Response from Authors to Reviewer Comments

We would like to thank the reviewer for his comments and suggestions and the editor for giving us an opportunity to revise our manuscript. We have revised the manuscript and we believe this revised version addresses all the comments made by the reviewer.

We follow the following convention in this document:

- Review comments: black italicized font.
- Response to review comments: black regular font.
- Modifications made in the paper: blue regular font.

Comments from Reviewer 1:

1. *A few minor items to correct: inconsistent spacing of colons in the Sec. 3 definitions, 'etc' without a period, "corresponding to it's influencer type" on pg 6, and a few other minor items that a careful reading will highlight.*

   The indicated inconsistency and minor grammatical errors have been corrected in the text.

2. *It's not clear why the GPlus experiment is sectioned into an appendix yet still included in the main body of the paper.*

   The authors believe that it is the style of the IEEE Transactions to position the appendix before the references but are willing to change this if the editor so requires.

3. *In the first paragraph of Sec 3, "...and interpreted as the probability of agent y persuading agent x...", I believe x and y should be switched, based on the rest of the paragraph.*

   The authors thank the reviewer for the correction, which has been made in the text as follows:

   <Page 3, Section 3, Paragraph 1 (left column), Line 12>This influence of x over y is represented by the edge-weight $w_{x,y}$ and interpreted as the probability of agent x persuading agent y to purchase a product.

4. *Sec 3a: It is non-standard to represent variance as \sigma instead of \sigma^2.*

   The text has been modified to represent variance as \sigma^2.

   <Page 3, Section 3A, Line 16>The overall interest of the population in a product is modeled as a Gaussian distribution with a certain mean (μ) and variance (σ^2).

   <Page 3, Section 3A, Line 19>A marketer might not know the interest of every agent, but have a general idea of the overall customers' interest of the population and can set the initial mean (μ) and variance (σ2) of the customers' interest prior to the simulation.

5. *zeta_x in Eq 5 is not explicitly defined.*

   The definition of zeta_x has been included:

   <Page 4, Section 3C.1, Line 13>Where the event of an agent *x* purchasing the product is denoted by ζ_x.

6. *In Sec 3c.1, you write that individuals not interested in purchasing a product do not participate in marketing. But in the paragraph containing Eq 9, you model non-purchasers as participating with low probability. Why the contradiction?*

   The authors would like to clarify that Eq9 incorporates the assumption that individuals not interested in a product are unlikely to participate in marketing [13]. Thus, the non-purchasers are modelled with a low probability. The following modification has been made in the text:

   <Page 4, Section 3C.1, Line 3>Purchasing a product is considered as a means of participating in the advertising campaign and deciding not to purchase the product means that the agent is unlikely to participate further in the campaign, since in reality individuals do not generally share, retweet, or forward product-related content if they are uninterested [13].

7. *Conversion ratio (Eq 13) is described as something we would want to know for an individual influencer, but it appears to be calculated as a single value for the system, i.e., all influencers get equal credit for converting customers. Is this interpretation accurate, and if so, is it desirable?*

   While the conversion ratio can be computed for an individual influencer too, this work focuses on modeling the group dynamics of influencers. As stated in the paper **<Page 7, Section 4.A, right column, line 11>**: *"...brand marketers generally hire multiple influencers across a social network.".* The authors focus on studying the dynamics of the influencers and their performance as a group.

The following clarification has been made in the text:

<Page 7, Section 4, left column, Line 21>The focus being on a set of influencers, the conversion ratio and the customer acquisition cost are calculated for a set of influencers of a particular kind instead of an individual influencer. This enables comparison among groups of different kinds of influencers.

8. *Using normal distribution to sample some parameters can yield values outside [0, 1]. I assume you use a truncated normal. Is that right?*

   Yes, a truncated normal is used to obtain values in the interval [0, 1]. The following modification has been made in the text:

   <Page 5, Section 3D.1, Line 3>The interest λ_x of an agent is initialized by sampling from the truncated normal distribution according to the input mean μ and variance σ^2 parameters, i.e., $\lambda_x \sim N(\mu,\sigma^2)$ in order to obtain values in the interval [0, 1].

