# Response to Reviewer 1

Major points:

- The language could be greatly improved throughout the paper, and I recommend that a native English speaker goes over the manuscript. The English makes it indeed sometimes very difficult to follow the reasoning of the authors.

RESPONSE: We are deeply embarrassed by this and have done our best to improve the language, including reading the paper several times over and getting feedback from several native english speaker peers.

- The Methods section is very long and should be either i) shortened, or ii) partially put in Appendix for the most technical points. The language combined to the very technical points mentioned in this Methods section make it difficult to follow even for a reader familiar with the concepts involved. Also, the various logical links between the steps involved in the reconstruction of the network should be clarified. Sometimes, in the Methods section as it is, it is not obvious to identify the links between the different parts.

RESPONSE: This is very good advice and has been taken into account. Namely, we believe the primary contribution of the paper is in terms of stimulation optimization. Therefore, we summarized all matters relating to network structure and estimation in the first section of the methods and moved the details to an appendix. We hope this makes the paper much more readable.

- The implementation of the stimulation effect in the reconstructed network is not sufficiently described and discussed. This is a critical point, since it is central to the claim of the paper that stimulation optimized for a specific reconstructed network is a viable approach.

RESPONSE: The following sentence was added after Eq. 11, which mathematically defines how stimulation is modeled, to build intuition on the aforementioned equation: *"Essentially, this model of stimulation allows electrodes to deterministically elicit spikes at precise times and thereby 'override' the neurons' ambient probabilistic mode of spiking defined in Eq. \ref{eq:simul}.".* Also, our assumptions of stimulation are discussed in the discussion section: "Furthermore, our model of electrical stimulation, which assumes a single electrode can illicit a spike in a single neuron, is overly simplistic. Any stimulation will affect at least dozens of surrounding cells (Wei and Grill, 2005; Desai et al., 2014)."

- There is no experimental validation of the model predictions, which would have greatly increased the paper's potential impact. In this version of the manuscript, the authors have reconstructed a network, showed that an optimal stimulation can be found, but this has not been tested or validated experimentally. Therefore, for the purpose of this paper, any neuronal network could have been used (not necessarily the one specific for this patient), which would have led to similar conclusions regarding the algorithm use and optimal stimulation identification. Maybe the article structure should be revised in order to form a more consistent whole, maybe either focusing only on network reconstruction, or on the identification of optimal stimulation parameters for a given neuronal network with seizure dynamics.

RESPONSE: This is a good point, and one we have struggled with when writing the paper. Human data was used in order to test our algorithm with the most realistic data. However, the focus and contribution is meant to be methodological/computational. We tried to follow your advice and make the paper more readable and focus more on identification of optimal stimulation parameters rather than model estimation or experimental approach. This was done by moving the model estimation stuff to an appendix. Also, we removed the figure focusing on human data acquisition as it was tangential to our point. Of the remaining 7 figures, only the first focuses on model estimation, with the remaining focusing on seizure initiation and stimulation. We hope this focuses the paper and makes it more readable.

- Some portions of the Results section should be moved in the Methods section (most of p. 8). This contributes to decrease the readability of the paper.

RESPONSE: This is very good advice and has been done. Now the results section is much more 'streamlined' and has much less discussion/methods in it.

Minor points:

- on page 2, an "in press" paper from the authors is cited but not accessible to the reviewer, which makes it impossible to evaluate this point.

RESPONSE: this has been changed with the proper citation (Sandler et al., 2015a)

- From the description of the Neuropace device functioning, it seems that the authors confuse interictal events and the seizure activity. This should be clarified.

RESPONSE: This is a great point. We added the following qualification when describing our control strategies: "Notably, this is the same control algorithm currently used in the Neuropace RNSR device (RNSR System User Manual). (However, it should be noted that the Neuropace broadens our definition of 'seizure state' to include interictal activity, with the hope that stimulation will prevent the network from ever entering a full seizure.)"

- Why were 6 Laguerre basis functions used? What is the rationale? Were different numbers of Laguerre functions used?

RESPONSE: These parameters have been explored in previous work and L=6 was chosen based on that. This has been clarified in the text and relevant citations added: "Based on previous studies, L=6 Laguerre basis functions were used (Sandler et al., 2015, Song et al., 2007, Song et al., 2009)."

