

Hidden Causal Manifolds in Human Brain Activity

Itay Baruchi⁽¹⁾, Vladislav Volman⁽¹⁾, Vernon L. Towle⁽²⁾ and Eshel Ben-Jacob^{(1),(3),*}

(1) School of Physics and Astronomy
Tel Aviv University
69978 Tel Aviv Israel

(2) Department of Neurology, Surgery and Pediatric
University of Chicago Hospitals
Chicago, Illinois USA

(3) The Center for Theoretical and Biological Physics
University of California San Diego
La Jolla, California 92093 USA

* Corresponding author eshel@tamar.tau.ac.il

Abstract

We propose that information processing in computational bio-networks from the genome to the central nervous system (CNS) is performed in the space of functional correlations. To test this hypothesis, a novel functional holography (FH) approach is introduced for analyzing the complex activity of biological networks in the space of functional correlations. Our approach is guided by the "whole in every part" nature of a holograms – a small part of a hologram will generate the whole picture but with lower resolution. The analysis starts by constructing the space of functional correlations from the similarities between the activities of the network components by a special collective normalization, or affinity transformation. Using dimension reduction algorithms like PCA, a connectivity diagram is generated in the 3-dimensional space of the leading eigenvectors of the algorithm. The network components are positioned in the 3-dimensional space by projection on the eigenvectors and connected with colored lines that represent the level of the similarities. Temporal (causal) information is superimposed on the resulted diagram by coloring the node's locations according to the temporal ordering of their activities. By this analysis, the existence of hidden manifolds with simple yet characteristic geometrical and topological features in the complex biological activity was discovered from cultured networks through modeled neural networks to recorded human brain activity. We propose that our findings hint that the functional holography analysis is consistent with a new holographic principle by which biological networks regulate their complex activity to perform information processing in the space of functional correlations.

1. Introduction

The activity of complex, multi-component computational biological networks is often represented in terms of similarity matrices and their corresponding connectivity diagrams. Examples range from metabolic pathways, through gene expression, to recorded brain activity. In general, the matrix element $S_{i,j}$ is the computed similarity between the activities of components i and j). The similarity can be based on different measures such as cross-correlations, coherences and mutual information, depending on the studied network. In the case of gene-expression measurements using DNA-microarrays, the similarity is usually the inter-gene expression correlation [1], while in recorded brain activity (e.g. EEG, MEG and ECoG), it is the inter-channel coherences [2].

A common approach in the studies of similarity matrices is to apply clustering algorithms to identify underlying sub-groups (clusters) of components that have higher inter-similarities [3]. Many advanced algorithms have been devised according to the specific systems and the assumed motifs that are looked for. In the dendrogram clustering algorithm, for example, the matrices are reordered to place together components with higher similarities. In the principal component algorithm (PCA) and its various variants, like SVD and ICA, the components are projected on a low dimension manifold whose axes are evaluated to capture most of the relevant information in the similarity matrix. The clustering algorithms are based on the implicit notion that the similarity matrices represent a high dimension space of similarities: N -dimension space for a network composed of N components.

Often, the similarity matrices are also visualized by the construction of their corresponding connectivity diagrams, in which a line is drawn between each two component whose similarity is above some threshold. Usually, the lines colors or gray levels represent the level of the similarities. For hard-wired networks (e.g. cultured neural networks), the diagram is constructed according to the components positions in the real network, by placing them on the diagram according to the physical distances. Both for hard-wired networks and distributed networks like gene-networks, the connectivity diagrams can be presented as similarity circles in which the components are placed along a circle and linked with lines that represent the similarities.

In many cases, one can also extract information about the temporal ordering in the activity of the different components, like phase coherences in recorded brain activity or timing between neurons firing in cultured networks [4]. This temporal information can be represented as a temporal ordering matrix whose $T_{i,j}$ describes the relative timing or phase difference between the activities of components i and j .

There has been rapid progress in the fields of data mining and bioinformatics, with new and more advanced visualization approaches and clustering algorithms being continuously developed [5]. Yet, many of the fundamental issues related to the interpretations of the results or, more specifically, the “reversed engineering” from the observed activity to the underlying causes are still to be resolved. The development this method has been motivated by the following goals:

1. To relate the similarity in the activity of two components to their functional dependence.
2. To relate the similarity between two components to the similarity of each with all other components.
3. To compensate for the common limitation incurred by measuring the activity of a fraction of the network components only.
4. To reduce the effect of the inherent noise both in the measurement procedure and in the biological activity itself
5. To identify underlying simple functional motives in the observed complexity. This quest was motivated by the assumption that such motives must exist if the biological network is to be able to regulate its own complexity.
6. To connect the observed temporal ordering (activity propagation) to underlying causal motives (information propagation and causal connectivity).
7. To identify functional sub-groups (functional clusters) and to reveal the functional connectivity between these sub-groups.
8. To be able to compare the activity (similarity matrices) of two different networks or different modes of behavior of the same network.

The functional holography (FH) approach is a mathematical concept of visualizing the network in an abstract 3-dimensional space of functional correlations calculated from the similarity matrix. In other words, we construct a dual network by placing the components in the abstract space and linking them according to the similarities. The mathematical procedure involves the following steps:

1. Evaluation of the similarity matrix $S_{i,j}$ between component i and j .
2. Computation of the similarity distances $D_{i,j}$ – the Euclidian distance between the position of components i and j in the N-dimensional space of similarities.
3. Collective normalization of the similarities. The above defined distances are used to normalize the similarities and obtain functional correlations or affinities. That is,

$$A_{i,j} = \frac{C_{i,j}}{D_{i,j}} \quad (1)$$

In other words, we transform the similarity matrix to a new affinity matrix $A_{i,j}$.

4. Dimension reduction - projecting the N-dimensional affinity space onto a 3-dimensional space that captures the maximal information, using a standard clustering algorithm like PCA (used here), SVD or ICA. The axes of the space are the three principal eigenvectors of the PCA, and the components position is determined by the projection of the affinity matrix on these three eigenvectors.
5. Construction of a functional manifold – a connectivity diagram in the 3-dimensional space generated by linking the nodes (component positions) by the original (non-normalized) similarities. In addition (optionally), to capture also the topological properties of the manifold, a curved surface is interpolated between the nodes.
6. Inclusion of temporal (causal) information – the activity propagation on the manifold. When information is available, the temporal ordering matrix is evaluated. The relevant information can be added to the manifold in two ways: 1. Adding arrows to the links between the nodes according to $T_{i,j}$. 2. evaluating a collective temporal ordering of all the nodes and marking their relative timing by colors or gray levels.
7. Holographic comparison and superposition. The idea is to compare the activity of two networks by projecting the affinity matrix of one network on the PCA leading vectors computed for the other, and vice versa. Two modes of behaviors of the same network may

be compared the same way. By superposition, we refer to the projection of each affinity matrix on the mutual PCA leading vectors computed for the combined matrix.

8. Holographic zooming. If we want to magnify a part of the manifold in order to capture more details, simple re-scaling used to magnify a picture will not do. Instead, it is necessary to perform a new iteration of the analysis, starting with a sub-similarity matrix for the cluster of components at the corresponding part of the manifold to be magnified.

The stages (5) - (8) of this functional holography approach are generally applicable to the similarity matrices directly, i.e., they do not depend on the proposed functional normalization. The idea of holographic comparison is expected to be very efficient in many applications, like comparing gene-expression data from different populations of organisms or groups of people.

In the next sections each of the above stages in the FH method are illustrated and described in detail. The recorded spontaneous activities of cultured neural networks are employed as a model system to illustrate the power of our new approach in providing a satisfying framework for achieving the goals specified above. We also employ our method on a model neural network activity and show its power in detecting inhibitory neurons.

Some preliminary results of analyzing ECoG recording of human brain activity during seizure (Ictal) and between seizure events (inter-Ictal) are presented. In the discussion, it is proposed that the success of the method in capturing the underlying functional and causal motives might imply that the activity of neural networks is self regulated by an underlying simple, low dimension manifolds in the space of functional correlations.

2. Evaluation of similarity matrices for cultured networks activity

Cultured networks provide relatively simple and well-controlled model systems for investigating long term (weeks), individual neurons activity at different locations by using a multi-electrode array [6,7]. The networks whose activity is analyzed here are spontaneously formed from a dissociated mixture of cortical neurons and glia cells from one-day-old Charles River rats. The cells are homogeneously spread over a lithographically specified area of Poly-D-Lysine for attachment to the recording electrodes. Consequently, the neurons send dendrites and axons to form a wired network. Although this process is self-executed with no externally provided guiding stimulation or chemical cues, a relatively intense dynamical activity is spontaneously generated within several days. The spontaneous activity of cultured networks is marked by the formation of synchronized bursting events (SBEs) - short (~200ms) time windows, during which most of the recorded neurons participate in relatively rapid firing. The SBEs are separated by long intervals (above seconds) of sporadic neuronal firing.

As can be seen from Fig. 1, each SBE has its own internal pattern of neuronal firing. Both the firing rate and the time-series statistical properties can greatly vary from neuron to neuron. While some neurons fire only 1-2 spikes/SBE, others can fire ever 20 spikes/SBE. In addition, the individual neuron activity also varies from SBE to SBE. Hence, to take into account these variations in the firing rate, one can define the averaged density of neuronal firing during the SBE, by averaging of the activity over many synchronized bursting

events. An example of such averaged density representation is shown in Fig. 1. It is also seen that the averaged density reflects fairly well the general structure of the SBE.

