Spatio-temporal dynamics in nanowire networks*

Ann Author† and Second Author‡
Authors' institution and/or address
*This line break forced with *
(MUSO Collaboration)

Charlie Author§
Second institution and/or address
This line break forced and
Third institution, the second for Charlie Author

Delta Author
Authors' institution and/or address
*This line break forced with *
(CLEO Collaboration)
(Dated: June 5, 2020)

An article usually includes an abstract, a concise summary of the work covered at length in the main body of the article.

Usage: Secondary publications and information retrieval purposes.

Structure: You may use the `description` environment to structure your abstract; use the optional argument of the `\item` command to give the category of each item.

I. INTRODUCTION

II. METHOD

A. Graph Generalization

Talk about how to map the original nanowire network on to a graph. Cite Zdenka here.

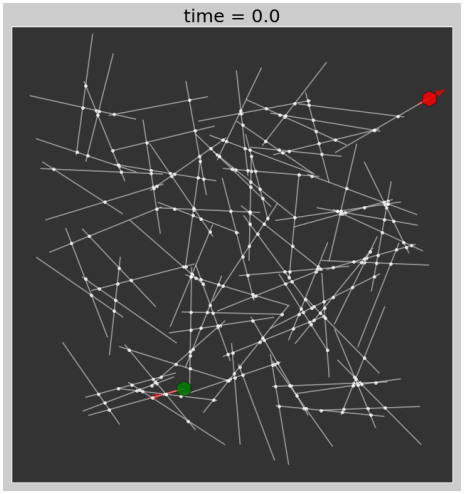

FIG. 1.

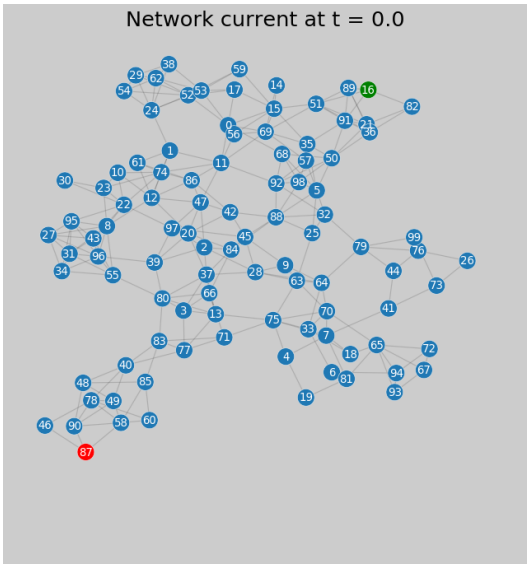

FIG. 2.

B. Simulation

When connected to external voltage biases, nanowire networks behave like traditional electrical networks. Each component in the network obeys Kirchhoff's law. Hence the voltage distribution across the network can be obtained by solving ([1]):

$$\mathcal{L}^{\dagger} V = I, \tag{1}$$

where $\mathcal{L}^{\dagger}$ is the expanded graph Laplacian of the network, composed as:

* A footnote to the article title
† Also at Physics Department, XYZ University.
‡ Second.Author@institution.edu
§ http://www.Second.institution.edu/~Charlie.Author

$$\mathcal{L}^{\dagger} = \left[\begin{array}{c|c} \mathcal{L} & C \\ \hline C^T & \mathbb{0} \end{array}\right], \tag{2}$$

in which $\mathcal{L}$ is the graph Laplacian and C represents the nodes (wires) connected to external electrodes.

$$\mathcal{L} = D - W. \tag{3}$$

Here W is the weighted adjacency matrix of the network. The weights on edges are determined based on their conductance:

$$\mathbf{W}_{ij} = \mathbf{A}_{ij} G(i,j). \tag{4}$$

And D is the weighted degree matrix generated from W:

$$d_i = \sum_{k=1}^{N} \mathbf{W}_{i,k}, \tag{5}$$

$$\mathbf{D} = \mathbf{diag}(d_i). \tag{6}$$

C. Centrality

Centrality **Put some reference here** is an important measure that helps understand the fundamental structures and connectivities of networks. Betweenness centrality of nodes and edges in networks can demonstrate .... Meanwhile, closeness centrality of nodes can be used to interpret the ..... [2].

A variation of centrality measures based on current flow model proposed by Brandes and Fleischer [3] is employed here. As the electrical dynamics in our networks fall in the class of **random walks** (not sure about here).

The betweenness centrality of a edge in the networks can be determined by:

$$c_{CB}(e) = \frac{\sum\limits_{s \neq t \in V} \tau_{st}(e)}{(n-1)(n-2)}, \tag{7}$$

where $\tau_{st}(e)$ is the current flow through edge e between nodes s and t, while n stands for number of nodes in the network.

The closeness centrality of a node in this context is structured in the same way as normal closeness centrality, with distance measured based on effective resistance rather than graphical distance. Therefore it is given by:

$$c_{CC}(v) = \frac{n-1}{\sum\limits_{v \neq w \in V} R(v,w)}. \tag{8}$$

In the context of a unit st-current in the network, $R(v,w)$ represents the effective resistance between nodes v and w.

D. Communicability

Communicability in a network represents .... [4].

