Author Responses for “Learning Erdős-Rényi Random Graphs via Edge Detecting Queries”

We are very pleased to have received these positive comments on our paper, and we are grateful to all of the reviewers for their feedback and suggestions.

Response to Reviewer 3: By freeing up some space as suggested, as well as making use of the 9th page allowed in the camera-ready version, we can include the details of the sublinear-time decoding algorithm in the main body. We will also elaborate on as many proof outlines as possible.

Response to Reviewer 4: We are grateful for the feedback, but are admittedly a little surprised at the final score entered, given the reviewer’s positive comments on all four of Originality, Quality, Clarity, and Significance.

Regarding the minor clarity issues, we will adjust Figure 1 according to these suggestions and fix the typos stated.

If we understand correctly, the reviewer’s main concerns are that the numerical results are not comprehensive. We appreciate this type of concern in general, but we would like to emphasize that this is a theory paper, and that the experiments are only meant to serve as a simple proof-of-concept rather than being central to the paper. (This is supported by Reviewer 3’s suggestion to in fact make the experiment section *more compact*.) Please also note that the majority of related works on this problem and group testing (see our References section) do not include experiments. In light of all this, we hope that the final decision is made based on the theory.

We compared COMP/DD/SSS/LP experimentally because these all use the same test matrix (i.i.d. Bernoulli) so give more directly comparable results. GROTESQUE is not meant to compete with these in terms of the number of tests, as a factor $O(\log \overline{k} \cdot \log n)$ increase is typically significant in practice. (We would be happy to highlight this further in Section 6). Again, we believe that adopting the current format is appropriate for a theory paper, and that improved sublinear-time decoding algorithms (e.g., optimal scaling laws, improved constants, and/or competitive empirical performance) are better left for future work, especially given that the number of results and their level of detail is already quite high for a conference paper.

Comparing the empirical number of tests to the theoretical thresholds is a good idea, though in accordance with Reviewer 3’s suggestions, we believe it would belong in the supplementary material and not the main body. The following figure illustrates an example for COMP (green / right), DD (blue / middle), and LP (red / left):

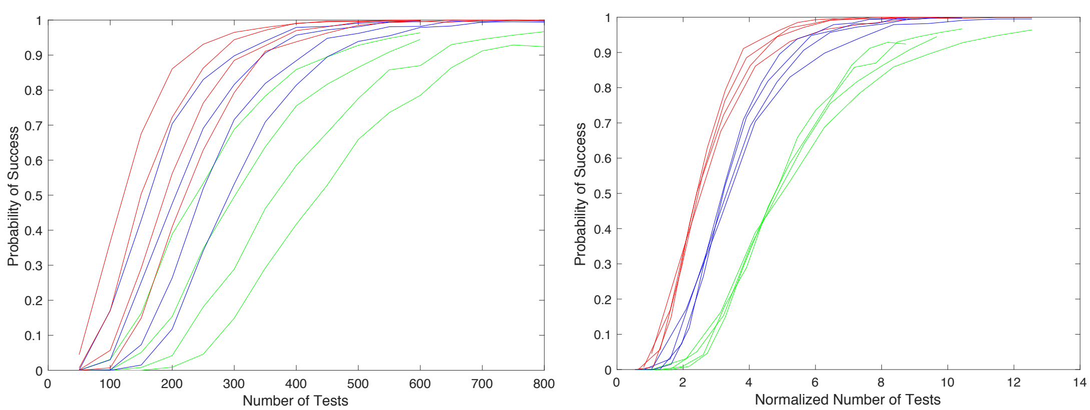

Figure 1: (Left) Number of tests for four different $(n, \overline{k})$ pairs described below; (Right) Normalized number of tests after division by $\overline{k} \log \frac{1}{q}$, where q is the probability of each edge in the graph.

Here, we have run the algorithms with $n \in \{80, 100, 120, 140\}$ and $\overline{k} = \frac{n}{10}$, and normalized the horizontal axis of the second plot to $\#\text{tests}/(\overline{k} \log \frac{1}{q})$. As predicted by our theory, the resulting 4 curves for each algorithm nearly overlap, with slight deviations due to noise and a sharper threshold as n increases. Moreover, the theory (see Figure 1 of the paper with a sparsity level close to zero) suggests an asymptotic threshold of roughly 1 for the optimal algorithm (which LP approximates), roughly 2 for DD, and slightly over 2 for COMP. The above figure is consistent with these numbers, though they are slightly increased because of the penalty incurred for finite n (as opposed to $n \to \infty$).