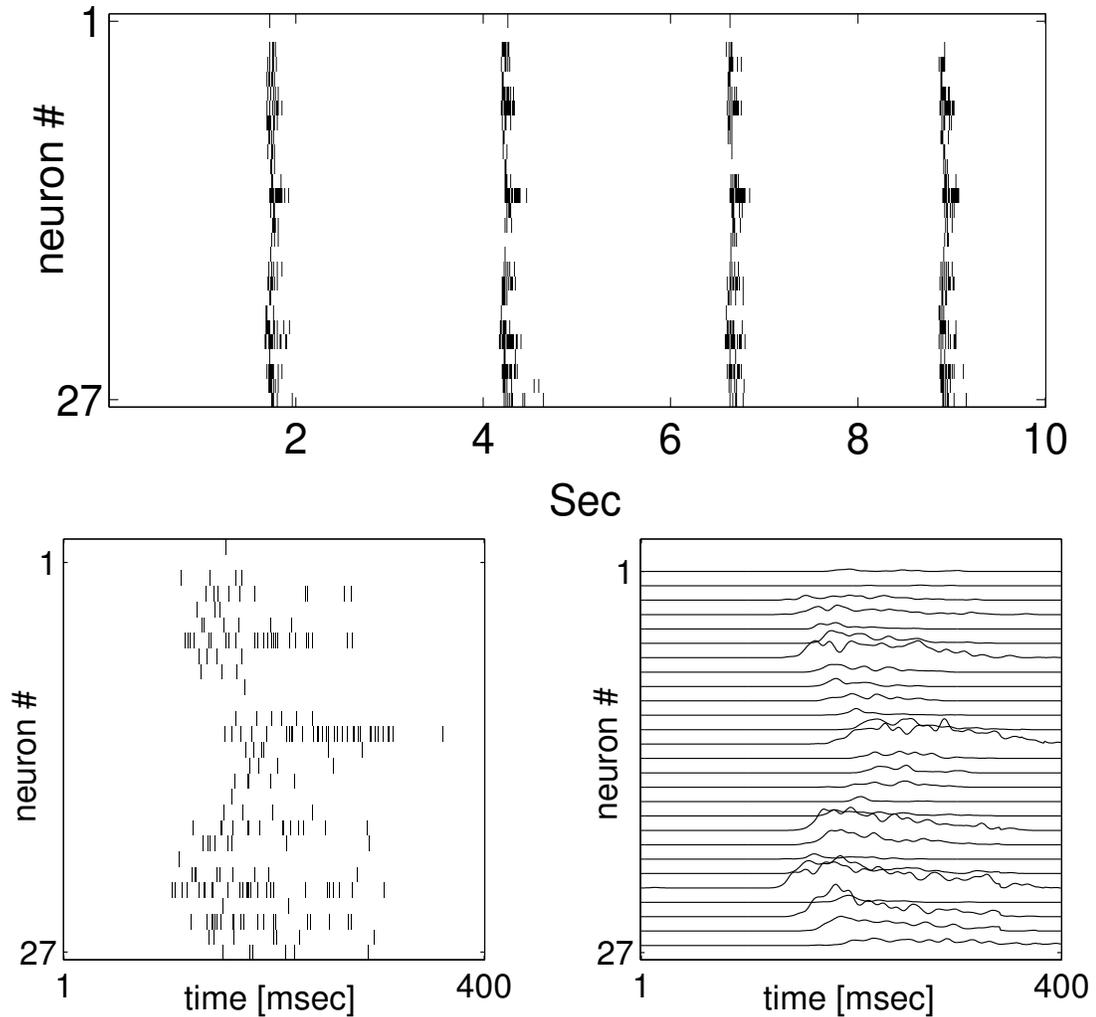


Figure 1: Top: Formation of SBEs in the recorded activity of cultured networks. The time axis is divided into 10^{-2} s bins. Each row is a binary bar-code representation of the activity of an individual neuron, i.e. the bars mark detection of spikes. Note that each neuron has its' own pattern of activity within a given SBE, and each SBE has its' own internal pattern of neuronal activity. Bottom left: zoom into the synchronized bursting event. During the SBE, each neuron has its specific spiking profile. Bottom right: the averaged density representation of a SBE.

To evaluate the similarity matrices each SBE can be described as a matrix composed of N vectors, one for each neuron. The vector $X_i(t)$, represents the temporal activity, or firing rate of neuron i , during the time window of the SBE. See ref [8] for more details.

To evaluate the inter-neuron similarity matrix, the Pearson Correlation coefficient [9] is calculated between the bursts of each pair of neurons according to the standard definition:

$$C_{i,j} = \frac{\langle (X_i(t) - \mu_i)(X_j(t) - \mu_j) \rangle}{\sigma_i \sigma_j} \quad (2)$$

Where X_i and X_j are the activities of neurons i and j with the corresponding means μ_i and μ_j and sample standard deviations (STD) σ_i and σ_j . The correlation coefficients for all pairs of channels are computed, creating the correlation matrix $C_{i,j}$. Note that, by definition, the correlation matrix is symmetric ($C_{i,j} = C_{j,i}$) and its diagonal is 1 ($C_{i,i} = 1$). Next, the similarity matrix is evaluated by averaging the inter-neuron correlations over a sequence of SBEs. That is,

$$S_{i,j} = \langle C_{i,j} \rangle_{SBEs} \quad (3)$$

A typical example of such inter-neuron similarity matrix is shown in Fig 2a. Frequently the correlation matrix is mapped into a connectivity diagram in which the locations of the neurons in real space are connected by colored lines whose colors indicate the level of the correlations (Fig 2b).

To test the self-consistency of the above definition, we evaluate a standard deviation matrix whose $STD_{i,j}$ element is the standard deviation of $C_{i,j}$ over the corresponding sequence of SBEs. The deviation is typically smaller than ~ 0.4 and its average is ~ 0.2 . Moreover, using normality tests like the Lilliefors normality test [10] is obtained. For most pairs of neurons the correlations are distributed normally.

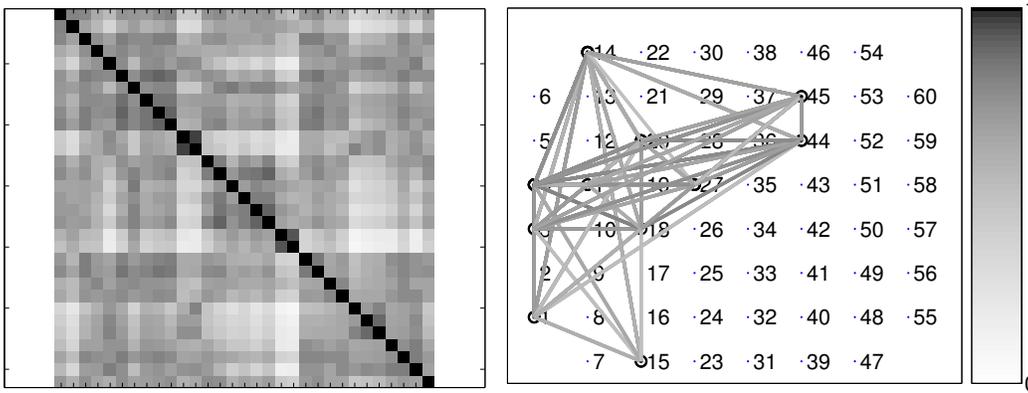


Figure 2: Left: A typical example of inter-neuron similarity matrix. Right: The corresponding standard deviation matrix computed as explained in the text. Left: The connectivity network in physical space – a projection of the inter-neuron correlation matrix gives information about the dependence of correlation on the physical location of the neurons. In the network, each pair of electrodes is connected by lines whose colors indicate the level of the correlation between the two neurons. In the figure we show only lines above some threshold.

3. The affinity transformation - collective normalization

The similarity matrix of N recorded neurons from a cultured network describes an N -dimension space. The position of neuron i in this space is set by the vector \vec{S}_i - its similarities $S_{i,j}$ with all other neurons j . Accordingly, the similarity distance $D_{i,j}$, between

neurons i and j , is simply the Euclidian distance between their positions in the similarity space. That is¹,

$$D_{i,j} = \|\vec{S}_i - \vec{S}_j\| = \sqrt{\sum_{m=1}^N (S_{i,m} - S_{j,m})^2} \quad (4)$$

As was mentioned in the introduction, the similarity distances are used for collective or functional normalization of the similarities. The reason the term “collective” is used has to do with the fact that the similarity distance is a functional of the differences between the similarities of the two neurons with the other neurons. The distance is larger if they have different similarities, and vice versa.

The functional normalization, or affinity transformation, is the evaluation of the new affinity or functional correlation matrix defined as:

$$A_{i,j} = \begin{cases} \frac{S_{i,j}}{D_{i,j}} & i \neq j \\ A_0 & i = j \end{cases} \quad (5)$$

Formally, $D_{i,i}=0$ and hence the diagonal elements $A_{i,i}$ are ill-defined. However, due to the inherent neuronal plasticity and to noise, there should be uncertainty in the auto-correlation and therefore in the neurons positions in the similarity space. The latter can be translated into position uncertainty and thus into finite $A_{i,i}$, which is related to the effective noise level and neuronal plasticity, and thus has to be properly adjusted to the noise level as will be detailed elsewhere. Here, for simplicity, $A_0 \approx I/C_0$, where C_0 is taken as the estimated noise level in the neuron correlations.

