Functional holography analysis: Simplifying the complexity of dynamical networks

Itay Baruchi, Danny Grossman, Vladislav Volman, and Mark Shein
School of Physics and Astronomy, Tel Aviv University, 69978 Tel Aviv, Israel

John Hunter and Vernon L. Towle
Department of Neurology, MC-2030, University of Chicago Hospitals, Chicago, Illinois 60637

Eshel Ben-Jacoba
School of Physics and Astronomy, Tel Aviv University, 69978 Tel Aviv, Israel
and The Center for Theoretical and Biological Physics, University of California San Diego,
La Jolla, California 92037

(Received 7 October 2005; accepted 3 February 2006; published online 31 March 2006)

We present a novel functional holography (FH) analysis devised to study the dynamics of task-performing dynamical networks. The latter term refers to networks composed of dynamical systems or elements, like gene networks or neural networks. The new approach is based on the realization that task-performing networks follow some underlying principles that are reflected in their activity. Therefore, the analysis is designed to decipher the existence of simple causal motives that are expected to be embedded in the observed complex activity of the networks under study. First we evaluate the matrix of similarities (correlations) between the activities of the network’s components. We then perform collective normalization of the similarities (or affinity transformation) to construct a matrix of functional correlations. Using dimension reduction algorithms on the affinity matrix, the matrix is projected onto a principal three-dimensional space of the leading eigenvectors computed by the algorithm. To retrieve back information that is lost in the dimension reduction, we connect the nodes by colored lines that represent the level of the similarities to construct a holographic network in the principal space. Next we calculate the activity propagation in the network (temporal ordering) using different methods like temporal center of mass and cross correlations. The causal information is superimposed on the holographic network by coloring the nodes locations according to the temporal ordering of their activities. First, we illustrate the analysis for simple, artificially constructed examples. Then we demonstrate that by applying the FH analysis to modeled and real neural networks as well as recorded brain activity, hidden causal manifolds with simple yet characteristic geometrical and topological features are deciphered in the complex activity. The term “functional holography” is used to indicate that the goal of the analysis is to extract the maximum amount of functional information about the dynamical network as a whole unit. © 2006 American Institute of Physics. [DOI: 10.1063/1.2183408]

We propose a new mathematical conceptualization for analyzing the complex activity of information-based, task-performing biological networks [e.g., gene networks, the immune system and the central nervous system (CNS)]. Our approach is guided by the “whole in every part” nature of holograms and the “simple complexity” assumption about the nature of the biological system dynamics—the idea that simple motives are hidden in the observed behavior. The mathematical basis for our approach is a novel functional holography (FH) analysis in which the complex activity of biological networks is unfolded in the space of functional correlations between the activities of the network constituents. Using dimension reduction algorithms, a connectivity diagram is generated in the three-dimensional space that captures the maximal relevant information. Temporal (causal) information is superimposed on the resulted diagram by coloring the elements locations according to the temporal ordering of their activities. By this analysis, the existence of hidden causal manifolds with simple yet characteristic geometrical and topological features in the complex biological activity can be deciphered. 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.

I. INTRODUCTION

The activity of task-performing biological and man-made dynamical networks (networks composed of dynamical elements) is often observed in terms of multichannel recordings of the dynamics of its components. Gene networks and neural networks are two representative examples of task-performing biological networks composed of dynamical elements (genes and neurons, respectively). The activities of
these networks are measured using different observables—the levels of gene expression and the time series of neuronal action potentials, respectively. Yet, the activity in both examples, as in many other networks, is often mapped onto a common abstract presentation—The similarity matrix $S$ or the matrix of correlations. In general, the matrix element $S_{ij}$ is the computed similarity (correlation) between the activities of components $(i)$ and $(j)$ of the network. The similarity can be based on different measures such as correlations, 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 intergene expression correlation, while in the case of recorded brain activity (e.g., EEG, MEG, and ECoG) it is the interchannel coherences.