In this work communicability is calculated based on the weighted connections [5]. A symmetric matrix $\mathbf{M}$ is generated to depict the communicability distribution, where $\mathbf{M}_{ij}$ represents the communicability between nodes i and j. Therefore, $\mathbf{M}$ reads:

$$\mathbf{M} = \exp\left(\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}\right). \tag{9}$$

The communicability of a node can be defined as how communicable the node is to the rest of the network, which will have the expression as:

$$m_i = \sum_{k=1}^{N} \mathbf{M}_{ik}, \qquad k \neq i. \tag{10}$$

E. Modularity

Modularity is a network metric that ..... **talk about it here** [6].

Recent studies in neuro-science demonstrated that modularity plays an important role in .... **talk about it here** [7].

Modularity ($\mathbf{Q}^w$) in a weighted network can be calculated by the Louvain method [8]:

$$\mathbf{Q}^w = \frac{1}{\mathbf{D}} \sum_{i,j} \left[\mathbf{W}_{ij} - \frac{d_i d_j}{\mathbf{D}}\right] \delta(m_i, m_j), \tag{11}$$

in which $w_{i,j}$ is the weight of the edge between nodes i and j, I^w represents sum of all weights in the network, k_i^w stands for the weighted degree of node i, m_i is the community where node i is located, and δ is the Kronecker delta function.

F. Information Dynamics

Transfer entropy is an effective metric for information dynamics. It explains ..... **cite Joe here**.

$$T_{Y\to X} = \sum_{u_n} p(u_n) \log \frac{p(x_{n+1}|x_n^{(k)}, y_n^{(l)})}{p(x_{n+1}|x_n^{(k)})}. \tag{12}$$

n is a time index, u_n represents the state transition tuple $(x_{n+1}, x_n^{(k)}, y_n^{(l)})$, $x_n^{(k)}, y_n^{(l)}$ represent the k and l past values of x and y up to and including time n. **rephrase here**.

The transfer entropy across an edge $(e_{i,j})$ is calculated by: **Have to write these equations in a better manner.**

$$TE_{i,j} = T_{V_i \to V_j} \tag{13}$$

And therefore the average outward TE and inward TE of a node can be calculated by:

$$\langle TEout_i \rangle = \frac{\sum\limits_{i,j,A_{i,j}\neq 0} TEout}{\#\text{ of edges connected to i}}. \tag{14}$$

Meanwhile, active information storage was introduced by Lizier et al. (**cite Joe)** to measure the effect from the past process $\mathbf{X}$ to the next observation X_{n+1}, which is given by:

$$A_X = \lim_{k\to\infty} A_X(k), \tag{15}$$

$$A_X(k) = I(\mathbf{X}_n^{(k)}; X_{n+1}) \tag{16}$$

The active information storage in context of edge voltage distribution is thus calculated by:

.....

G. Benchmark

Regular reservoir computing realizations are typically initialized with homogeneous or randomly generated states. The system then evolves based on specific tasks. **WHY we want different initial states?**

A nanowire network enters different stages during its activation. Its states at different stages typically have different voltage distribution, unevenly grew filament states and therefore diverse conductance levels. In this section the network's states at different activation stages are extracted and used as initial states for benchmarking. Starting from these pre-activated states, memory capacity of the network is measured to estimate the network's ability of reconstructing its short-term memory **Cite Jaeger**. Non-linear transformation is used to test the network's learning ability **Cite Kevin**.

A randomly generated time series from uniform distribution in the interval $[-2, 2]$ is used as the input signal for memory capacity test. A linear combination of the network's state at t (node voltage) is applied to reconstruct the previous input at $t-k$. $k = 1...100$ are used to investigate different delayed intervals. The $k-$delay memory capacity is defined as:

$$\mathrm{MC}_k = \frac{\mathrm{cov}^2(\mathbf{u}_{t-k}, \mathbf{o}_t)}{\sigma^2(\mathbf{u}_{t-k})\sigma^2(\mathbf{o}_t)}. \tag{17}$$

The generic memory capacity for the network can thus be calculated as:

$$\mathrm{MC} = \sum_{k=1}^{\infty} \mathrm{MC}_k. \tag{18}$$

The non-linear transformation task makes use of sinusoidal wave as the input signal and try to transfer the signal to other periodic signals in the shape of square or sawtooth. Similar to memory capacity test, linear combinations of the network's states are done to fit the target signal. The accuracy is calculated by $1 - \mathrm{RNMSE}$:

$$\mathrm{RNMSE} = \sqrt{\frac{\sum(\mathbf{Y}-\mathbf{O})^2}{\sum \mathbf{O}^2}}. \tag{19}$$

General questions about result part

- Do we need the current flux vs node centrality? (basically a straight line)
- Do we need communicability? Since COMM = $\sum G^k$ in my definition.
- For the coloring of TE scatter, ON time might be a more straight forward scheme, e.g. the color-way in AIS vs centrality plot.
- How to name different regimes of activation and what will be the time points to distinguish them.
- What will be the best incentive for different initial states in benchmarking. (As far as I know, most of the reservoir computing realizations are based on homogeneous or randomly generated initial states. This pre-activated way might be a good alternative.)
- The word usage of node-edge and wire-junction. (In the context of graph/network, seems like node and edge are more understandable. But when talking about ON/OFF and filament states, I feel like junction makes more sense. Should we just use one pair of representation in the paper?)

