cannot just blindly apply them when we come to another graph. If the dK sequence is so important, why generate alternative graphs at all: just give people the dK sequence and let them work with it.

## Reviewer #5

**Strengths:** - Addresses an important and difficult problem with some novel angle and non-trivial observation.
- combines design with concrete validation that provides more ground (even if it does not close the topic).

**Weaknesses:** - The reader should be ready to accept semi-proof and lack of complete rigorous formal model and proof, for the sake of exposition of interesting observations.
- Limitations are severe (only releasing k=2 reduces the k series to its single most expression: assortativity). The extension is conceptually interesting but far from being as attainable as the authors say.
- The levels of privacy proposed are not ready for prime time. Most \epsilon-differential privacy work recommends \epsilon at most 0.1 this paper considers the parameter between 5 and 100!

**Comments to Authors:** I think this is an interesting candidate for publication. The use of dk series seems promising to address the need to manipulate realistic graphs, this paper is making one step for it. This has potential to impact an important problem.

We are far from a contribution solidly written in the marble given the choices made by the authors in terms of high \epsilon and a rather poor exposition of the methods, which seem to follow more high-level intuition than precise first principles. Nevertheless the point that partitioning improves sensitivity and brings guarantee is a good catch, and the paper makes the case serious.

p.1 “We take a different approach to address the above question, by making the observation” the current claim of ownership of this observation is strange. This tension is the very much at the core of much research including k-anonymity, differential privacy.

The limitation of k=2 might appear a bit hard to swallow after such a grand claim. I still think it is an interesting step because that is the way to go. Who knows how sensitivity behave for k=3 and if any partitioning make sense.

“In contrast, our goal is to inject changes into all aspects of the graph topology, instead of focusing on a single graph metric.” You do not reproduce all aspects, and you will likely never. You propose to recover what the dk series can obtain, starting with k=2. It does already make your approach original (and promising) but stating it so broadly is a countersense.

“Unfortunately, the author asserts there are incorrect results in the paper 1.” This is perhaps unfortunate but it does not explain how your method is intrinsically better. It would much stronger to highlight the difference first and then mention this point.

Lemma 1 is a partial result. You only provide an upper bound, which does not prove that the real sensitivity is necessarily high.

The statement of error measure is very vague. How does random noise alter the actual structure of the graph?

Please clarify what happens between clusters? Are the data lost with some forms of random generation of links between them?

Have you ever seen a single paper advocating \epsilon between 5 and 100? You are essentially saying the users “Do not worry, your chance of being identified by joining the database are only multiplied between 148 and about 10^43”. What kind of guarantee is that?

## Response from the Authors

We thank the reviewers for their insightful comments. Several comments were results of ambiguous text in the paper, which we have addressed by clarifying our claims and assumptions and providing deeper explanations of our findings. In particular, we explain that the omission of the dK-PA was simply because it generated so much noise that the dK-generator failed to generate matching graphs. Two additional key points stood out in the comments, and we address them in detail below.

First, on the issue of dK-2 as a graph statistical representation, we modified text to more clearly explain the advantages and the limitations of our choice. We explain that we require a statistical representation of a graph that can be converted to and from an unique graph. The dK-series is ideal for this. We use the dK-2 series, because it is the most detailed dK-series that has a corresponding graph generator (e.g. there is currently no known dK-3 series graph generator that works on large graphs). While the choice of dK-2 limits the accuracy of our current model, our methodology is general, and can be used with higher order dK-series when their generators are discovered (e.g. we are currently working on developing a scalable dK-3 generator). It is possible that providing privacy on higher order dK-series may require more severe noise, which could consequently destroy their higher accuracy. Therefore, our conclusion is that higher order dK-series will become a practical solution only if we are able to preserve their accuracy through the perturbation process and when a generator will be invented.

Second, we address via text edits questions on the choice of \epsilon: smaller \epsilon indicates stronger privacy. We use moderate to high values of \epsilon in our tests for two reasons. One, we wanted to find the \epsilon value that contributes to the smallest noise such that it produces a graph statistically similar to the synthetic dK-2 graph with no privacy. Thus we can indirectly quantify the level of privacy inherent in a synthetic graph without additional privacy constraints. We show that this property is achieved when \epsilon is equal to $100$. In addition, the dK-2 series is a very sensitive function and naturally requires high level of noise to guarantee strong privacy. Our primary goal was to identify the feasibility of this approach, and leave further optimizations to achieve high fidelity graphs for lower \epsilon values as goals for future work.