A common approach in the studies of similarity matrices is to apply clustering algorithms to identify underlying subgroups (clusters) of components that have higher similarity between their activities. Many algorithms have been devised according to the specific systems and the assumed motifs that are looked for. The various clustering algorithms like dendrogram reordering of the matrix and the widely used principal component analysis (PCA) algorithm are based on the implicit notion that the row/column vectors of the similarity matrices span a high dimension space of similarities: $N$-dimensional space for a network composed of $N$ components.

It has also become popular to visualize the similarity matrix by the construction of a corresponding connectivity diagram (connectivity network or correlation network)—a graph whose nodes are the network components and the links between the nodes represent the activity similarities. That is, the link between components $(i)$ and $(j)$ is the element $S_{ij}$. Usually, a color or a gray level code of the links (or simply the lines connecting the nodes) are used to represent the level of the similarities. For hard-wired networks (e.g., neural networks), the diagram is presented as a connectivity diagram in real space—the graph is organized according to the locations of the electrodes in the cultured network or in the brain, respectively.

In many cases, one can also extract information about the activity propagation (temporal ordering) by using various methods such as the cross correlations, the phase coherences (e.g., in recorded brain activity), or the timing between neuron firings in cultured networks. This temporal (causal) information can be represented as a temporal ordering matrix whose $T_{ij}$ element describes the relative timing (time lag or phase lag) 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. Yet, many of the fundamental issues related to the interpretations of the results, or, more specifically, the “reverse engineering” from the observed activity to the underlying functional connectivity, are still to be resolved. The development of the present method has been motivated by the above general goal with the aim to reach the following specific objectives:

1. To compensate for the common limitation incurred by measuring the activity of only a fraction of the network components.
2. To reduce the effect of the inherent noise both in the measurement procedure and in the biological activity itself.
3. To identify underlying simple functional motives in the observed complexity.
4. To relate the observed temporal ordering (activity propagation) to underlying causal motives (propagation of information and causal connectivity).
5. To identify functional subgroups (functional clusters) and to reveal the causal relations between them.

The functional holography analysis approach is a visualization of the network activity in an abstract, reduced three-dimensional space that captures the maximum amount of information about the functional and causal relations between the network’s components. In other words, the analysis is aimed to extract the maximum amount of information about the network as it functions as a whole unit. (Note that the term hologram stands for “whole”—holo in Greek, plus “information” or “message”—gram in Greek. The term was used for the familiar holograms to indicate that the photographic plates can capture the whole information about the 3D image.)

The mathematical procedure of the method involves the following steps:

1. Evaluation of the similarity matrix $S_{ij}$ between components $i$ and $j$.
2. Clustering (by sorting or reordering) of the similarity matrix.
3. Collective normalization of the similarities (or affinity transformation) to construct a matrix of functional correlations.
4. Dimension reduction—projecting the $N$-dimensional matrix of functional correlations onto a three-dimensional space that captures the maximal information.
5. Retrieval of information from the correlation matrix (that is lost in the dimension reduction) onto the 3D principal space. This is done by linking the components locations in the reduced space with lines color coded according to the similarities.
6. Inclusion of temporal (causal) information that describes the activity propagation in the network.
7. Holographic zooming and holographic comparison.

The structure of this chapter is as follows: The functional holography method is described in the next section. The potential applicability of the method to decipher hidden functional and causal motives in the activity of dynamical networks is demonstrated in Sec. III. We analyze the activity of both cultured neuronal networks (Sec. III A) and modeled neural networks (Sec. III B). In Sec. IV we present some preliminary results of applying the method in the analysis of ECoG (subdural EEG) recordings of human brain activity.

Section V is devoted to the description of potential extensions of the method, such as analysis of topological structure (Sec. V A), holographic comparison between two net-
works (Sec. V B), and the concept of holographic zooming (Sec. V C).