III. RESULTS

A. Centrality and Dynamics

As the network evolves, junction filament states will grow faster at more "central" positions. A winner-takes-all current path **Cite Dublin paper** will form at the most central positions first, then branch out to the rest of the network. At a specific time t, current-flow betweenness centrality of junctions can be determined by Eq. 7. The junctions with higher centrality are more likely to exhibit higher filament state and voltage. Meanwhile, voltage ($\vec{V_e}$) and filament state ($\vec{\lambda}$) are strongly coupled on the junctions. Higher voltage on a junction will lead to faster growth of its filament state. In return, higher filament state means the junction will have a higher conductance (**refer to equation here**), which increase the chance for the junction to achieve higher voltage.

a. Filament state The absolute values of filament states $|\vec{\lambda(t)}|$ as a function of edge current-flow betweenness centrality are plotted on fig. 3 with its data points colored by corresponding conductance states. The plot shows that junctions at more preferable positions will have faster filament growth. The color spectrum on the plot demonstrates the conductance transition of junctions based on their filament states. These high centrality junctions are having significantly higher conductance **(WORDING) since their filament states are greater than the threshold**, and thus form the first current path.

b. Voltage distribution Voltage distribution on junctions $|\vec{V_e(t)}|$ as a function of centrality is also included in fig. 3 with the same coloring scheme. The data points can be clearly classified into three regimes. The two linear regimes at the left and right ends correspond to the OFF and ON junctions. The transition regime in between which represents the junctions whose conductance are at some middle values in the tunneling model. The plot indicates that within each regime, junctions with higher centrality will have higher voltage. The earlier current paths consist of these higher centrality junctions.

c. Current flow **Not sure if we have to keep this** The cumulative current flow through nanowires shows are linear relationship versus node current flow betweenness centrality.

B. Centrality and Functionality

NOT sure if we want to keep communicability. Since COMM $= \sum G^k$

WORDING HERE as well. Intuitively one would think certain nodes in the network will have a more important role for specific functionalities. Here analysis of the network's functionality such as communicability and transfer entropy are presented. The results suggest that more central wires and junctions have higher importance in information dynamics ...**MAYBE communicability**.

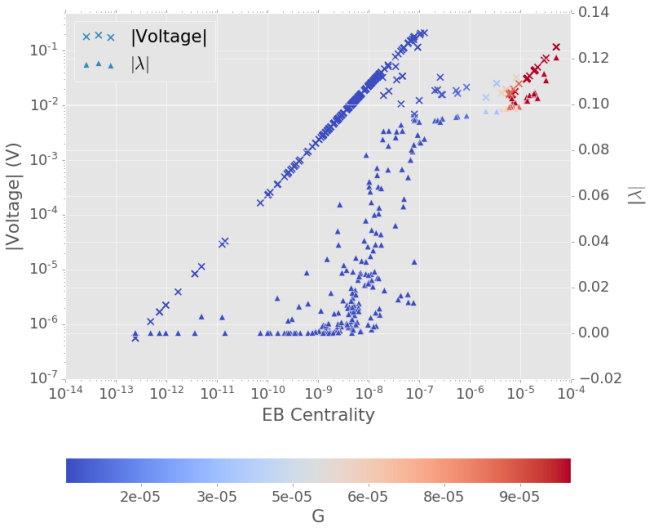

FIG. 3. Absolute values of voltage and filament as a function of centrality at T = ?? s. Current path was formed at ??. The data points are colored based on the junctions' corresponding conductance. **ASK Zdenka. Maybe junction on time will be a more obvious metric.**

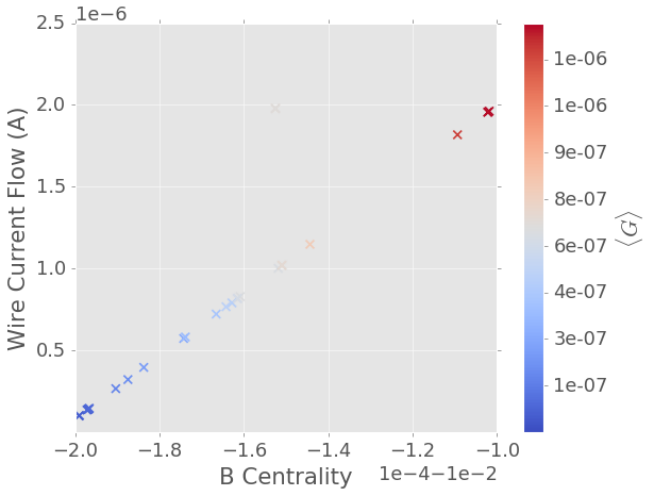

FIG. 4. Current flow on a nanowire vs node centrality

The communicability of nodes at t can be calculated based on eq. 10. At an arbitrary time point during the activation (Right after first current path formation here), when plotted as a function of current-flow closeness centrality on log-log scale, communicability shows as a close to exponential curve, indicating that the nodes at more central positions are more communicable to the rest of the network as this time point.