Since the similarity distances represent a collective property of all channels, the above procedure can help to capture hidden collective motifs related to functional connectivity in the network behaviors.

4. Dimension reduction

In the previous section it was mentioned that the similarity matrix can be associated with an N-dimension space in which the neuron i is positioned according to its vector \vec{S}_i . Similarly, the affinity matrix is also associated with an N-dimension space of functional correlations. Correspondingly, the position of neuron i in this space is set by a vector \vec{A}_i - the vector of affinities of this neuron with all other neurons. The analysis is, guided by the assumption that most of the relevant information can be obtained in a low dimension space embedded in the larger one. To test this hypothesis we evaluate the PCA eigenvalues of an affinity matrix that has been evaluated for the observed activity of a cultured network as is shown in Fig 3a. The PCA provides a method for identifying new N directions in the N-dimensional space that are ranked (ordered by their eigenvalues) according to the level of standard deviations in the neuron positions [10]. The spectrum shown in Fig. 3b indicates

¹ Note that in ref [4], in the definition of the distances $D(i,j)$ there is division by \sqrt{N} ; both definition are equivalent.

that most of the information in the N-dimensional space of the functional correlations can be captured after dimension reduction to a 3-D space whose axes are the three leading PCA eigenvectors. The position of neuron i in this reduced space (X_i, Y_i, Z_i) is set by the projection of \vec{A}_i on the first three PCA eigenvectors (Fig 3a).

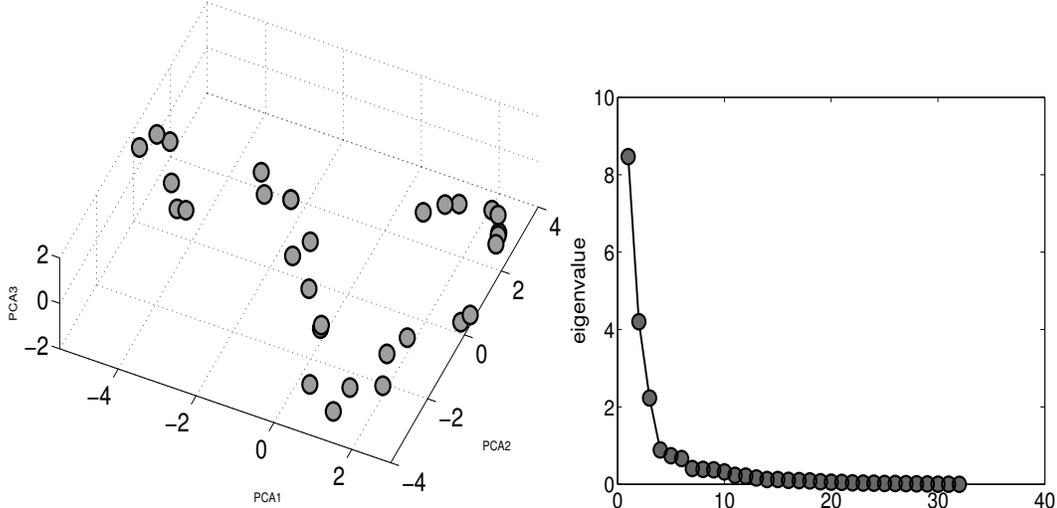


Figure 3: The dimension reduction -projection of the neurons on the 3-D space of the PCA leading eigenvectors of network (I). Left – the projection for the affinity matrix. Right – The PCA eigenvalues of an affinity matrix for cultured neural network activity whose similarity matrix is shown in Fig 2. We will refer to this network activity as (I) in the next figures. In this network the number of recorded neurons is $N=33$. The eigenvalues are for the PCA eigenvectors that are ordered according to the values.

5. From clusters to manifolds

Projection of the matrix elements on a low dimensional space (usually 1, 2 or 3), is common in dimension reduction using clustering algorithms. Since the motivation is to look for sub-groups or clusters in the collection of components, identification of the clusters is the end result. Guided by the assumption that there are underlying functional manifolds, (stage 5 of our method) the neuron positions are next linked by the ‘naked’ (un-normalized) similarities above some threshold, as shown in Figure 4. The idea is to add information about inter- and intra-clusters, information which is lost by the projection of the original matrix on its corresponding low dimension space. The result is a functional connectivity diagram in the abstract space of leading PCA eigenvectors. The term functional is used to emphasize that the diagram includes information about the functional correlations which determines their position in the diagram. This is to be distinguished from the ordinary connectivity diagrams in which the neurons are positioned according to their physical locations in the network. Interestingly, the strong links (higher levels of similarities) form a bundle with simple geometrical organization and simple yet characteristic topology. The results demonstrate that by bringing back some of the information lost in the dimension reduction we find additional features beyond clustering. Keeping in mind that the recorded neurons are just a small sample of the neurons in the network, we suggest interpolating between the neuron positions according to the similarity links to extract the underlying backbone or the functional manifold of the diagram. It is important to stress that the affinity transformation (the functional normalization) is

essential for capturing the hidden manifold - a similar connectivity diagram constructed directly in the 3-D PCA space of the similarity does not show such underlying manifold.

The topology of the manifold of network (II) is that of three perpendicular horseshoes such that the middle one has a joint segment with each of the other two. It is also topologically equivalent to a one turn helix but with a curvature that is not uniform, i.e. it is higher at three specific locations that correspond to the locations of the clusters. It is emphasized that functional manifolds with the same topology, which is simple yet of very specific characteristics, can be identified also in recorded *in vivo* activity, including from the human brain.

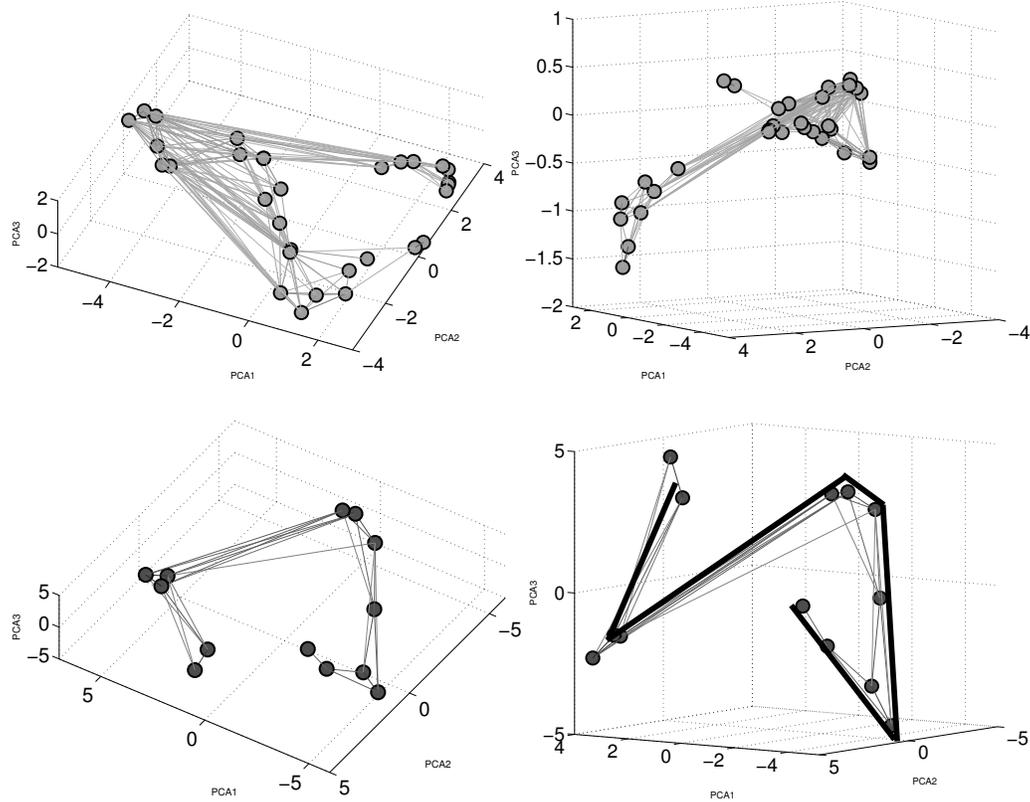


Figure 4: Functional manifolds in the 3-D space of leading PCA eigenvectors. The diagrams at the top show two angles of view of the manifold for network (I) shown in Fig 3. By linking the nodes with similarities (above 0.5) new motifs are revealed, as explained in the text. The most pronounced one is the identification of a quasi 1-D manifold with a specific yet simple topology, as shown in the bottom right figure. The manifold at the bottom left illustrates that this motif is unlikely to be a coincidence as it is exhibited by other networks. The black line is the extracted quasi 1D manifold (backbone), as explained in the text.

6. From manifolds to causal manifolds

As was mentioned in the introduction, the similarity matrices do not include essential information about the networks behavior – the temporal propagation of the activity relative timing between the components. This additional information (when available) is usually presented in temporal ordering matrices whose $T_{i,j}$ element describes the relative timing or

phase difference between the activity of components i and j . There are various methods to evaluate the temporal ordering matrices. In studies of ECoG recorded brain activity, usually it is calculated in terms of ‘phase coherences’ – the relative imaginary parts of the Fourier transform of the activity of channels i and j .

