First, we would like to thank the reviewers for the careful reading of the paper and for helpful/thoughtful remarks. All the recommendations made by the reviewers will be taken into account in the revised version.

Response to Reviewer # 1 Thank you very much for your very enthusiastic report. Let us briefly comment on the question related to the choice of the parameter \( \lambda_0 \), and on its impact on the quality of estimation. The basic intuition included in your report is correct: the fact that \( \lambda_0 \sqrt{n} \) tends to infinity when \( n \to \infty \) implies that all the outliers for which the outlyingness \( y_i - X_i^\top \beta^* - \xi_i \) is smaller than \( c \sigma \sqrt{\log n} \) will be concerned only by the squared loss. As a consequence, if we were aware that all the outliers satisfy the condition \( |y_i - X_i^\top \beta^* - \xi_i| \leq c \sigma \sqrt{\log n} \) for some \( c > 0 \), there would be no need of using the Huber loss; the standard (penalized) least-squares estimator would have the statistical precision described in Theorem 3. However, one can never really check whether this condition is satisfied or not, since, for instance, \( \beta^* \) is unknown.

Response to Reviewer # 2 Thank you very much for your very positive and encouraging report. We agree with all the remarks/comments/suggestions you made. Concerning your remark (6), the trace should be removed. In an initial version of the paper, Lemma 3 was stated in a more general case in which \( b \) was a matrix. This is why the trace operator was used. We apologize for this typo.

Response to Reviewer # 3 Thank you very much for your very positive and encouraging report. We agree with all the remarks/comments/suggestions you made. Concerning your remark (6), the trace should be removed. In an initial version of the paper, Lemma 3 was stated in a more general case in which \( b \) was a matrix. This is why the trace operator was used. We apologize for this typo.