- In the Methods section, clarifications should be made regarding many steps (e.g., "The SA algorithm was run twice using different definitions for neighboring states", this is too vague).

RESPONSE: We did our best to make the methods clearer. The entire paragraph dealing with simulated annealing neighboring states was rewritten to be more clear: "The SA algorithm was run in two consecutive rounds. In both rounds one electrode was chosen whose stimulation parameter value would randomly transition to a neighboring parameter value with equal probability; however, the definition of neighboring values was different for both rounds. In the first round, the neighboring values were defined as the immediately higher frequency, the immediately lower frequency, or OFF.If the selected electrode was already OFF, all 7 of the frequency values were defined as neighbors (i.e. an OFF electrode would transition to a random frequency with equal probability). Allowing an electrode to turn off at each iteration encouraged sparsity and facilitated the identification of a subset of electrodes to more carefully optimize in the second round. In the second round, the OFF electrodes from the first round were not adjusted and the remaining electrodes transitioned either to the higher or lower frequency with equal probability."

- On page 7, "Each electrode could take" should be replaced by "Each electrode frequency parameter could take" for example.

RESPONSE: this has been corrected

- The last sentence of the 4.3 section should be revised since it seems out of place "All that needs to be done is to improve the software." Furthermore, I disagree with this, it is also to significantly improve our understanding of how electrical stimulation modulates the excitability of epileptic tissue, and the involved mechanisms. There are still numerous details to be understood.

RESPONSE: The last two sentences were significantly toned down to: "Furthermore, since this approach is fundamentally an algorithmic one and does not heavily rely on specific electrode design and placement, it may be backwards compatible with current generation devices and significantly improve their efficacy."

# Response to Reviewer 2

COMMENTS TO THE AUTHOR(S)

Sandler and colleagues have presented a ”closed-loop” neurostimulation paradigm for abating epileptic seizures. They have presented a well-designed approach to identify patient-speciﬁc neurostimulation patterns. This study tries to address one of the clinically most important contemporary challenges in epilepsy treatment. The overall aim of the manuscript is sound. The methodological/engineering part of the paper is done with a good care. While I feel that any attempts in this front are highly welcome, there are some issues that I think need more attention in this particular manuscript. I have some technical criticisms. Here is a list that is not quite complete.

Major comments:

• Although the authors have mentioned in the discussion that the emphasis of their work is more on proof-of-concept, their method seems to have a limited range of applicability. They have only used ten minutes of continuous data to estimate their models. It is very difficult to believe that the system would provide the same performance under real conditions. First of all, the main assumption in this study is that the dynamics of non-ictal data remains unchanged in epileptic patients. Most of the time, this assumption is not valid. From technical point of view, the hippocampus is not a system isolated from the rest the brain. It receives lots of information and its dynamics can change instantly. On the other hand, seizures don’t have the same dynamics even within the same patient. Therefore, if the RNN is trained on a segment of data, its estimated parameters will become suboptimal for other portions of data. To test if this assumption is correct, I recommend that the authors analyze more data from the same patient within different time intervals, and show that practically they can replicate them using their RNN. Another assumption that seems to be unrealistic is that each patients have seizures with a unique dynamics. This is not true even within the same patient. Please refer to the following paper on seizure prediction using computational modeling. In 25% of patients, the dynamics of seizures and background activities changes over time and a trained network with a fixed configuration would not provide a good performance under this circumstance. Therefore, an adaptive approach is necessary to increase the performance. Please discuss this issue in the discussion of the paper.

Aarabi and He, Seizure prediction in hippocampal and neocortical epilepsy using a model-based approach, Clinical Neurophysiology 125 (5), 930-940

RESPONSE: We admit that due to the difficulty of recording single units in humans we were unable to record an actual seizure as we had a limited amount of time with the patient. Furthermore, in order to make our main problem of estimating optimal stimulation tractable, we assumed a stationary model with stationary neuronal and seizure dynamics. We admit this is a simplification; however, as in any modeling project one needs to start somewhere, and stationarity is a very common assumption when modeling seizures (for example see Mina et al., 2013). In order to clarify this, we have added the following paragraph in the discussion:

"The current model assumes stationarity of neuronal dynamics, while in reality recorded neurons exhibit strong nonstationarities due to synaptic plasticity and electrode drift.