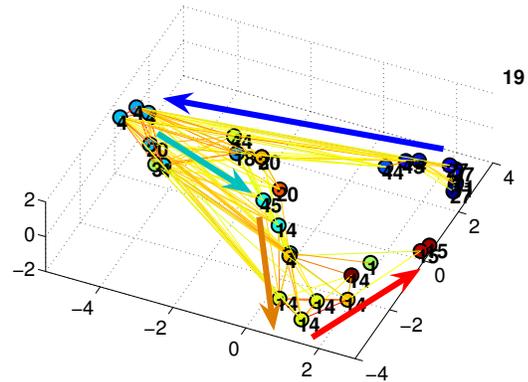
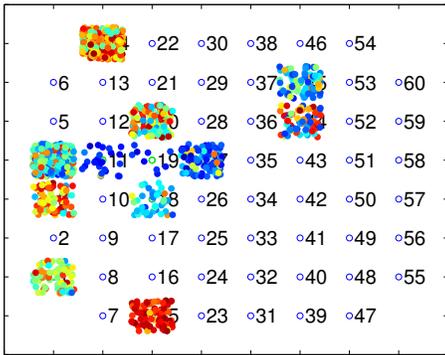
Recently, a new notion – the "temporal center of mass", or temporal location of each neuron in the SBE time window – was proposed [4]. The idea is to regard the activity density of each neuron i as a temporal weight function so that its temporal center of mass, T_i^n , during a SBE (n) is given by:

$$T_i^n = \frac{\int (t - T_n) D_i^n(t - T_n) dt}{\int D_i^n(t - T_n) dt} \quad (6)$$

Where the integral is over the time window of the SBE, and T_n marks the temporal location of the n -th SBE which is the combined "center of mass" of all the neurons. As shown in Fig 5, the temporal center of mass of each neuron can vary between the different SBEs. Therefore we define the relative timing of a neuron i to be $T_i \equiv \langle T_i^n \rangle_n$ – the average of the sequence of SBEs. Similarly, we define the temporal ordering matrix as:

$$T_{i,j} = \langle T_i^n - T_j^n \rangle \quad (7)$$

Looking at the activity propagation in real space (Figure 5), it can be seen that it starts in the center of the network near electrodes (11, 19, 27) and propagates in time in two directions – electrodes (3,4) and (36,37) to terminate on two opposite sides near electrodes 14 and 15. Interestingly, when the temporal information is superimposed on the neurons in the 3-D space of leading PCA eigenvectors, the activity propagates along the manifold in an orderly fashion from one end to the other (Fig 5). For this reason, it is proposed to view the resulted manifold which includes the temporal information as a causal manifold².



² In some cases, the functional connectivity diagrams are more complex, and additional visual perception about the causal features of the activity is obtained by interpolation of a curved surface between the neuron positions, as described in ref [4]. Doing so resulted in a 2-D causal manifold in which the activity propagates mostly along the edges.

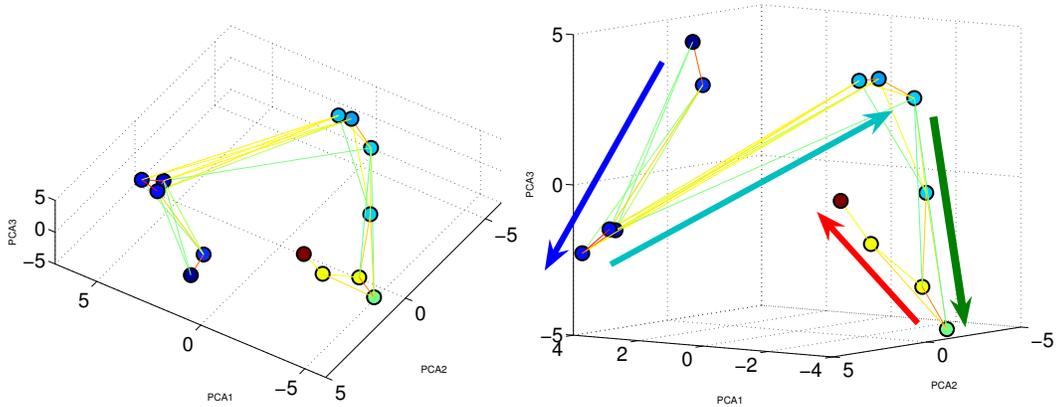


Figure 5: Inclusion of the temporal information. Top Left – illustration of the activity propagation in the physical space of the networks. At each neuron position, we mark by gray level the values of its corresponding T_i^n for all the SBEs in the sequence. Top right – illustration of the neuron timing T_i superimposed on the functional connectivity diagram in the 3-D space of leading PCA eigenvectors. The timing is marked by the gray level at the neuron position. In the bottom right picture we present a schematic illustration of the notion about the activity propagation along the quasi 1D manifold. The bottom left picture shows an additional example of real network (the one shown in the bottom left picture in Fig 5), with arrows that indicate the temporal ordering added to emphasize the directionality. As one can see, there is clear correspondence between the shape of the manifold, the connectivity of the neurons and the propagation of the signal. Bottom – more examples of networks that exhibit causality in the affinity manifold.

7. Functional holography of simulated activity

The generation of synchronized bursting events series, as well as some of its salient statistical properties, is successfully captured by the generic dynamical model of neuronal networks devised by Volman et al. [19]. In this model network, both neurons and synapses connecting them are described as dynamical elements (Appendix A). It has been demonstrated that within this modeling framework, it is possible to recover many salient features of cultured networks activity [19]: The model network spontaneously generates synchronized bursting events, which are separated by long (above seconds) periods of sporadic activity. The internal dynamics of neurons during the SBE are also well- captured by such a model neuronal network with uniform synaptic connectivity. In particular, the inter-neuron correlation matrix, when projected onto a physical space of a model network, yields a network with uniform connectivity. The above construction may also be compared with the actual matrix of synaptic strengths projected onto a physical space of networks' electrodes. For a pair of mutually connected neurons, however, one must consider two distinct cases, as the strength of the connection from i -th to j -th neuron is not necessarily the same as the connection from j -th to i -th neuron. Both of these cases are shown in Figs 6a and 6b, to demonstrate the asymmetry in the neuron-neuron connectivity. The corresponding inter-neuron correlation matrix of the network's simulated activity is shown in Fig 6c and its projection on a connectivity network in Fig 6d.

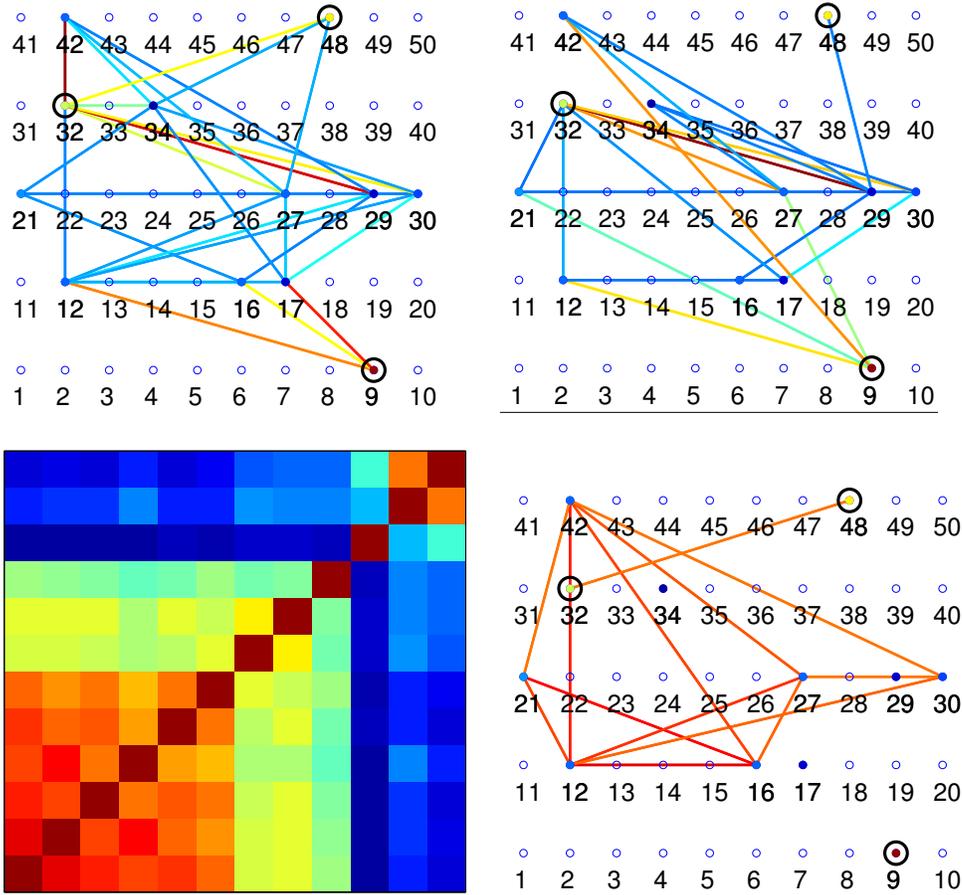


Figure 6. The internal dynamics of neurons during the SBE, for a simulated network with all-to-all synaptic connectivity. We show the results for 12 neurons randomly sampled from a model network. Because pairs of neurons in the model network are reciprocally connected, we show the maps of synaptic connectivity for the cases of a) A_{ij} matrix and b) A_{ji} matrix. c) The inter-neuron correlation matrix for the model unitary network. Note that the structure of the matrix is block-like, with larger block corresponding to the excitatory model neurons, and smaller (upper) block - to the inhibitory ones. d) Physical space projection of the inter-neuron correlation matrix, obtained for the activity of model unitary network.