In Sec. VI we reflect on the possibility that the success of the method in capturing the underlying functional and causal motives might imply that the activity of biological dynamical networks (neural networks in this case) is self-regulated by underlying simple, low dimensional manifolds in the space of functional correlations.

To make this chapter self contained, we present in Appendix A (Appendixes A–D are included in Ref. 35) concise summary of the widely used methods for evaluation of correlation matrices, the basic dendrogram clustering algorithm for sorting the matrix (into subgroups of highly correlated components) and of dimension reduction onto a lower dimension space using the principal component analysis (PCA) algorithm.

In Appendix B we present two methods for calculating the activity propagation (temporal information). Appendix C is devoted to a more detailed investigation of the mathematical properties of the affinity transformation. The dynamic synapses and soma model of neural networks used in Sec. III B is presented in Appendix D.

II. THE FUNCTIONAL HOLOGRAPHY ANALYSIS METHOD

A. Common analyses of correlation matrices

Various clustering algorithms have been devised for deciphering the existence of subgroups of elements that have higher intercorrelations between the nodes that belong to the same subgroup. The dendrogramed clustering algorithm is one of the most widely used methods for sorting the correlation matrix according to the underlying subgroups (Appendix A). In Fig. 1 we show the analysis of a simple example that is presented in greater details in Appendix A. We generate 25 signals to imitate a multichannel recording of the ac-

FIG. 1. (Color) Correlation matrix of synthetically produced signals. Left (a), synthetic signals that include three groups—the first subgroup of nine signals (signals 7,8,10,11,13,15,17,22,24) was generated by harmonic signals with the same periodicity, a phase shift of about 2π/10 and added noise. The second subgroup (signals 3,5,6,14) is another set of harmonic signals, with a different frequency. The other signals just have pure noise with no correlations. Right top (b), the corresponding similarity matrix—the correlation matrix in this case that was computed using the Pearson’s correlations (Appendix A). Right bottom (c), the sorted correlation matrix using the dendrogramed clustering algorithm (Appendix A). In this matrix the two subgroups form distinct clusters.
tivity of a network of 25 components. The signals [Fig. 1(a)] include two subgroups of periodic signals with higher correlations and a group of random signals. In Fig. 1(b) we show the corresponding correlation matrix computed using the Pearson correlations (Appendix A). Applying the dendrogrammed clustering algorithm on the correlation matrix, the subgroups are deciphered in the resulted sorted (reordered) matrix [Fig. 1(c)].

The dimension reduction algorithms were developed to extract the maximum significant information from the $N$-dimensional correlation matrix (for measurements from $N$ components). Motivated by the assumption that the data contains relatively high level of noise, the idea is to define leading (or principal) vectors in the $N$-dimension space of correlations. Usually, the principal vectors are associated with the directions in the correlation space with maximal variations in the distribution of the locations of the components in this space (Appendix A). Next, the components are projected onto a low dimension space (typically from one to three dimensions) that is defined by the principal eigenvectors of the algorithm. In the projection, each component is located according to its eigenvalues for the principal vectors. In Fig. 2 we illustrate the projection of the correlation matrix shown in Fig. 1 on a principal three-dimension space using the PCA algorithm.

In the functional holography approach presented here, we also use dimension reduction. However, first we perform collective normalization of the correlations (similarities) as is described in the next section. Only then the dimension reduction algorithm is applied on the new matrix of normalized correlations.

B. Collective normalization—the affinity transformation

A new stage in the functional holography analysis is to perform a collective normalization of the correlations and thus compute a new matrix of normalized correlations (referred to as the affinity transformation and the affinity matrix). There have been attempts to go beyond the two point correlations to capture mutual or relative effects between several components. The idea of the collective normalization is to normalize the correlations between every pair of nodes according to their correlations with all the other recorded components. The collective normalization is not uniquely defined operation as it can be performed using different metrics in the correlation space.