9. *Why were sigma, a, gamma, and c kept constant, and why the chosen values?*

   The following clarification has been added to the text to indicate the sources that were used to determine parameters values for simulation:

   <Page 7, Section 4, Left Column, Line 31>The above values were chosen for these parameters in order to simulate a general circumstance. $\sigma = 0.2$ was chosen in order to not let a large number of samples cluster around the mean, i.e., to incorporate diversity in consumer interest. $a = 0.9$ ensured that around 90% of the network is active [28], and the gamma of 0.01 was chosen to maintain a significantly lower spread of negative influence [13]. Assuming influencer interaction with followers generates a large influence impact on the follower [26], we use the influence update $c = 0.7$.

10. *Figs. 3 and 4 have a categorical horizontal axis, so using connecting lines between data points might be a little confusing. It might be worth a note in the figure caption so the reader doesn't try to find meaning in the slope of the lines.*

    The captions of Fig. 3 and 4 have been edited to indicate that the horizontal axis represents categorical data.

    <Page 9, Figure 3 caption>Simulation results for a non-luxury with respect to different types of influencer sets, and their corresponding size n. The horizontal axis represents categorical data.

<Page 9, Figure 4 caption>Simulation results for a luxury with respect to different types of influencer sets, and their corresponding size n. The horizontal axis represents categorical data.

11. *A strong methodological concern I have is that each piece of output data appears to come from a single replication of the simulation, despite having several sources of randomness in the initialization phase. Different initial assignments of agent interest and edge weights may affect your results, so a single data point is likely inadequate. Multiple replications need to be ran to begin to account for the stochastic inputs of the model, and some measure of variability reported and analyzed (e.g., point estimate of the mean with confidence intervals). This is the only reason I recommend a major, rather than minor, revision.*

The Authors thank the reviewer for making this point. The authors have addressed it in the following manner:

Considering the stochastic nature of the model, **<Page 7, section 4, line 41>** as mentioned in the paper, "Each experimental result showcased in the following sections is found by taking an average of 10 trials".
Due to the enormity of the studied graph dataset, with the number of edges in the order of million, the randomness incorporated in the initialization process does not affect the final conclusion for a given set of initialization parameters.

The following modification has been made to the text:

<Page 7, Section 4, Left Column, Line 42> For each experiment, the graph was initialized once using the initialization parameters and due to the enormity of the studied graph datasets with the number of edges in the order of millions, multiple replications of the same initialization parameters did not alter the conclusions and findings which was observed during simulational testing.

12. *The Background section seems rather short, but that may be due to the amount of background material in the Introduction.*

The background section discusses various existing influence, propagation models. As rightly noted by the reviewer, the authors have made an effort to provide appropriate background in the introduction as well.

13. *In Sec 3c.1, you write that individuals not interested in purchasing a product do not participate in marketing. I think it's a fine modeling assumption, but I often forward marketing information to friends that are interested in things that I'd never buy!*

As mentioned in response to Comment 6, the authors acknowledge that an individual who does not purchase a product might participate in its marketing with a low probability, but make a reasonable assumption that it is unlikely to occur when compared with individuals who purchase a product.

Comments from Reviewer 2:

1. *What I could not agree on was how the network graph represents reality. The authors should clarify (if I am wrong) that why it is enough to use to random graph? From the construction of the random graph, a set of influencers (with a higher degree) are identified. Even this is acceptable to an extent. But, when the same graph is used for follower- follower interaction, the whole network becomes unrealistic (in my view). How come, a follower is connected to a follower randomly, instead, it should be an anomaly. How come, a random set of nodes connected to an influencer will be connected to each other in the same way? There is very little chance of this happening... Typically in a network of a large number of nodes. So, practically follower-follower interaction should be quite negligible. In fact, I am in favor of having a different network for that which is quite thin not the same random network. The graph representing the influencers and followers could have been more representative so that the mapping with the real world is obvious.*

The authors used a small-world random graph since it has been shown in the literature that social networks have small-world properties [10][33][24]. In addition, the authors wish to show the working of the model on different kinds of graphs. The following lines have been added to the text to make the clarification:

<Page 10, Section 4D, Lines 3> To show how the model works on a different kind of graph, we also experiment on a synthetic small world graph generated using the Watts Strogatz model [40]. This model was chosen because social networks are known to possess small-world properties [10][33][24].