We apply the function holography to the models activity in the same way we do it for the recorded network activity (described in the previous sections). In figure 7 we show the functional-connectivity diagram of the model activity. As can be seen in the top images, the method can detect the difference between the inhibitory neurons (colored red) and the excitatory ones (colored blue). The inhibitory neurons create a sub-structure which is separated from the large structure of all the other neurons. Also, a causal structure is detected although much less regulated than in the cluttered network activity (bottom images).

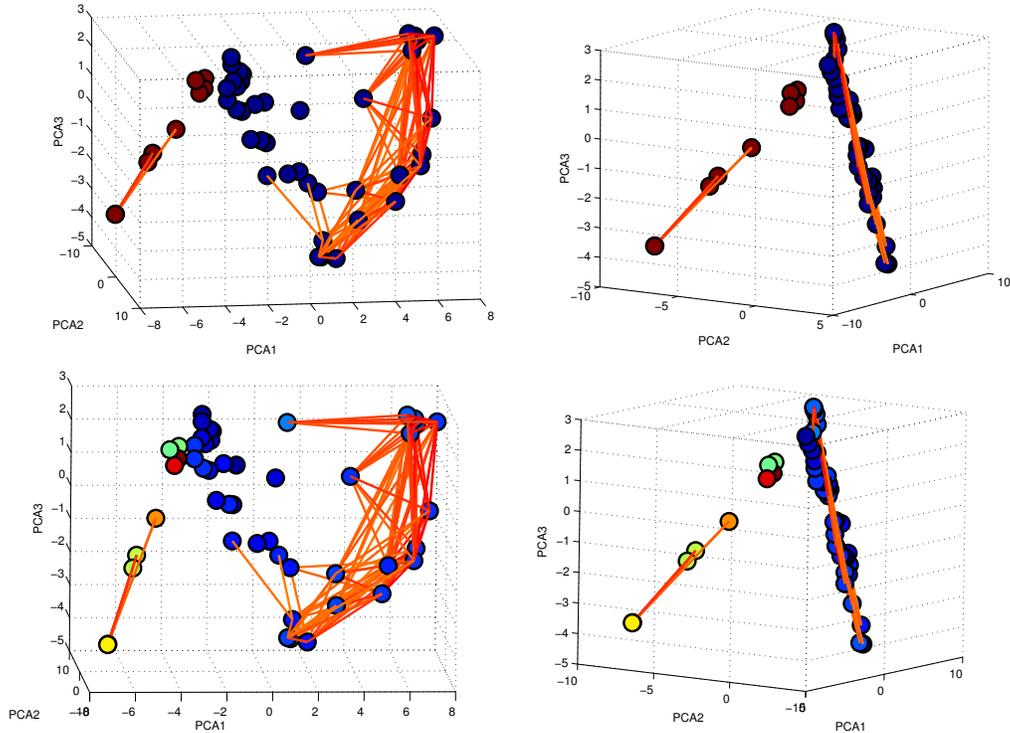


Figure 7: The causal manifold of the uniform network model. Top – two different viewing angles of the models activity manifold. Color represents the type of neuron – where red neurons are inhibitory neurons and blue ones are excitatory. As can be seen – the manifold exhibits the difference between the two neuron types. Bottom – Super imposing the activity timing T_i of each neuron.

8. Functional holography of human brain activity

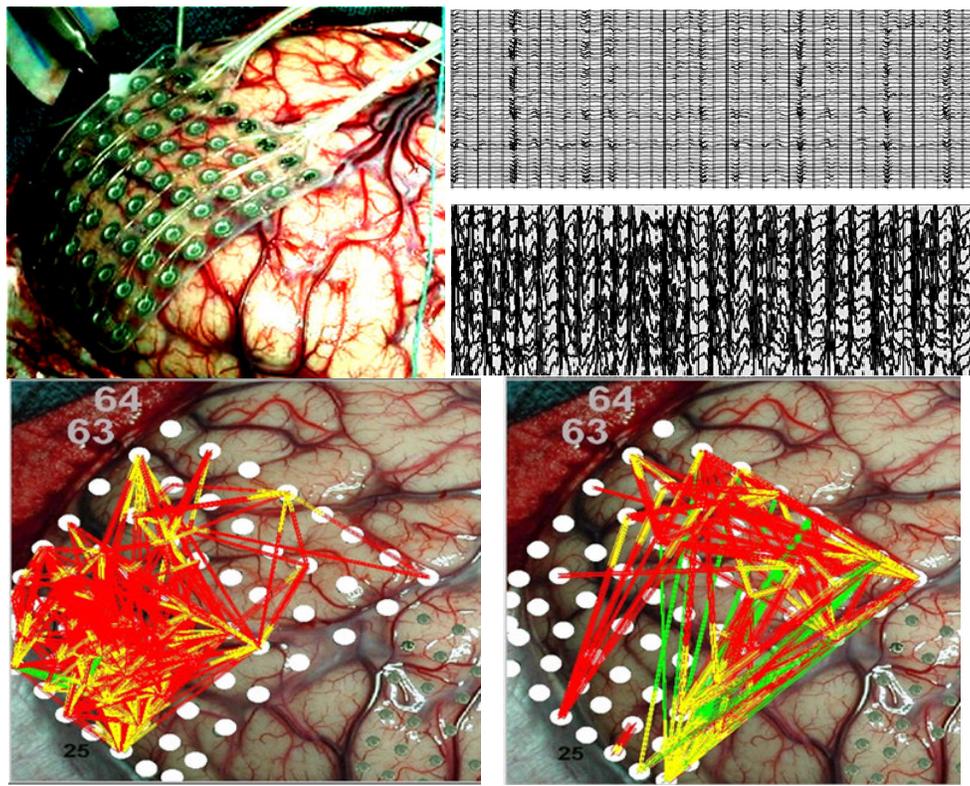
In the above sections, we presented examples of manifolds computed for the recorded activity of cultured neural networks. Despite the artificial nature of these networks, the manifolds exhibit simple yet characteristic geometrical and topological features. In an attempt to rule out that these are mere artifacts generated by the method, the recorded activities of chemically stressed networks were analyzed. The results published in ref [4] are reassuring – the generated manifolds lost the simple characteristic features of the normal activity. In another, more direct test, a network was dissected into two active networks, revealing that the manifold in the affinity space also split into two. An even more convincing feasibility test of the new method described here involves analysis of recorded human brain activity from epileptic patients who are candidates for brain surgery. The method can also be applied to experimental seizure studies that are gained much attention [11].

The occurrence of epilepsy is rising and is estimated to affect, at some level, 1%-2% of the world population [12]. Due to availability of many anti-epileptic drugs, approximately 80% of all epileptic patients can be kept seizure-free. But for the remaining 20%, the only cure is surgical resection of the seizure focus [13,14]. One of the most challenging tasks facing epileptologists is precise identification of brain areas to be removed so that the problem can be cured with minimal damage and side effects. Often, the precise location of the

epileptogenic region remains uncertain after obtaining conventional, non-invasive measurements such as electro-encephalogram (EEG) and magneto-encephalogram (MEG) can not provide sufficient information because of the relatively low spatial resolution of these methods. In these cases, the activity is directly recorded by the electrocorticography (ECoG) procedure in which the recording electrodes are placed directly on the brain surface as is shown in Figure 8. The common approach to analyze the ECoG recording is by evaluation of the coherences between each pairs of electrodes. These coherences (the similarities for this case) are the overlap of the Fourier transform of the recorded voltage. Next a connectivity diagrams are constructed and the similarity matrices are analyzed using clustering algorithm as is illustrated in Figure 8. The idea is to compare the resulted connectivity diagrams (or the similarity matrices), during epileptic seizure (Ictal) and between episodes (Inter-Ictal) to learn more about the cause of the epilepsy. At present, the functional interpretation of these advanced methods is still not clear especially since the outcome matrices and connectivity diagrams appear more complex than the raw data. Hence, much effort is devoted to improve these methods and to the search for new ones.

In view of the above, the idea that functional holography can reveal the existence of hidden causal manifolds embedded within the complexity of the recorded brain activity was tested. And if so, whether it can also provide an efficient tool to distinguish between the inter-Ictal and the Ictal activities. Typical results are presented in Figure 9. The first four pictures show the manifolds for the similarity matrices in Figure 8. The manifold of the inter-Ictal activity has a very simple topology of almost circular horseshoe like part and another bar perpendicular to its plane and position at the center of the horseshoe. During Ictal the quasi 1D property of the manifold gives way to a quasi 2D topology on the surface of a sphere. Albeit the new manifold has more complex topology as could be expected, it retains some of the features of the inter-Ictal one when viewed from specific angle.

This example demonstrates the power of the new method to identify hidden motifs in the complex brain activity. Preliminary analyses also indicate that causal features are captured when the temporal (i.e. phase coherences) information is imposed on the manifolds. These results bear the promise that functional holography might become a valuable epileptogenic diagnostic procedure as well as research approach.