The simplest metric is defined by the Euclidian distance in the correlation space. The latter are the Euclidian distances between the locations of the components in the correlation space, that is

$$D_{ij} = \|\mathbf{S}_i - \mathbf{S}_j\| = \sqrt{\sum_{k} (S_{ik} - S_{jk})^2}.$$  

Using this metric, the normalized correlations $A_{ij}$ are given by

$$A_{ij} = \begin{cases} \frac{S_{ij}}{D_{ij}}, & i \neq j, \\ 0, & i = j. \end{cases}$$

Note that when the similarities are defined to be the correlations, by definition $C_{ij}=1$ which implies that the metric must be defined on an $N-1$ surface (an $N-1$ dimension cube). It is reflected also by the fact that formally, $D_{ij}=0$ and hence the diagonal elements $A_{ij}$ are ill-defined.

In Refs. 4 and 6 it was proposed that this difficulty can be overcome by setting the diagonal elements to finite values $A_0$ that correspond to the level of noise. Alternatively, one can replace the equalities $C_{ij}=1$ by the correlations between the activities of the $i$th component at two different time windows. In both cases, for the normalization to be self-consistent, the diagonal terms must be higher than all the off-diagonal ones. (Moreover, from the cases we studied, when off diagonal terms with high values are detected it is usually an indicator of artifact in the recorded data and should be carefully investigated as is further discussed in...
Appendix C. On the practical level, one should investigate the histogram of the correlation distances and impose a short distances cutoff.

The consistency of the normalization can also be guaranteed by using different metrics. For example, one can calculate the metacorrelations $MC_{ij}$, that is, the correlations between the sorted rows of the correlation matrix of the components $(i)$ and $(j)$ as is defined in Eq. (3). To avoid confusion, we mark in the figures the functional correlations that are computed using this metric by AMC$_{ij}$. That is,

$$MC_{ij} = \frac{\sum_{k\neq i,j} (C_{ik} - \mu_{c_i})(C_{jk} - \mu_{c_j})}{\langle \hat{C}_i^2 \rangle \cdot \langle \hat{C}_j^2 \rangle}^{1/2},$$

$$AMC_{ij} = C_{ij} \cdot MC_{ij}.$$  

For this metric AMC$_{ij} = 1$, hence there is no problem with the diagonal elements. This “correlation of correlations” metric is defined on the sorted rows of the matrix such that the elements $C_{ii}, C_{jj}, C_{ij},$ and $C_{ji}$ are taken out. Namely, it is a measure of the similarity between the similarities of each of the components $(i)$ and $(j)$ with all other components but not with each other. Hence, this metric is defined on a lower dimension manifold in the $N$-dimension space of correlations. In Fig. 3 we present the affinity matrices computed for the correlation matrix shown in Fig. 1, using two different metrics. Additional examples are given along this section and in Appendix C.

C. Retrieval of lost information and the holographic network

The next step involves dimension reduction of the affinity matrix and locating the components in the principal 3D space spanned by the three leading eigenvectors of the matrix. To retrieve some of the information lost in the collective normalization (see Appendix C for details), we simply link each pair of nodes by color coded lines according to the original (non-normalized) similarities between them. The retrieved information can be made even more transparent when pairs of nodes that have similarities above (or within a range of) some level are drawn as is shown in Fig. 4. The resulted
connectivity diagram in the abstract space of leading principal eigenvectors can be viewed as a “holographic network.” The term reflects the fact that information from the higher dimension space is retrieved. In Secs. III and IV we will illustrate that functional motives of the networks’ activity are revealed in the holographic network.

As is illustrated in Fig. 4 the holographic network captures essential motifs in the analyzed signals. In the shown example, the nine signals for which the correlation matrix was analyzed have a phase shift of about $2\pi/10$ so that the last signal has almost the same phase as the first one. For this reason the shape formed is a closed shape (closed by the high correlation red lines). The channels reside on a closed circle, with each channel highly correlated to the channel with the smallest phase difference and loosely correlated to more distant channels. Channels with a phase shift will have negative correlation with each other, and should be most distant from each other in correlation space. Since these arguments hold symmetrically for all channels, the simplest way to abide these constraints is a closed circle.