The stationarity assumption was necessary in order to define a tractable optimization problem. It is unknown to what extent nonstationarities in neuronal dynamics lead to nonstationarities in optimal stimulation parameters. Future work may aim to address this issue by simulating nonstationary networks such as in (Robinson et al., 2016)."

The reviewers point regarding the heterogenuity of seizures is well taken and accurate. Unfortunately, as we were not able to record actual seizure activity we had to rely on synthetic homogenous seizure dynamics. In order to be transparent about this limitation, we added the following: "Synthetic seizure dynamics were homogenous, while there is evidence that actual seizures exhibit diversity even within the same patient (Bower et al, 2012; Aarabi and He, 2014). In future work, estimating network dynamics directly from recorded seizure data will lead to more conclusive results as less assumptions about seizure initiation will need to be made."

Finally, we fully admit that the work here is still in the proof-of-concept stage and is not yet clinically viable - nor is it meant to be. Nonetheless, we believe that this work represents a significant step forward in tackling the very important problem of stimulation design. In future work, we hope to use the conceptual and computational framework established here to do more clinically relevant work.

• The second main issue is about the model itself. As mentioned in page 8 (line 18), the firing probability of each neuron is determined based on specific timing parameters of spikes from the neuron itself and those of other neurons. Mathematically talking, the approach is quite adequate for neurons with well-defined dynamics. However, the dynamics of neurons and connectivity patterns vary over time. This issue needs to be addressed by analyzing more data from the same patient. The model can be fit to two data segments within different time intervals and a comparison between the model parameters might show consistency or inconsistency between the results.

RESPONSE: See previous point regarding stationarity and the extra paragraph added to the discussion to address this issue. Note that in our model the stationarity of seizures leads directly from stationarity in neural dynamics and thus we have addressed both in the same paragraph.

• The third issue is about the way seizures have been simulated in this paper. As indicated in page 10 (line 20), the method presented to create seizures is so much simplified. It is unclear why the authors believe that simulated seizures have the same dynamics as those usually recorded from patients. It seems that by changing the baseline and lowering alpha one can generate simple seizures. But it is not specified if seizures’ temporal evolution remains the same over time. I suggest that the authors can generate seizures with different dynamics by changing these parameters and use their approach to abate them. The results should be reported accordingly.

RESPONSE: We agree that our mechanism for seizure generation makes many assumptions and simplifications, including that of stationarity: while the seizures emerge spontaneously, the underlying dynamics of the neurons remain the same throughout the simulation (both during seizure and nonseizure states). This is a reasonable model of many experimental models of seizures including high K+/low Mg (Kang et al., 2010, J. Biological Physics) and 4-aminopyridine (Barbarosie and Avoli, 1997) where pharmacological manipulations change the biophysical properties of neurons and cause spontaneous seizures to emerge. In fact these many of these manipulations such as high K+, low Mg-, raise neuronal baseline potential, which is exactly what our model does. Nonetheless, in our model there is no 'temporal evolution over time'. Rather, it is more like a Markov process which spontaneously enters and exits the seizure state. We thank the reviewer for pointing out that this is not fully accurate in real life and we will aim to address this in future work.

As far as the reviewers suggestion of changing the seizure parameters, we find that the network is very sensitive to them. There is a relatively narrow range where one gets spontaneous seizures. If the parameters are too low, then the network stays stable. If they are too high, the network saturates with spikes. Thus, these parameters are hand chosen and fixed for all simulations. We did not get into details of this as it is tangential to the paper, we are glad to discuss this more if the reviewer feels it is a good idea.

• The authors are suggested to specify the reason why only 8% of seizures simulated in the study could not be abated by their system.

RESPONSE: The following sentence was added: "In the four unsuccessful cases, it was found that while the stimulation did push the network temporarily out of the seizure state, the seizure restarted after the stimulation was turned off. "

In fact, this results somewhat expected by chance since in the 50 trials, there was 50\*.75s=37.5s of nonstimulation time. We found that the network enters the seizure state roughly once every ~11 seconds. Thus, we would expect ~3.4 seizures to in the OFF time. We did not discuss this in the paper in order to make the paper more readable, but certainly can do so if the reviewer feels it's appropriate.

• (page 8, line 11) the authors have argued that the RNN diﬀers from other artiﬁcial neural networks which have stereotyped connections between neurons. Like the RNN which is a blackbox, there are other networks like recurrent neural networks that can model any nonlinear functions provided that the number of neurons and layers are determined correctly. Please discuss this issue.