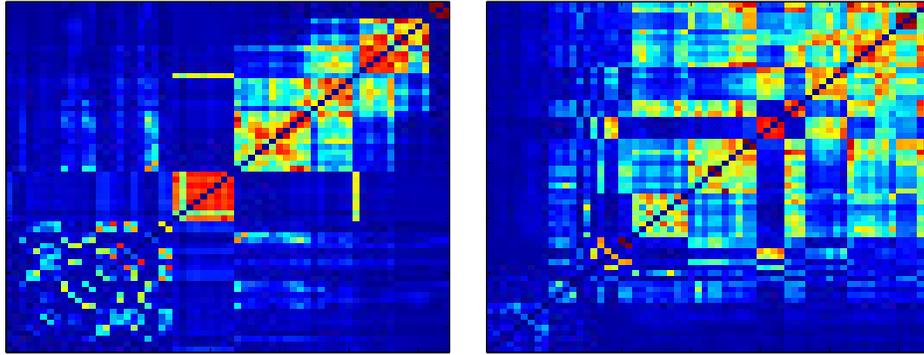
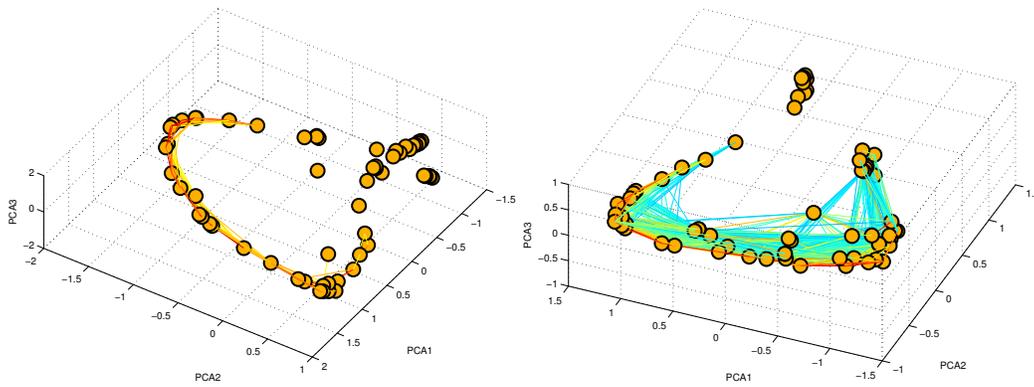


Figure 8: Illustration of the common approach in analyzing ECoG recording of human brain activity. Top left picture shows a set of electrodes placed on the surface of the brain (the frontal lobe in this case). The two pictures in black and white show the voltage recorded from each electrode as function of time. The top one is for inter-Ictal activity and the bottom one is for Ictal (during seizure). The middle two pictures show the connectivity diagram constructed according to the coherences - the color of each link between two electrodes indicate the level of coherence (blue low and red high). The picture on the left is for inter-Ictal activity and the one on the right is for Ictal. The bottom pictures show the corresponding dendrogrammed similarity matrices. Note that in this case the coherences are used as the measure of the similarities. See ref [2] for more details.



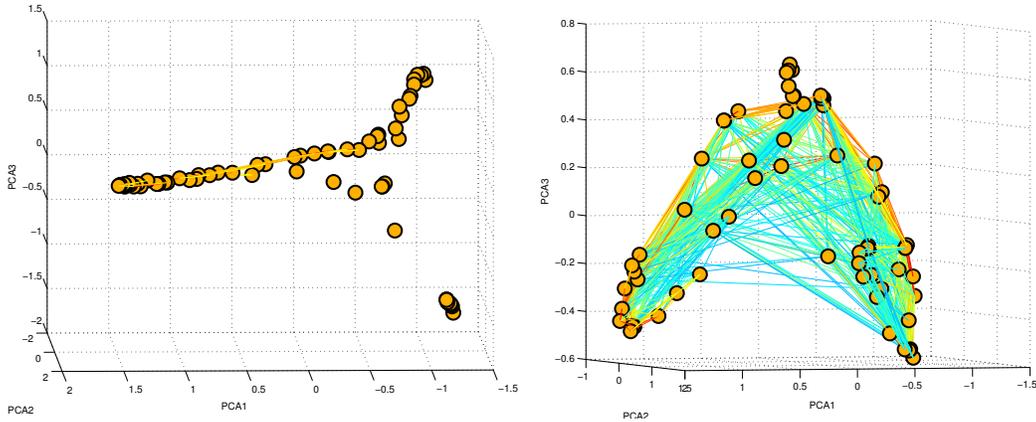


Figure 9: Examples of causal manifolds for ECoG recorded human brain activity. All four pictures are for the inter-Ictal and Ictal activities shown in Fig 8. The ones on the left are for inter-Ictal and those on the right are for Ictal. For each case we show the manifold from two angles of view.

9. Holographic zooming

Often, one is interested in more details about a part of the manifold. Such ‘zooming in’ can be performed but not simply by re-scaling of the axes as done, for example, when a part of a picture is magnified. The idea of the holographic zooming is to take advantage of the collective normalization in the following way: 1. Identify the part of the manifold to be magnified, i.e., identify the cluster of relevant component; 2. Isolate the sub-similarity matrix for the cluster; 3. Perform a second iteration on this matrix, i.e., the affinity transformation, dimension reduction and construction of a manifold. The latter is the magnified part of the manifold.

In Figure 10 another example of ECoG recoding analysis is presented. In this case (a different patient) the recording was from a larger area (a different set of electrodes that covers a larger surface was used). Usually, recording from a larger surface is needed when it is harder to locate the seizure focus or there are multiple foci. The manifold for the larger area recording is composed of two perpendicular manifolds, each with a horseshoe like topology. Using the holographic zooming on one of them reveals that its topology is that of a two perpendicular ‘Siamese horseshoes’ (one joint arm).

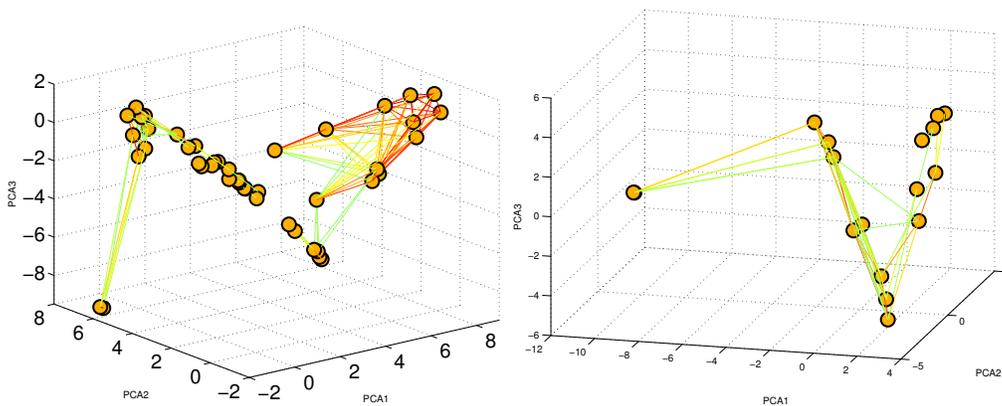


Fig 10. Holographic zooming. The pictures show the holographic zooming for a different set of recordings from a different patient. The recordings in this case were from a larger area of the brain. The similarity in this case was calculated by evaluation of the correlations between the voltage traces from each electrode. Note that the manifold shown on the left picture is composed of two perpendicular horseshoes like manifolds. The picture on the right shows a holographic magnification of the left one. The magnification reveals that this manifold is also composed of two perpendicular horseshoes but with a joint leg.

10. Concluding remarks

A new method for analyzing the complex activity of biological networks is presented, guided by the notion of holograms. In a holographic process, the information describing a 3-D object is encoded on a two-dimensional photographic film, ready to be regenerated into a holographic image or hologram. A characteristic feature is the ‘whole in every part’ nature of the process – a small part of the photographic film can generate the whole picture, but with fewer details. Another property is high tolerance to noise and high robustness to lesion: even with many imperfections or with several pixels removed, the image of the object as a whole is still retained in the hologram. To magnify a part of the original 3-D object, one needs to produce a new photographic film for the part to be magnified. Another related feature is the holographic superposition – when illuminated together, (placed side by side) two holograms can generate a superposition of the corresponding two 3-D objects. Superposition of objects can also be made by imprinting the images of the two (or more) 3-D objects on the same holographic film. These and other special features of hologram are due to the way the information is encoded on the films – not a direct projection of the picture in real space but in the correlations between the pixels. These are converted back to a picture in three dimensions by proper illumination.

The above properties of holograms guided the development and are the rationale behind the functional holography method presented here. The term functional is to indicate that the analysis is in the space of functional correlations that serve the analogue role to the long-range correlations imprinted on the photographic film (by the use of the interference of coherent lights). The methods shown here share the special features of holograms – tolerance to noise, robustness to lesion, holographic superposition and holographic zooming.

As feasibility tests the method ability to capture hidden motifs in complex activity was used to analyze two extreme examples of neuronal networks – the spontaneous activity of cultured neural networks and recorded human brain activity. In both cases the analyses revealed the existence of hidden low dimension manifolds in the high dimension space of functional correlations. The manifolds have surprisingly simple yet characteristic geometrical topological and causal features. Using holographic superposition different modes of the network activity resulted in an entangled manifold composed of a superposition of the individual modes manifolds. Using holographic zooming on recorded brain activity we demonstrated how additional hierarchical motifs can be revealed.