FIG. 5. (Color) Inclusion of temporal information. We show the inclusion of the causal information for the holographic network shown in Figs. 3 and 4. The activity propagation is added by coloring the nodes location according to the relative phases or time lag between them. Blue is for early times (negative phases) and red for late times (positive phases). Note that adding this information helps to reveal the phase shifts imposed in the generation of the signals.

FIG. 6. (Color) The expansion effect of the affinity transformation. Top left (a) is a synthetic correlation matrix that has two groups. Top right (b) is the corresponding affinity matrix. The pictures below (c), (d) show the projection of the matrices on the corresponding 3D space of principal PCA vectors (for the correlation matrix on the left and for the affinity matrix on the right). The red lines show the links with the highest correlations. Note that for both cases the two subgroups form distinct clusters. However, for the affinity matrix we see that each of the clusters is “stretched” in the plane that is perpendicular to the line connecting the two clusters. This effect helps to reveal internal information of each of the subgroups while keeping the clusters apart.
D. Inclusion of temporal information

An additional important step in the functional holography analysis is the inclusion of information about the temporal propagation of the activity (or relative timing between components). This essential information is currently not included when clustering algorithms are used. The temporal information (when available) is usually presented separately as temporal ordering matrices whose $T_{ij}$ element describes the relative timing or phase difference between the activity of pairs of components $i$ and $j$. In Appendix B we describe methods to extract information about the global temporal ordering. One method is to use the time ordering analysis described in Refs. 4 and 6 and another method of phase detection using the cross-correlation functions (Appendix B2). Next, in order to superimpose temporal information about the system onto the holographic network, the nodes are color coded according to the computed temporal ordering (activity propagation) as is illustrated in Fig. 5.

The temporal propagation pattern of the group (in this case the relative phases) can be seen in the diagram by starting with the dark blue (most negative phase) and moving counter-clockwise to the dark red (most positive phase). In this simple example it is easy to see that the temporal information is properly captured and it agrees with the causal motive of the activity propagation along the holographic network.

E. Interpretations—contraction and expansion of features

Since the affinity transformation represents a collective property of all channels, it can help to capture hidden collective motifs related to functional connectivity in the network behaviors. The transformation leads to the amplification of subgroups within the data set and attenuation of intersubgroup correlations and noncorrelated channels as is illustrated in the very simple examples shown in Figs. 6 and 7.
We emphasize that we selected these two very simple examples to make the properties more transparent. Illustration of the properties for a “real case”—45-dimensional correlation matrix constructed from ECoG recorded brain activity (as is explained later) is shown in Appendix C.7

III. ANALYZING THE ACTIVITY OF NEURAL NETWORKS

A. Cultured neural networks

Cultured networks provide relatively simple and well-controlled model systems for investigating long term (weeks), individual neurons activity at different locations by using a multielectrode array.8,9 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 (~200 ms) 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.10

As can be seen from Fig. 8, 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. In addition, the individual neuron activity also

FIG. 8. The recorded activity of cultured neural network. Top (a), 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. Bottom left (b), zoom into the synchronized bursting event. During the SBE, each neuron has its specific spiking profile. Bottom right (c), the averaged spikes density (firing rate) of the neurons during the SBEs.