The node closeness centrality also shows an interesting correlation with transfer entropy. For any arbitrary node in the network, the out TE is considered as the average of transfer entropy measured from this node to its neigh-

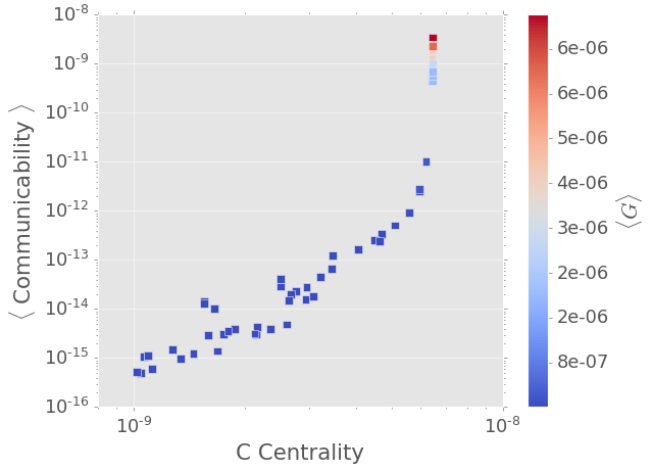

FIG. 5. **Node communicability vs closeness centrality at $t = 2s$.** Node communicability is calculated by eq. 10. Closeness centrality is calculated by eq. 8. The plot is done on log-log scale. As the nodes at more central positions are having much higher communicabilities (and higher conductance). And the other nodes follows a close to linear trend.

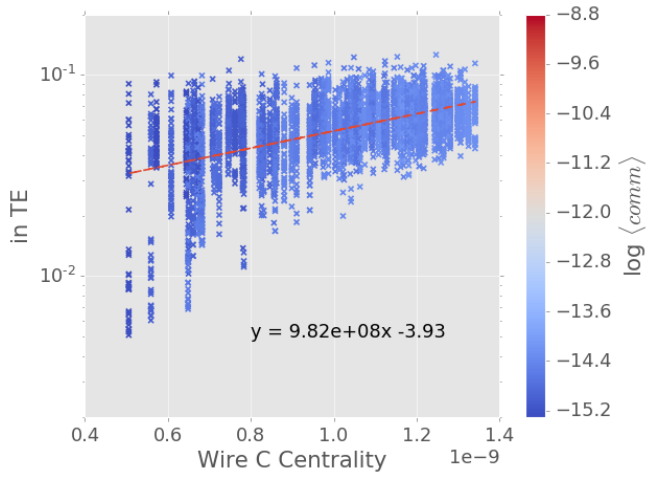

FIG. 6. **Might not need color here. Node in-TE as a function of centrality. (Out TE in supplement)** 50 repetitions with different source-drain pairings on the same 100 nw network are plotted. Each data point represent one junction in one repetition. TE is averaged across the whole time-series. Node centrality is calculated based on the current-flow closeness algorithm at $t = 0$. Communicability is calculated by Eq **equation** at $t = 0$ as well. All the source-drain pairings are controlled to have same graphical distance. Identical Mackey-Glass signals are applied to these simulations to activate the network. Each data point on the plot represent one node in one simulation. A close to linear correlation between TE and centrality can be determined from the plot, which means nodes with higher centralities tend to have richer dynamics for taking in and sending out information. **Intrinsic property of the network that does not have any dynamics. We have structural information here.**

bors. And similarly the in TE is thus the transfer entropy from all the neighbors to this node. **(Have to write equations to define in and out)**. With 50 repetitions of different source/target pairing (while the graphical distance in between is kept constant), TE averaged across the whole time-series shows a close to linear relationship with centrality on a semi-logy scale. Centrality here is calculated at $t = 0s$, which reflects the effect from the structural properties of the network (as all junctions are having same weight at $t = 0s$). Nodes at more central positions have the potential to send out and take in more information flows and exhibit richer dynamics.

Local observations of TE around the formation of winner-takes-all current path ($T = ..$ here) are studied as well. Fig. 7 shows the in-TE vs closeness centrality around activation. On a semi-logy scale, the high centrality nodes are exhibiting high TE and communicabilities. These nodes are connected to junctions that are turned ON already. On the other hand, the group at the left of the figure consists of low centrality nodes. The transition from low TE to high TE with respect to centrality increase validates the analysis in fig. 6.

Furthermore, fig. 8 suggests that junctions with higher betweenness centralities tend to have lower active information storage and will be turned on relatively earlier. This kind of behavior is consistent with our expectation since the junctions at more central locations tend to get higher voltage, thus the filament growth will be faster and the high-conductance state can be reached earlier. In addition, these rapid filament growth and conductance increase leads to stronger information dynamics on central junctions. Thus their future states are less predictable with their own history, which means the AIS will be lower.

C. Time-series analysis

Time series analysis of the network's activation can be even more intriguing. Fig. 9(a) shows the network's collective conductance time-series. The activation of the network mainly have three regimes - the resting (**maybe another word** ?ground ?low-conductance) regime, the transition regime and the stable (?equilibrium ?high-conductance) regime. Key events during the activation coincide with each other on time scale.

In the resting regime, all junctions are OFF. Collective conductance of the network remains low. Filament states on junctions start to grow based on their voltage. Increase of conductance will first take place at junctions with higher centralities. Information flow is high when the signal first came in, but in general at a low level in this regime. The modularity level stays constant since all junctions are at low-conductance state. A peak in modularity will then emerge when several junctions in the network starts to have higher conductances as they are forming "highlands". When more junctions experience conductance increase, modularity starts to drop.