The results presented here aimed to demonstrate the potential of the method as a new valuable procedure for diagnosis of biological networks activity. Such diagnostic procedures are needed for the interpretation of recorded brain activity using EEG, MEG,

ECoG and fMRI [15]. For example, the purpose of ECoG is to localize a suspected seizure focus in the cerebral cortex for patients who are candidates for surgery. The decision of whether to remove or to leave a marginally active area of cortex intact elicits a wide range of opinions from clinicians. Although it is expensive and manpower intensive, ECoG remains the cardinal method for determining the anatomic site of seizure onset, yet even this method of direct recording is not always conclusive. When the seizures involve the medial temporal lobe or are associated with a demonstrable neoplastic process, surgery is successful in about 60-70% of cases. However, when the seizures are suggestive of extra-temporal foci, or no lesion can be identified by MRI and fMRI, the success rate of surgical interventions drops precipitously to less than 50%. This poor cure rate for extra-temporal seizure implies that the current conceptualizations about the disorder are inadequate [16,17,18,27], and more accurate models of epileptic processes and better diagnostic procedures are needed. As was mentioned earlier, functional holography might provide satisfactory solution to this need.

To further test the general applicability of the method in collaboration with other researchers, a preliminary analysis of recorded spinal activity from lamprey was performed, of fMRI measurements of human brain activity and of DNA-microarrays measurements of gene-expression. Again, hidden manifolds with simple geometrical and topological features are discovered. These results give rise to some intriguing questions.

Although can not be ruled out, it is unlikely that they are just accidental. Assuming they are not, they can still be a consequence of some inherent artifact of the method itself. To test that this is not the case we applied the method to analyze the activity of simulated networks whose structure of synaptic connectivity and the nature of neurons (e.g. inhibitory vs. excitatory), are controlled [19,20]. The method can identify for example the inhibitory neurons and the existence of sub-networks when the network is composed of overlapping ones. The efficiency of the method was also tested in comparison between modeled and real networks and found that it can identify additional self-regulation motives in the real ones.

Clearly, additional detailed tests of the method on a variety of systems are needed, but let assume that the discovered hidden causal manifolds are real rather than accidental or arbitrary. The most fundamental question then is why it is so effective. It is proposed that the reason is that the analysis is consistent with the manner in which the biological networks regulate their complex activity. In ref [21] it was shown that even for cultured neural networks the activity is self-regulated to operate at maximal complexity.

Higher complexity is proposed to afford the network with elevated plasticity to perform a wide spectrum of tasks [4,21,22,23]. Motivated by the above it is proposed that the holographic principle for the self-regulation of the networks complex activity. The basic assumptions are: 1. that the networks activity is performed and regulated in the space of functional correlation. For neural networks it implies that the information is processed (encoded decoded and computed) in functional correlations rather than in rates or timing. 2. These high dimension space is regulated by hierarchically organized low dimension manifolds with simple geometry and topology. The hierarchical organization is according to the holographic zooming described earlier. 3. The different modes of behavior and the activity of different sub-networks are mutually regulated by the holographic superposition which keeps entangled yet perpendicular manifolds. 4. Holographic creativity and a continuous space of manifolds. Using a discrete set of photographic films it is possible to generate a continuous spectrum of holograms that are constructed by different

combinations of the photographic films with continuous adjustment of the relative illumination. In an analogous manner the biological networks can create new modes of behaviors from a discrete set of fundamental sub-networks each with its own manifold.

Biological networks do not have photographic films to store holograms nor do they have illuminating lights to for imprinting and regeneration. Hence one can doubt the possible reality of the above holographic principle beyond being just an intellectual metaphor. Indeed, as will be explained in detail, for the principle to be relevant to biological networks the activity must be regulated by at least two complementary mechanisms like the glia regulation in neural networks [24,25,26]. Simply phrased, neuro-glia fabrics provide the 'photographic films' for the holograms which can be imprinted and retrieved by calcium and other chemical waves regulated by the glia cells when act as excitable media.

If correct, the holographic schemata provide the brain with an entirely new ways of coding decoding and processing of information. For example, to sustain associative memory what is needed is the equivalent of superposition of two small portions of the photographic films of each memory. The most attractive is that the holographic schemata provides in principle simple solutions to the fundamental question of creation new images or new meanings of texts. These can be sustaining by the holographic superposition and holographic creativity.

Acknowledgments

The method presented here has evolved from the joint study with Dr. Ronen Segev, Eyal Hulata and Yoash Shapira [7]. Much insight has been gained from analyzing the activity of cultured networks in collaboration with Vladislav Volman and from analyzing the structure of artificial and neural networks in collaboration with Pablo Blinder and Dr. Danny Baranes. One of us (E. B-J.) is most thankful to Prof Steven Schiff for many illuminating conversations, guidance into the foundations and the literature about Epilepsy and constructive comments and advices during the development of this research. We thank Prof. Robert Benzi, Prof. Eytan Domany, Prof. Sir Sam Edwards, Prof. Herbert Levine, Dr. John Hunter and Prof. Itamar Procaccia for constructive comments about the mathematical basis of the method. Preliminary analyses of fMRI measurements are done in collaboration with Dr. talma Handler and Dr. Yaniv Asaf. This research was partially supported by a grant from the Israel Science Foundation, the Maguy-Glass chair in Physics of Complex Systems, and NSF Grant # PHY 99-07949. One of us (E. B-J.) thanks the KITP at University of California Santa Barbara, the Weitzman Institute and the Center for Theoretical and Biological Physics for hospitality during various stages of this research.

Appendix A. The dynamical soma and synapse model

Computer modeling can serve as an equally powerful research tool in the studies of biotic systems [28,29], provided it is utilized and analyzed in proper manners adopted to the special autonomous (regulating) nature of these systems. Guided by the above realization, we have developed a new model in which both the neurons and the synapses connecting them are described as dynamical elements [19]. To model the neurons, we have adopted the Morris-Lecar [30] (M-L) dynamical description. The reasons for this choice are several: (1) in the M-L element, memory can be related to the dynamics of potassium channels; (2) it has been shown by Abbott and Kepler [31] that the M-L equations can be viewed as a reduction (to two variables) of the Hodgkin-Huxley model; (3) The M-L dynamical system has a special phase-space, which can lead to a generation of scale-free temporal behavior in neuronal firing, when fed with a simple noise current.

The simple version of Morris-Lecar model reads:

$$\begin{aligned}\dot{V} &= -I_{ion}(V, W) + I_{ext}(t) \\ \dot{W}(V) &= \phi \frac{W_{\infty}(V) - W(V)}{\tau_w(V)}\end{aligned}$$

with $I_{ion}(V, W)$ representing the contribution of the internal ionic Ca^{2+} , K^+ and leakage currents, with their corresponding channel conductivities g_{Ca} , g_K and g_L being constant

$$I_{ion}(V, W) = g_{Ca}m_{\infty}(V)(V - V_{Ca}) + g_K W(V)(V - V_K) + g_L(V - V_L)$$

The additional current I_{ext} represents all the external current sources stimulating the neuron, such as signals received through its synapses, glia-derived currents, artificial stimulations as well as any noise sources. In the absence of any such stimulation, the fraction of open potassium channels, $W(V)$, relaxes towards its limiting curve $W_{\infty}(V)$, which is described by the sigmoid function. The limiting dynamics of calcium channels are described by $m_{\infty}(V)$.

In our numerical simulations, we have used the following values: $g_{Ca}=1.1\text{mS/cm}^2$, $g_K=2.0\text{mS/cm}^2$, $g_L=0.5\text{mS/cm}^2$, $V_{Ca}=100\text{mV}$, $V_K=-70\text{mV}$, $V_L=-35\text{mV}$, $V_1=-1\text{mV}$, $V_2=15\text{mV}$, $V_3=10\text{mV}$, $V_4=14.5\text{mV}$, $\pi=0.3$. With such a choice of parameters, $I_c=0$.

According to the theory of neuronal group selection, the size of brains' basic functional assembly varies between 50 and 10^4 cells. Motivated by this, and by the notion of unitary networks (as explained in text), we study the dynamics of networks composed of 20-60 cells. To follow physiological data [32], 20% of the cells are usually set inhibitory.

The neurons in the model network exchange action potentials via the multi-state dynamic synapses, as described by Tsodyks et.al. [33]. In this model, the effective synaptic strength evolves according to the following equations:

$$\begin{aligned}\dot{x} &= \frac{z}{\tau_{rec}} - ux\delta(t-t_{sp}) \\ \dot{y} &= -\frac{y}{\tau_{in}} + ux\delta(t-t_{sp}) \\ \dot{z} &= \frac{y}{\tau_{in}} - \frac{z}{\tau_{rec}}\end{aligned}$$

Here, x , y , and z are the fractions of synaptic resources in the recovered, active and inactive states, respectively. The time-series t_{sp} denote the arrival times of pre-synaptic spikes, τ_{in} is the characteristic time of post-synaptic currents (PSCs) decay, and τ_{rec} is the recovery time from synaptic depression.

The variable u describes the effective use of synaptic resources by the incoming spike. For facilitating synapses, it obeys the following dynamic equation:

$$\dot{u} = -\frac{u}{\tau_{facil}} - U_0(1-u)\delta(t-t_{sp})$$

Where the parameter U_0 determines the increase in the value of u with each spike. If no spikes arrive, the facilitation parameter decays to its baseline value with the time constant τ_{facil} . For the depressing synapses (as is the case when post-synaptic neuron is excitatory) one has $\tau_{facil} \Rightarrow 0$, and $u \Rightarrow U_0$ for each spike.