FIG. 9. (Color) Applying the functional holography analysis on the recorded activity of cultured neural network. Top, the interneuron correlation matrix. Middle, the activity connectivity network in real space. Bottom, the causal manifold. As can be seen the holographic network has a relatively simple geometry and topology. Moreover, the activity propagates along the backbone of the resulted manifold.
organize to generate nonarbitrary wiring and driven currents. That is, averaging the interneuron correlations over a sequence of SBEs. That is, 

\[ S_{ij} = \langle C_{ij} \rangle_{\text{SBEs}}. \]

A typical example of such interneuron correlation matrix and its corresponding activity connectivity network in real space and the corresponding holographic network are shown in Fig. 9. As can be seen in this example, the holographic network has a relatively simple geometry and topology. Moreover, the activity propagates along the backbone of the resulting manifold. These results are surprising considering that the cultured networks are grown from an arbitrary mixture of cells and in the absence of electrical stimulations or chemical cues. Yet, in light of the simple examples presented earlier we propose that the results are not arbitrary but reflect some inherent regulating mechanisms. To further test this idea we analyze in the next section the activity of modeled networks and show that simple manifolds are obtained only for networks with imposed geometry and or structured driving currents.\(^{13,14}\) It implies that the cultured networks self-organize to generate nonarbitrary wiring and driven currents.

### B. Modeled neural networks

We analyze here the activity of artificial neural networks using the generic dynamical model devised by Volman et al.\(^ {14}\) In this model network, both neurons and synapses connecting them are described as dynamical elements (see Appendix D for detailed description of the model). It has been demonstrated that within this modeling framework, it is possible to recover many salient features of cultured networks activity: 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 Fig. 10 we present the ordered matrix of interneuron correlations and its corresponding connectivity network in real space.

Next we construct the corresponding holographic network as shown in Fig. 11. As can be seen the analysis helps to detect the difference between the inhibitory neurons and the excitatory ones. The inhibitory neurons create a substructure which is separated from the large structure of all other neurons. This result follows because in the model network, the activity-dependent strength of inhibitory and excitatory synapses evolves according to different dynamical equations. Namely, the synapses projecting onto inhibitory neurons obey slow facilitating dynamics. This leads to the delayed activation of inhibitory neurons (see also Appendix C for more details). Looking at the pattern of temporal activation, we notice that in the case of model networks with uniform distribution of synaptic contacts (as is the case here) the neurons with similar dynamics are expected to be activated almost simultaneously during the bursting event, due to the isotropic fashion in which the activity spreads through the network. We observe that model neurons fall into two clusters (each cluster activated at a different time), corresponding to excitatory and inhibitory neurons (recall that inhibitory neurons are activated later due to their delayed facilitating synapses). Clearly the holographic network of such modeled network has different features from the ones of the real cultured network.
We proceed to test the effect of structured (nonuniform) wiring and structured driving current on the network activity. For that we constructed a network of a 100 model neurons in the following manner: we position the neurons at the nodes of the square lattice, and assume that the synaptic contacts are established only between the nearest neighbors. [The spatial connectivity function used, was a step function for the radial distance between neurons, \( C(r) = \theta(r - \beta) \), where \( \beta \) was chosen so that the network will have 30\% of the connections active.] Further, we set all the neurons in the networks slightly below the threshold for firing. An exception is made for two neurons located at the farthest corners of the lattice—we make these two cells rapidly firing by setting them above the threshold. The existence of such rapidly firing cells (henceforth termed spikers) is well established for cultured cortical networks, where it has been observed that the activity of spiker neurons persists even in the absence of network activity, indicating that spikers might serve as a trigger of bursting events.

We have earlier shown that the activity of these cells possess features indicative of self-regulated homoclinic chaos (Ref. 20). Here, we utilize our previous findings to investigate the possible interplay between localized source of network’s excitation and the causal spread of signal through the network.

The model network described above generates two kinds of synchronized bursting events, each with distinct well-defined internal structure. The two types of SBEs correspond to the two sites of activity initiation. Furthermore, due to the fact that the connectivity of a network is distance-dependent (only nearest neighbors), the bursting event develops in a well-defined temporal ordering. This fact is clearly seen from Fig. 12(b), where we show the causal manifold for one of the burst types.