The transition regime starts when the first junctions is turned on (**Here it's the tunneling model so maybe**

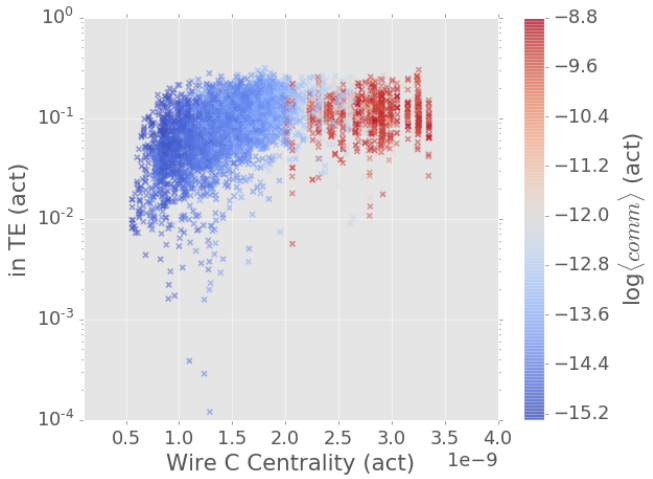

FIG. 7. **In-TE vs closeness centrality around activation.** The same activation test is done as fig. 6. TE is averaged in a 0.2s window right before the activation (mainly to cancel out the randomness). Centrality is measured at the current path formation time.

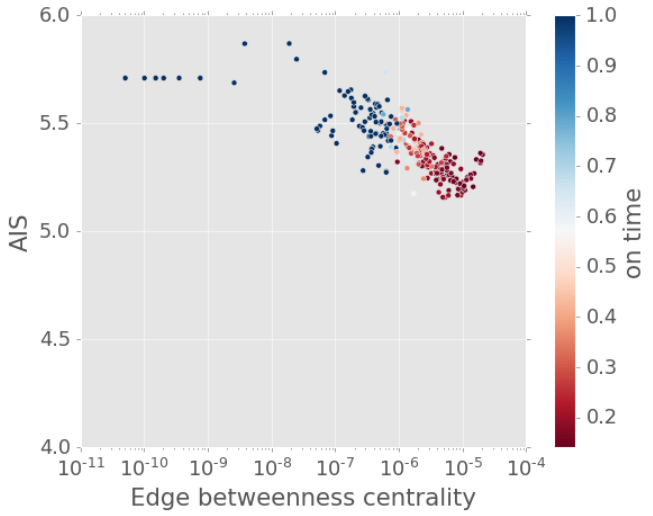

FIG. 8. Active information storage of junctions as function of betweenness centrality.

some other wording. Like some switches are exhibiting higher conductance). Collective conductance of the network starts to increase significantly as the time derivative will have a spike. Modularity of the network reaches the bottom level **wording here**. With more junctions turned ON, the first current path (winner-takes-all) will be formed and the conductance keeps increasing with a lower path. Modularity starts to increase again as the nodes on the current path stands out from the rest of the network. Information flow will increase to a higher scale (A spike in Gaussian data).

The transition regime ends when most junctions are at high-conductance states (the rest are not reachable with this specific source-drain pairing). Collective con-

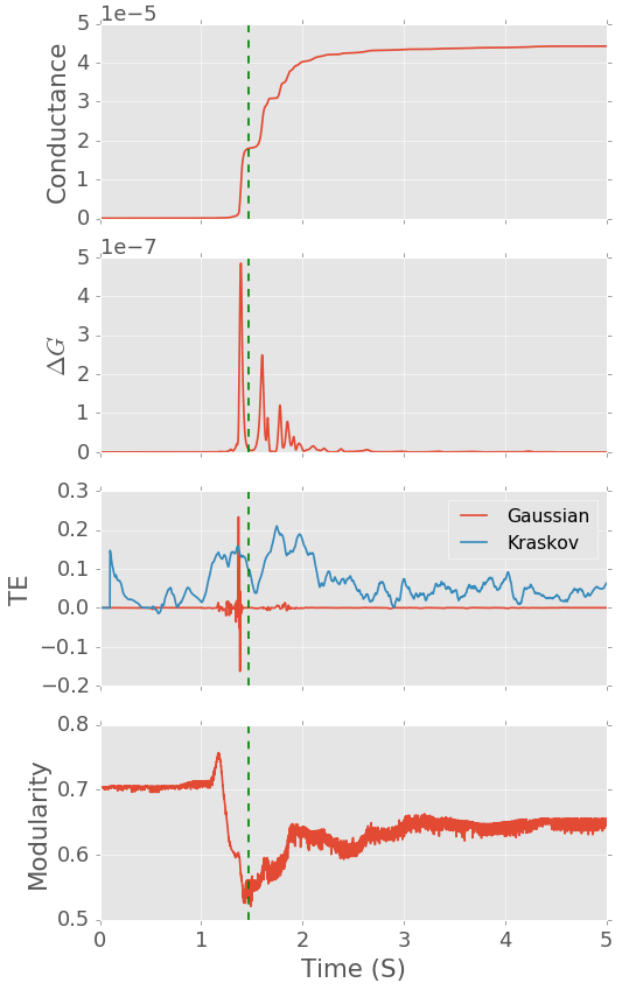

FIG. 9. **Time series data for network measures.** The dashed vertical green line represents the time of first current path formation.
(a) Collective conductance of network as a function of time. The collective conductance started increasing when the first junction is turned on. The growth will slow down after the first current path is formed.
(b) The time derivative of conductance. The derivative is calculated using second order accurate central differences on the collective conductance time series data. The growth regime of conductance shows a great correspondence of ΔG's spikes.
(c) Transfer entropy as a function of time. The transfer entropy time series data here is calculated with the Kraskov estimator and averaged across the network. A moving average with window size of 100 is applied to smooth the curve. Transfer entropy in the activation period is significantly higher than the rest.
(d) Modularity as a function of time. The weighted modularity of network is calculated by Louvain method(Eq. 11) As the network evolves, modularity will first increase as high conductance junctions are spread across the network and having their own communities. With the formation of the first current path, the modularity will have a significant drop since these isolating communities are connected. Then the modularity will keep having such fluctuations with smaller scales as more junctions are turned on and forming new current paths.
Gaussian -linear inter, KSG - nonlinear, sensitive