RESPONSE: For the sake of brevity and focus, we decided to remove the sentence comparing RNNs with artificial neural networks. However, the reviewer makes a good point: many complex neural networks can indeed approximate arbitrary connections and definitely the connections expressed here. In fact, our group has done work to explicitly compare the laguerre volterra model used here with neural networks (Geng & Marmarelis, 2016). However, such networks are not nearly as common as the more traditional neural nets which are not sparse and cannot express laguerre-volterra type nonlinear dynamics.

Geng, Kunling, and Vasilis Z. Marmarelis. "Methodology of Recurrent Laguerre-Volterra Network for Modeling Nonlinear Dynamic Systems." IEEE Transactions on Neural Networks and Learning Systems (2016).)

Minor comment:

***• Please specify the age and type of epilepsy for the patient and add at least one more figure illustrating a seizure from the patient.***

• The sampling rate for data collection has not been specified.

RESPONSE: This has been clarified now and the following sentence has been inserted: "Single unit neural activities (i.e., spike trains) were recorded and isolated using the Blackrock Cervello Elite electrophysiological recording system with a raw data acquisition frequency at 30k samples/sec without filtering, and a spike sorting frequency at 30k samples/sec with 500-5,000 Hz bandpass filtering."

• Please specify why the parameters were divided into 2R groups in the estimation procedure (page 5 line 17).

RESPONSE: Every input neuron has R potential input neurons (R being the total # of recorded neurons). Each recorded neuron could influence the output neuron either linearly via the 1st order kernel or nonlinearly via the 2nd order kernel. These interactions were treated as 2 different groups in order to be able to disentangle linear and nonlinear effects. The following sentence was added to clarify: "Thus, for the R=24 recorded neurons, there were 48 groups: 24 1st order kernel groups and 24 2nd order kernel groups. "

• (page 6, line 45), MFR has not been defined.

RESPONSE: this has been corrected: "To apply the clustering algorithm, the instantaneous mean firing rate (MFR) was first computed for each neuron using a 100 ms moving average filter."

***• Fig. 2, plot C and D need to be reproduced, the labels are missing or unclear.***

• Page 8, lines 11-43, this section should be moved to methods or discussion, not in the results.

RESPONSE: Good point. this section has been removed from the results section

• Page 10, line 42 and page 11 line 56, please check the grammar.

RESPONSE: First sentence the extra and was removed: "As can be seen, this method is able to reliably detect when the RNN enters and leaves the seizure state, thus effectively making it a prototype seizure detection algorithm"

Second sentence adjusted for grammar: "As can be seen from these figures, the optimal stimulation induces the network to leave the seizure state in under 250ms, and more importantly enables the network to stay in the nonictal state after the 250 ms stimulation ends."

• Page 12, last paragraph might be moved to discussion.

RESPONSE: we believe that the beginning of this paragraph is useful to provide a brief context for the results presented in the remainder of the paragraph. If the reviewer believes it should still be moved, we don’t mind fulfilling this request.

• Page 16, lines 39-42, this statement is an overinterpretation of the results, the authors should first prove that each patient has a unique network topology and dynamics.

RESPONSE: In order to stay away from overinterpretation, the sentence has been changed to: "However, our most important finding was that optimized stimulation over multiple electrodes significantly outperformed any of the previously mentioned stimulus styles by exploiting the unique connectivity and dynamics of the RNN."

• Page 17, lines 16-17, “in future …”, it is unclear based on what the authors think their results would be better with real seizures.

RESPONSE: We hypothesize our results would be better because less assumptions about seizure dynamics would need to be made. Here because we were unable to obtain real seizure data we had to simulate a synthetic one out of necessity (see earlier comments). Obviously this makes alot of assumptions about the nature of seizures which may be false. If we had actual seizure data, we could estimate the RNN directly from it without relying on any assumptions, which we hypothesize would improve results. In order to state this more clearly, we have changed that sentence to: "Synthetic seizure dynamics were homogenous, while there is evidence that actual seizures exhibit diversity even within the same patient (Bower et al, 2012; Aarabi and He, 2014). In future work, estimating network dynamics directly from recorded seizure data will lead to more conclusive results as less assumptions about seizure initiation will need to be made."