Taken together, the results for the model neuronal networks (with either uniform or structured connectivity) demonstrate the potential applicability of functional holography method to detect the spatio-temporal patterns of dynamical network’s activity. When put on the proper grounds, the

\[ \text{FIG. 11. (Color) The FH manifold of the uniform network model. We show the corresponding holographic network for all the 50 neurons in the modeled network for two different angles of view. There is a clear separation into two clusters. The large one composed of the excitatory neurons (colored blue) and the smaller one (colored red) is composed of the inhibitory ones.} \]

\[ \text{FIG. 12. (Color) Analyzing the activity of structured modeled network. (a) The ordered correlation matrix for the structured model network. (b) The FH manifold for the activity of the model network. Note that the holographic network has a relatively simple geometry and topology that resembles that of the real network. Another common feature is that, the activity propagates along the backbone of the resulted manifold.} \]
method can also be useful to recover the synaptic connectivity given the correlations between neurons activities.

IV. ANALYZING ECoG RECORDED HUMAN BRAIN ACTIVITY

A 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 have gained much attention.\textsuperscript{15} The occurrence of epilepsy is rising and is estimated to affect, at some level, 1%–2% of the world population.\textsuperscript{16} Due to availability of many antiepileptic 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.\textsuperscript{17,18} 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, noninvasive measurements such as electroen-

FIG. 13. (Color) Illustration of the common approach in analyzing ECoG recording of human brain activity. Top left (a) picture shows a set of electrodes placed on the surface of the brain (the frontal lob in this case). The two pictures in black and white (b) show the voltage recorded from each electrode as a function of time. The top one in (b) is for inter-Ictal activity and the bottom one is for Ictal (during seizure). The two pictures (c), (d) 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 in (c) is for inter-Ictal activity and the picture in (d) is for Ictal. The pictures (e), (f) show the corresponding dendrogramed similarity matrices. Note that in this case the coherences are used as the measure of the similarities. See Ref. 2 for more details.
cephalogram (EEG) and magnetoencephalogram (MEG) cannot 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 Fig. 13. 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 Fig. 13. 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 methods is still not clear especially since the resulting 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.4,6 Typical results are presented in Fig. 14. Notably, the manifold of the inter-Ictal activity has a very simple topology of almost circular horseshoe like part and another subgroup perpendicular to its plane and position at the center of

FIG. 13. (Continued).

FIG. 14. (Color) Holographic networks of recorded brain activity. The holographic networks are for the ECoG recorded human brain activity for the inter-Ictal and Ictal activities shown in Fig. 13—inter-Ictal (a) and Ictal (b). The pictures show the manifolds from different angles of view. In the analysis we included only electrodes whose correlations with the other electrodes are above noise level. Note, that the locations change their functional role during seizure (Ictal) relative to those during the inter-Ictal durations.
the horseshoe. During the Ictal phase 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 indicates 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.

V. EXTENSIONS OF THE ANALYSIS

A. Topological characterization of the manifolds

In some cases important information can be revealed from analyzing the topological properties of the manifolds. To extract the topological information one has to use interpolation algorithms to define the 2D surface that interpolates between the nodes as is shown in Fig. 15. In this case in which the manifolds are evaluated for gene expression recordings the surface was interpolated when equal weight was assigned to all the nodes. In other cases one can consider to assign different weights to the nodes according to various properties, e.g., their total connectivity to other nodes, the strength of their activity, their known biological roles, their locations in real space, etc.

B. Holographic comparison and superposition of networks activity

Often, clustering algorithms are used for comparison between the activities of different networks, e.g., gene expression in two groups (positive and negative) of patients, or between two modes of behavior of the same network, e.g., during and between epileptic seizures of the same patient.

We propose the following holographic comparison between networks: (1) Compute the PCA leading eigenvectors of the affinity matrix for each network. (2) Project the affinity matrix of each network on the leading eigenvectors of the other one. Clearly, this approach can also be used for comparison between different modes of activity of the same networks, like the above-mentioned case of brain activity in between and during seizure. The holographic comparison can also be used to compare different clusters identified in a given matrix. Once the clusters are identified, the similarity matrix for each is isolated from the combined matrix and the above two stages are applied.