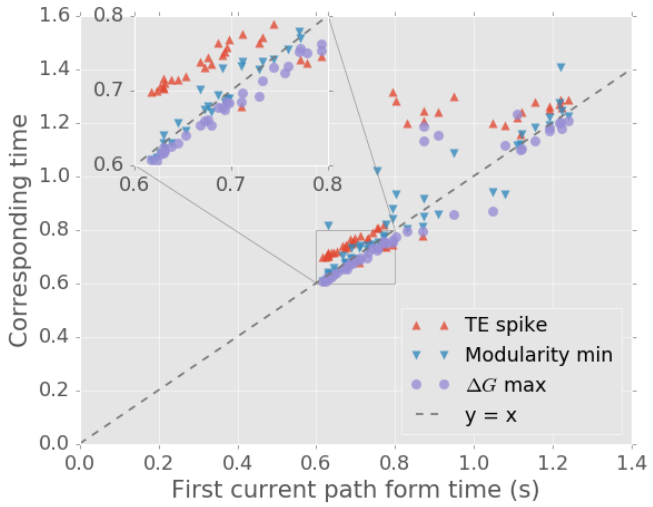

FIG. 10. **Key events time as functions of first current path formation time.** The same 100 network is used for 50 simulations with different source-drain pairings. The graphical distances between sources and drains are controlled to be constant. Identical Mackey-Glass signal with random amplitudes ranging between 2-5 V are applied to these simulations. Simulations with higher amplitudes will form current paths earlier.
X-axis represents the current path formation time in different simulations. Y-axis represents the corresponding time for key events to take place in these simulations. As shown in this figure, key events such as ΔG maximum, TE spike (in Gaussian estimation) and modularity drop will happen around the first current path formation time.
Rephrase here Thus the network will be optimal for information processing and specific task around the transition phase when the first current path is forming and the network is turning on.

ductance will be stable at a high level. Information flow will decrease to a lower scale similar to the resting regime. The modularity will also be stable in this regime.

Snapshots of the network in different phases shown in fig. 11. Nodes are colored with the sum of inTE and outTE averaged in the last $0.2s$ time window. Similarly, junctions are colored by the sum of the TE flows in both direction in the last $0.2s$ time window. (Kraskov estimator is used here.) At $t = 0.5s$, the whole network is at rest. None of the nodes or junctions are having significant information transfer. When the first current path is formed at $t = 1.465s$, some nodes and junctions start to have significant information flow, but most flows are happening around the central nodes. When $t = 2.0s$, the information flow is even stronger as more nodes are having high in and out TEs. Meanwhile, even nodes at peripheral positions are having significant information flow. After the network is in the equilibrium regime for a while ($t = 2.5s$), the information flows decay as the local TE values are significantly lower than the transition regime.

A more thorough validation of the hypothesis is done activating the same network with different source-drain pairings. The graphical distance between source and drain in each repetition is controlled to be constant. Identical Mackey-Glass signals with arbitrary amplitude (**maybe another word like strength/ xxx parameter**) between $2 - 5V$ are delivered to the network. The times when the key events take place are recorded and plotted as fig. 10. **Wording HERE** With the current path formation time on the x-axis, times for other key events line up fairly well. When the amplitude is larger, the first current path will be formed earlier and similar for other events. With amplitude ? greater than 3.5V, the key events line up linearly. With small amplitudes, the activation takes more time and thus more fluctuations in the dynamics are captured. These fluctuations contributes to the outliers in the plot.

D. Pre activation task

Why we care about different initial states?
The same 100-nw network is activated with Mackey-Glass signal. The states of the network at different stages during the activation process are extracted and employed as initial states for benchmarks tasks. Memory capacity and non-linear transformation are used to test the network's performance at different activation levels (details in method part).

Fig. 12 (a) shows the memory capacity of the network with respect to different initial states. The tests with pre-activation between 1 to 2 seconds are showing significantly better performance compared to the rest. The average active information storages corresponding to these tests are shown in fig. 12 (c). The states with better performance are exhibiting higher active information storage during the tests. Since memory capacity of the network is indicating its capability of short term memory, higher MC coincide with higher AIS is not surprising.

Meanwhile, fig. 12 (b) shows the non-linear transformation performance with different pre-activated states. Similar to the results in MC tests, the states with 1 to 2 seconds of pre-activation perform better than others. However, with roughly around 1.5 seconds of pre-activation, the performance experience a notable drop, which could be caused by the formation of the first current path. As the degree of freedom of voltage distribution is low when the first current path is forming. The averaged transfer entropy in non-linear transformation tests is shown in fig. 12 (d). The better performing states show correspondence with higher averaged transfer entropy. Non-linear transformation test requires certainly level of degree of freedom to achieve relatively good performance. According to **TE equation**, those higher TE tests are exhibiting stronger information dynamics, which is mainly caused by the complexity and uncertainty of voltage distribution. Thus better NLT performance can be explained by higher transfer entropy values in those tests.

