1: "Figure 1 should indicate standard deviation"**. Please see Fig. 1b, which shows that our results are statistically significant. **Weakness 2: "Does MOM consistently achieve good reconstruction? Does direct MOM minimization do just as well as the MOM tournament in practice?"** Yes, they are consistent, and we find no statistically significant difference between them. In Fig 1c we show MSE as the number of measurements are varied and the fraction of outliers is set to 0.02. **Weakness 3: "Is it possible to achieve even better reconstruction by further fine tuning the hyperparameters of trimmed loss minimization?"** We do not know – the difficulty with trimmed loss minimization is that we do not know how to cross-validate hyperparameters. We were conservative and use 80% of the samples at each gradient step. **Weakness 4: "How sensitive is MOM to the selection of the number of batches M?"** Excellent question, thank you for the suggestion! In Fig 1d, we plot the MSE vs number of batches (M), under the setting of Section 5, Figure 3 in the main paper (1000 heavy-tailed measurements + 20 corruptions). We find that the method fails when M is too small, since the majority of batches are corrupted. As long as M is above a certain threshold, MSE increases slowly. In the main paper, we used a conservative value of $M = 340$, as the qualitative results do not change much with M. Note that there exists a cross-validation scheme for determining the optimal batch size, which we will include in the main paper due to space constraints in this rebuttal.

@R2: Thanks for the positive feedback. We agree with your valuable suggestions and opinion that generative models have limitations compared to classical approaches. However, given their constant progress, they have good potential to be useful for real problems. On Question (b): you are correct, we will include a description for L1 minimization.

@R1: Thanks for the positive feedback. **"The problem does not seem to be commonly encountered [..] the experimental results appear to solve several toy problems."** We respectfully disagree. We allow for outliers in the *measurement matrix and measurement vector.* Robustness to measurement corruptions and outliers is fundamental for compressed sensing theory (Comments from R2, R4 also support that). Practically, it can appear for mis-calibrated medical imaging, malicious or strategic observations or other cases where robust statistics is relevant. @R1: **"The paper focuses on ERM as a baseline despite other approaches for CS with generative models."** Indeed several approaches exist for linear compressed sensing with generative models, but most of them suggest different algorithms for approximately solving ERM (which is a non-convex and hard optimization problem). E.g., ADMM and PGD give improvements over Bora et al. These algorithms will fail as they rely on ERM and do not consider robustness. The only work on robust CS with generative models is by Wei et al. which does not propose a practical algorithm and shows no experimental evaluation. Our work is the first practical algorithm for robust CS with generative models. @R1: **"Statements (Ln. 154, 197, etc). are imprecise"** Indeed, we only want to build intution- we will try to improve the presentation. Precise statements are in Lemma 4.3 & 4.4, and Theorem 4.5.