The holographic superposition is designed as another method for comparison between different modes of activity of the same network. The idea is similar to the holographic comparison; only the projection is on the mutual PCA leading eigenvectors. That is, the leading eigenvectors of a combined matrix that includes the different modes. In Fig. 16 we show an example of holographic superposition for cultured networks whose activity is composed of distinct subgroups of SBEs—distinct modes of activity (see Refs. 4, 6, 9, and 14).

C. 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 rescaling 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 subsimilarity 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.

As explained in the discussion, the holographic zooming is directly connected to the question about the proper dimension reduction—the proper number of leading PCA eigenvectors to be examined.

FIG. 15. (Color) The topological characters of the manifolds. We show the interpolating surfaces between the nodes for two cases of inter-gene correlations extracted from Microarrays experiments (the details will be provided elsewhere).
VI. CONCLUDING REMARKS AND LOOKING AHEAD

A new method for analyzing the complex activity of dynamical (and in particular biological) networks is presented, guided by the notion of holograms. In a holographic photography, the information describing a 3D 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 3D 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 3D objects. Superposition of objects can also be made by imprinting the images of the two (or more) 3D 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.

We illustrated the ability of a method to capture hidden motifs in the complex activity of neural networks and in recorded brain activity. In all these cases the analysis revealed the existence of hidden low dimension manifolds in the higher dimensional 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.

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 (Sec. III B) to test that this is not the case. As we showed the method can identify, for example, the inhibitory neurons. In other studies (to be presented elsewhere) we show that it can reveal the existence of subnetworks 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.

The results presented here aimed to demonstrate the potential of the method as a new valuable procedure for diagnosis of the activity of biological networks. Such diagnostic procedures are needed for the interpretation of recorded brain activity using EEG, MEG, ECoG, and fMRI. For example, the purpose of ECoG is to locate the potential seizure foci 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. As was mentioned earlier, functional holography analysis might provide satisfactory solution to this need.

To further test the general applicability of the method we started collaborative preliminary studies in which the functional holography analysis is applied to fMRI measurements of human brain activity and of DNA-microarrays measurements of gene expression. Again, hidden manifolds with simple geometrical and topological features together with characteristic temporal propagations are discovered. These results give rise to some intriguing questions.

Clearly, additional detailed tests of the method on a variety of systems are needed, but let us 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. A far reaching possible reason to be explored in the future 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.

Motivated by the above, we propose the following putative holographic principle for self-regulation of task-performing biological networks: (1) the networks activity is performed and regulated in the space of functional correlations. For neural networks it implies that the information is processed (encoded, decoded and computed) in the functional correlations rather than in rates or timing. (2) This higher dimensional space is regulated by the 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 subnetworks are mutually regulated by the holographic superposition which keeps entangled yet perpendicular manifolds.

One might even speculate that the creativity of a human brain can be afforded by a continuous space of manifolds. The idea is motivated by the following simplified metaphor: 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 illuminations. In a metaphorically 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 form imprinting (coding) and regeneration of the images (decoding). Hence one can doubt the possible reality of the above beyond being just an intellectual metaphor. Indeed, for the above to be of any relevance to biological networks the activity must be regulated by at least two complementary mechanisms like the glia regulation in neural networks. Simply phrased, neuroglia fabrics might provide the photographic films for the holograms which can be imprinted and retrieved by calcium and other chemical waves and stimulations regulated by the glia cells when act as excitable media.

If correct, the holographic schemata provide the brain with 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 of 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 (Ref. 8). Much insight has been gained 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 Professor 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 Professor Robert Benzi, Professor Eytan Domany, Professor Sir Sam Edwards, Professor Herbert Levine, and Professor 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 No. PHY99-07949. One of the authors (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.

35. See EPAPS Document No. E-CHAOEH-16-045601 for Appendices A-D. This document can be reached via a direct link in the online article’s HTML reference section or via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html).