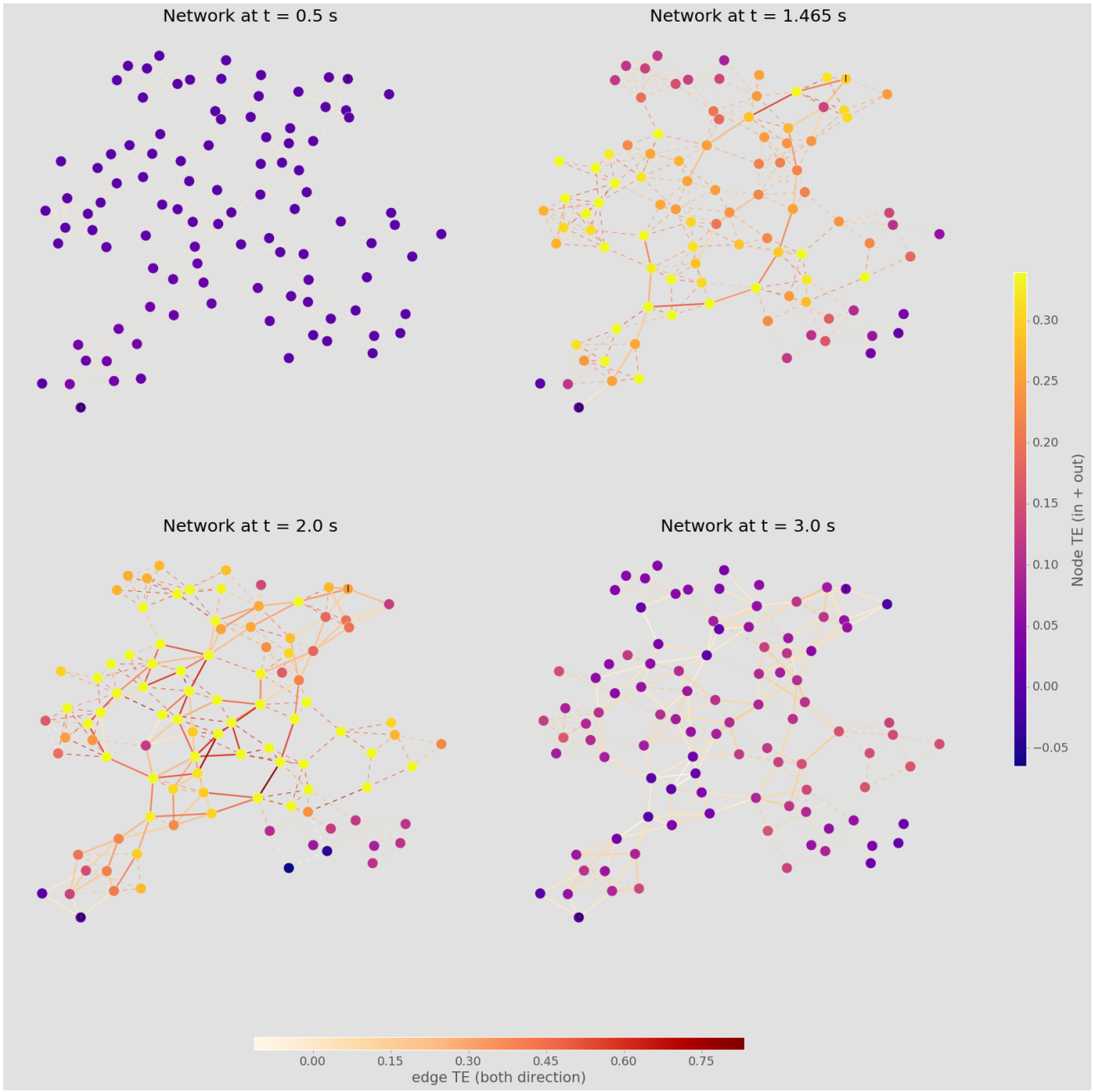

FIG. 11. **Network information flow snapshots at different time points.** Snapshots of the network are taken at time points before activation ($t = 0.5s$), when the first current path formation ($t = 1.465s$), when the network finishes large scale activation ($t = 2s$) and when the network is stable with high collective conductance ($t = 2.5s$). Nodes are colored with time-averaged TE flow (inTE + outTE) within the last 0.2 second window using the Kraskov estimator. junctions are colored by the sum of corresponding transfer entropy in both directions. TEs are increasing when the network is being activated and richer dynamics emerge. After the network reaches a stable state (t = 2.5 s), the TEs activities decay as well.