The effective synaptic current of a neuron i is obtained by summing all of its j synaptic currents:

$$I_{syn}^i = \sum_{j \neq i} A_j y_j(t)$$

Where the parameter A_j is the maximal value of synaptic strength.

The values of parameters control the ability of system to exhibit modes of correlated activity. In our studies (unless indicated otherwise), we assigned to the network the parameters specified below, using the following notations: I indicates inhibitory neurons and E - excitatory ones. For example, $\tau_{rec}(E \rightarrow I)$ refers to the recovery time of a synapse transmitting input to an inhibitory neuron from excitatory one. Hence, we set: $\tau_{rec}(I \Rightarrow I)=200\text{ms}$, $\tau_{rec}(E \Rightarrow I)=200\text{ms}$, $\tau_{rec}(I \Rightarrow E)=1200\text{ms}$, $\tau_{rec}(E \Rightarrow E)=1200\text{ms}$, $U_0(I \Rightarrow I)=0.5$, $U_0(E \Rightarrow I)=0.5$, $U_0(I \Rightarrow E)=0.08$, $U_0(E \Rightarrow E)=0.08$, $A(I \Rightarrow I)=9$, $A(E \Rightarrow I)=9$, $A(E \Rightarrow E)=2.2$, $A(I \Rightarrow E)=6.6$. Actual values for each neuron were then generated as reported in [33]. We set $\tau_{in}=6$ ms for all neurons. In addition, due to the small size of our simulated network, we chose $\tau_{facil}=2000$ ms for all inhibitory neurons.

To complete the picture, we need to provide a mechanism responsible for the generation of spontaneous activity in the isolated network. To simulate this, each neuron is subject to the fluctuating additional current

$$I_{ad}(t+1) = I_{ad}(t) + \xi \quad ; \quad \xi = \begin{cases} +\varepsilon, & p = 0.5 \\ -\varepsilon, & p = 0.5 \end{cases}$$

The fluctuating current may drive the neuron beyond the firing threshold, thus enabling it to generate spike and trigger the SBE. To keep a proper balance between the above current and inputs received from other neurons via the synaptic connections, the additional current I_{ad} is limited to the range $I_{low} \leq I_{ad} \leq I_{high}$. The total current seen by a neuron at any time is a sum of $I_{syn}(t)$ and $I_{ad}(t)$.

References:

- [1] Ihmels J.; Levy R.; Barkai N.; Principles of transcriptional control in the metabolic network of *Saccharomyces cerevisiae*. *Nature Biotechnology*, doi:10.1038/nbt918 (2003).
- [2] Vernon L. T.; Syed I.; Bergera C.; Grzesczukb R.; Miltona J.; Ericksonc R. K.; Cogenc P.; Berksona E.; Spirea J.P.; Identification of the sensory/motor area and pathologic regions using ECoG coherence. *Electroencephalography and clinical Neurophysiology* 1998,106: 30–39.
- [3] Brazier M.A.B. *Cluster Analysis*. London: Edward Arnold. 256, 1993.
- [4] Baruchi I.; Ben-Jacob E.; Functional Holography of Recorded Neuronal Networks Activity. *Neuroinformatics*, August 2004, vol. 2, no. 3, pp. 333-352(20).
- [5] Hoffman P.E.; Grinstein G.G. A survey of visualizations for high-dimensional data mining. *Morgan Kaufmann Publishers Inc.*(2002) 47--82.
- [6] Kamioka H.; Maeda E.; Jimbo Y.; Robinson H.P.C.; Kawana A.; Spontaneous periodic synchronized bursting during formation of mature patterns of connections in cortical cultures. *Neuroscience Letters* ,206 109-112, 1996.
- [7] Segev R.; Benveniste M.; Hulata E.; Cohen N.; Palevski A.; Kapon E.; Shapira Y.; and Ben-Jacob E; Long term behavior of lithographically prepared in vitro neuronal networks. *Phys. Rev. Lett.*, 88, 118102, 2002.
- [8] Segev R.; Baruchi I.; Hulata E. Ben-Jacob E.; Hidden Neuronal Correlations in Cultured Networks. *Phys. Rev. Lett.* 92, 118102 (2004).
- [9] Otnes R.K; Enochson L.; *Applied Time Series Analysis*. Wiley – Interscience Publication. 1978.

- [10] Conover, W. J.; Practical Nonparametric Statistics. New York, Wiley. (1980).; Krzanowski W.J.; Principles of Multivariate Analysis – A User’s Perspective. Oxford University Press. 1990.
- [11] Netoff, T. and Schiff, S. J. Decreased neuronal synchronization during experimental seizures > Neuroscience 2002 22, 7297-7307
- [12] Towle V.L.; Ahmad F.; Kohrman M.; Hecox K. and Chkhenkeli S., Electrographic Coherence Patterns of Epileptic Seizures in Epilepsy as a dynamic disease Ed. by P. Jung and J. Milton 2002 Springer, Berlin.
- [13] Doyle, W.K., Spencer, D.D. Anterior temporal resections. In: Epilepsy: A Comprehensive Textbook. J. Engel, Jr), Pedley, T.A. (Eds.), V. 2. Lippincott-Raven, 1998: 1807-1817.
- [14] Chkhenkeli S.A., Towle V.L., Milton J.G., Spire J.-P. Multitarget stereotactic surgery of intractable epilepsy. In: Abstr. of the XIII Congress of ESSFN. Freiburg. Germany. 1998, 4: 21.
- [15] Bandettini P.A.; Moonen C.T.W.; Aguirre G. K. ; Functional MRI. Springer-Verlag Telos; 1st edition (July 15, 1999).
- [16] Bickford R.G. The application of depth electroencephalography in some varieties of epilepsy. Electroencephalography and Clinical Neurophysiology, 1956, 8: 526-527.
- [17] Brazier M.A.B., Evoked responses from the depths of the human brain, Annals of the New York Academy of Sciences, 1964, 112: 33-59.
- [18] Brazier M.A.B.; Spread of seizure discharges in epilepsy: Anatomical and electrophysiological considerations, Experimental Neurology, 1972, 36: 263-272.
- [19] Volman V.; Baruchi I.; Persi E.; Ben-Jacob E.; Generative Modeling of Regulated Activity in Cultured Neuronal Networks. Physica A 335: 249-278 (2004).
- [20] Volman V.; Baruchi I.; Ben-Jacob E.; Self-regulated homoclinic chaos in neural networks activity. 8th Experimental Chaos Conference (in press).
- [21] Hulata E.; Baruchi I.; Segev R. ;Shapira Y.; Ben-Jacob E.; Self-Regulated Complexity in Cultured Neuronal Networks. Phys. Rev. Lett. 92(19):198105(1)-198105(4) (2004).
- [22] Ben-Jacob E.; Bacterial Self-Organization: Co-Enhancement of Complexification and Adaptability in a Dynamic Environment. Phil. Trans. R. Soc. Lond. A, 361:1283-1312, (2003).
- [23] Ayali A. ,Fuchs E. ,Zilberstein Y. ,Shefi O. ,Hulata E. ,Baruchi I. ,Ben Jacob E. Contextual Regularity and Complexity of Neuronal Activity: from Stand-Alone Cultures to Task-Performing Animals. Complexity, Vol 9, Issue 6, 25-32, (July/August 2004) .

- [24] Stevens B, Porta S, Haak LL, Gallo V, Fields RD (2002) Adenosine: A Neuron-Glial Transmitter Promoting Myelination in the CNS in Response to Action Potentials Neuron, 36, 855–868
- [25] P. R. Laming, H. Kimelberg, S. Robinson, A. Salm, N. Hawrylak, C. Müller, B. Roots and K. Ng., (2000) Neuronal-glia Interactions and Behavior Neuroscience and Biobehavioral Reviews 24 295-340.
- [26] Stout C.E., Constantin J.L., Naus C.G., Charles A.C.(2002) Intercellular calcium signaling in astrocytes via ATP release through connexin hemichannels J. Biological Chemistry 277: 10482-10488
- [27] Theoden I. Netoff and Steven J. Schiff. Decreased Neuronal Synchronization during Experimental Seizures. The Journal of Neuroscience, August 15, 2002, 22, 7297–7307.
- [28] E. Ben-Jacob, I. Cohen, and H. Levine. Cooperative self-organization of microorganisms. Adv. Phys., 49(4):395{554, 2001.
- [29] E. Ben-Jacob and H. Levine. Biological pattern formation: examples, issues and strategies. J. Physical Biology, in press, 2004.
- [30] C. Morris and H. Lecar. Voltage oscillations in the barnacle giant muscle fiber. Biophys. Journal, 35:193{213, 1981.
- [31] L.F. Abbott and T. Kepler. Model neurons: from Hodgkin-Huxley to Hopfield. Statistical mechanics of neural networks (Springer-Verlag, Berlin), 1990.
- [32] M. Abeles. Corticonics. Cambridge University Press, 1991.
- [33] M. Tsodyks, A. Uziel, and H. Markram. Synchrony generation in recurrent networks with frequency-dependent synapses. The Journal of Neurosci., 20, 2000.