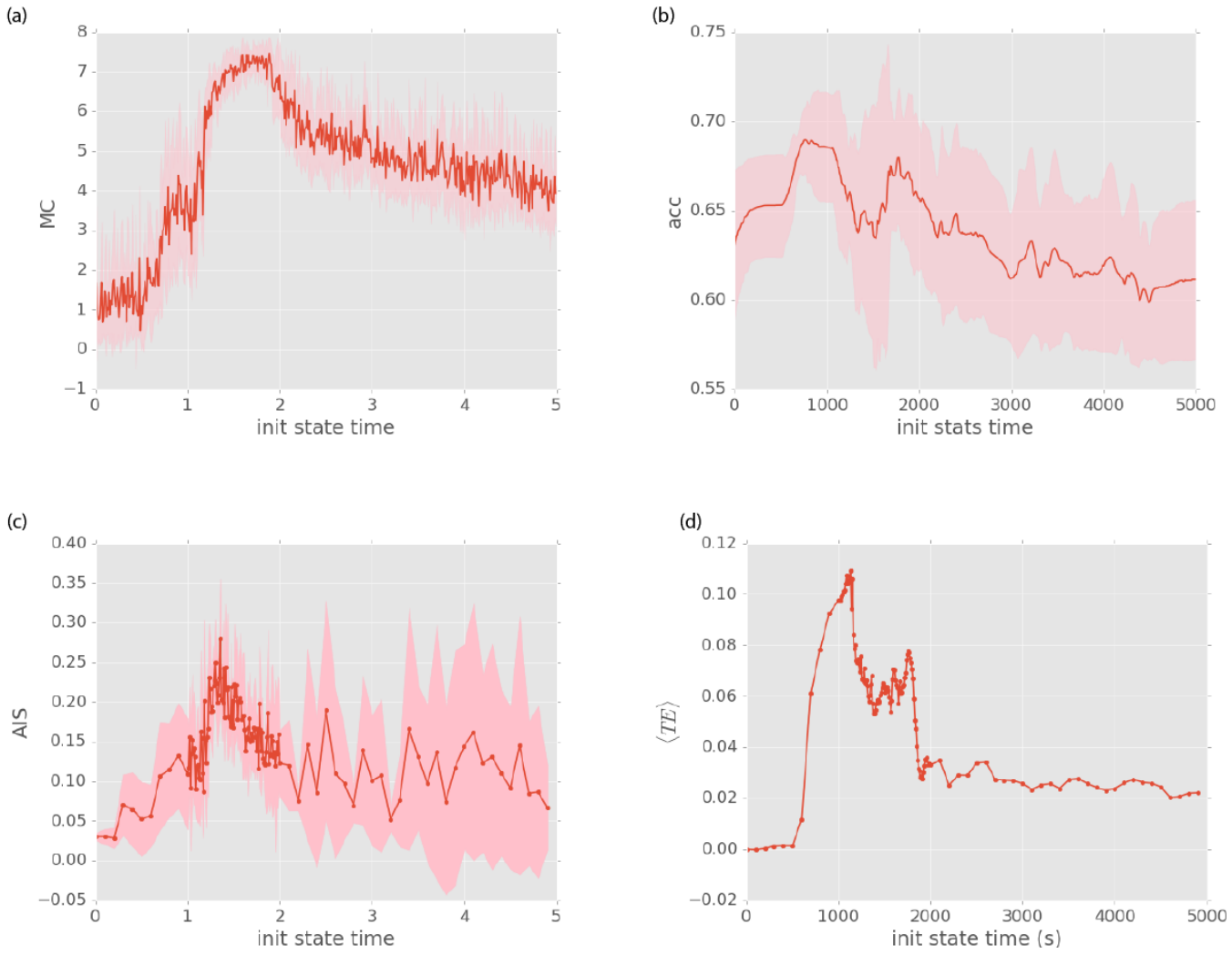

FIG. 12. **Benchmarks of the network with different initial states.** The network's voltage distribution and junction filament states are extracted every 10 time-step during the activation with a Mackey-Glass signal input. Memory capacity/ non-linear transformation tests are then performed accordingly initiating from those pre-activated states.
(a) Memory capacity result with respect to different initial states.
(b) Non-linear transformation performance with respect to different initial states.
(c) Average active information storage of memory capacity tests.
(d) Average transfer entropy of non-linear transformation tests.

IV. DISCUSSION

V. CONCLUSION

ACKNOWLEDGMENTS

We wish to acknowledge the support of the author community in using REVTEX, offering suggestions and encouragement, testing new versions, ...

Appendix A: Appendixes

Appendix B: A little more on appendixes

[1] F. Dorfler, J. W. Simpson-Porco, and F. Bullo, Electrical Networks and Algebraic Graph Theory: Models, Properties, and Applications (2018).
[2] M. Newman, *Networks: An Introduction* (2010) pp. 1–784.
[3] U. Brandes and D. Fleischer, Centrality measures based on current flow, in *Lecture Notes in Computer Science*, Vol. 3404 (2005) pp. 533–544.
[4] E. Estrada and N. Hatano, *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, Tech. Rep. 3 (2008).
[5] J. J. Crofts and D. J. Higham, A weighted communicability measure applied to complex brain networks, Journal of the Royal Society Interface **6**, 411 (2009).
[6] M. Rubinov and O. Sporns, Complex network measures of brain connectivity: Uses and interpretations, NeuroImage **52**, 1059 (2009).
[7] D. Godwin, R. L. Barry, and R. Marois, Breakdown of the brain's functional network modularity with awareness, Proceedings of the National Academy of Sciences of the United States of America **112**, 3799 (2015).
[8] V. D. Blondel, J. L. Guillaume, R. Lambiotte, and E. Lefebvre, Fast unfolding of communities in large networks, Journal of Statistical Mechanics: Theory and Experiment **2008**, 10.1088/1742-5468/2008/10/P10008 (2008), arXiv:0803.0476.