Published in final edited form as:

J Phys A Math Theor. 2016 September 30; 49(39): . doi:10.1088/1751-8113/49/39/395001.

Detecting Spatio-Temporal Modes in Multivariate Data by Entropy Field Decomposition

Lawrence R. Frank^{1,2} and Vitaly L. Galinsky^{1,3}

Lawrence R. Frank: lfrank@ucsd.edu; Vitaly L. Galinsky: vit@ucsd.edu

¹Center for Scientific Computation in Imaging, University of California at San Diego, La Jolla, CA 92037-0854, USA

²Center for Functional MRI, University of California at San Diego, La Jolla, CA 92037-0677, USA

³Department of ECE, University of California, San Diego, La Jolla, CA 92093-0407, USA

Abstract

A new data analysis method that addresses a general problem of detecting spatio-temporal variations in multivariate data is presented. The method utilizes two recent and complimentary general approaches to data analysis, information field theory (IFT) and entropy spectrum pathways (ESP). Both methods reformulate and incorporate Bayesian theory, thus use prior information to uncover underlying structure of the unknown signal. Unification of ESP and IFT creates an approach that is non-Gaussian and non-linear by construction and is found to produce unique spatio-temporal modes of signal behavior that can be ranked according to their significance, from which space-time trajectories of parameter variations can be constructed and quantified. Two brief examples of real world applications of the theory to the analysis of data bearing completely different, unrelated nature, lacking any underlying similarity, are also presented. The first example provides an analysis of resting state functional magnetic resonance imaging (rsfMRI) data that allowed us to create an efficient and accurate computational method for assessing and categorizing brain activity. The second example demonstrates the potential of the method in the application to the analysis of a strong atmospheric storm circulation system during the complicated stage of tornado development and formation using data recorded by a mobile Doppler radar. Reference implementation of the method will be made available as a part of the QUEST toolkit that is currently under development at the Center for Scientific Computation in Imaging.

1. Introduction

In a wide range of scientific disciplines experimenters are faced with the problem of discerning spatial-temporal patterns in data acquired from exceedingly complex physical systems in order to characterize the structure and dynamics of the observed system. Such is the case in the two fields of research that motivate the present work: functional magnetic resonance imaging (fMRI) and mobile Doppler radar (MDR). The multiple time-resolved volumes of data acquired in these methods are from exceedingly complex and non-linear systems: the human brain and severe thunderstorms. Moreover, the instrumentation in both

these fields continues to improve dramatically, facilitating increasingly more detailed data of spatial-temporal fluctuations of the working brain and of complex organized coherent storm scale structures such as tornadoes. MRI scanners are now capable of obtaining sub-second time resolved whole brain volumes sensitive to temporal fluctuations in activity with highly localized brain regions identified with specific modes of brain activity ([1]), such as activity in the insular cortex, which is associated with the so-called *executive control* network of the human brain (e.g., [2]). Mobile dual Doppler systems can resolve wind velocity and reflectivity within a tornadic supercell with temporal resolution of just a few minutes and observe highly localized dynamical features such as secondary rear flank downdrafts [3, 4, 5]. These are just two examples of the many physical systems of interest to scientists that are highly nonlinear and non-Gaussian, and in which detecting, characterizing, and quantitating the observed patterns and relating them to the system dynamics poses a significant data analysis challenge.

A variety of methods have been developed to analyze spatiotemporal data including random fields [6], principal components analysis (PCA) [7, 8], independent components analysis (ICA) [9, 10], and a host of variations on classical spectral methods [11]. The various methods have been applied in a wide range of disciplines, such as climatology [12, 13], severe weather meteorology [14], traffic [15], agriculture [16], EEG [17, 18, 19, 20], and FMRI [21, 22, 23, 24, 25]. And with the explosion of data now available from the internet, spatiotemporal analysis methods play an increasingly important role in social analytics [26], as anticipated by early social scientists [27]. However, despite the great variety of methods that have been developed, they generally suffer from significant limitations because they adopt (sometimes explicitly but often implicitly) *ad hoc* procedures predicated on characteristics of the data that are often not true, such as Gaussianity and linearity, or unsupported, such as the independence of the signal sources. Other procedures based on very specific models of dynamical systems (such as least-squared with the assumption that a system is near a critical point [28, 29]) lack the generality that enables their use in other applications. These deficiencies become more acute as the capabilities of the instrumentation increases and the measurements become more sensitive to complex and subtle spatio-temporal variations in the observed physical systems.

The goal of the current paper is to develop a general theoretical framework and a computational implementation, based upon probability theory (which we consider to be synonymous with the term Bayesian probability theory [30]), for the analysis of spatiotemporal signal fluctuations in time-resolved noisy volumetric data acquired from non-linear and non-Gaussian systems. In addition to the practical concerns of providing a method for analyzing our own data, the overarching goal is to provide a general framework that extends the utility of the method to a broad range of problems. An important aspect of our approach is the development of a theoretically clear and computationally sound method for the incorporation of prior information. While the role of prior information is explicit in probability theory in a very general way, its explicit practical implementation in any particular application is often not so clear, particularly in non-linear and non-Gaussian systems which often do not admit analytical solutions and thus require a logically consistent approach that facilitates well-defined approximation methods.

Recently, a reformulation of probability theory in terms of the language of field theory, called *information field theory (IFT)* [31], has provided a rigorous formalism for the application of probability theory to data analysis that is at once intuitively appealing and yet facilitates arbitrarily complex but well defined computational implementations. The IFT framework has the appealing feature that it makes explicit the important but often overlooked conditions that ensure the continuity of underlying parameter spaces (fields) that are to be estimated from discrete data. Moreover, employing the language of field theory has the important consequence of facilitating the efficient characterization of complex, non-linear interactions. And, as a Bayesian theory, it provides a natural mechanism for the incorporation of prior information.

However, while IFT provides a general and powerful framework for characterizing complicated multivariate data, many observed physical systems possess such a high parameter dimensionality that the estimation of subtle coherent phenomena becomes a problem too ill-posed to be solved, even approximately, without some constraints being imposed to reduce the parameter space of the problem. Unfortunately, constraints imposed for this purpose are almost universally *ad hoc* and either explicitly or implicitly make strong assumptions about the structure of the data (e.g., smoothness of some parameter space) and the noise (e.g. independent, Gaussian, etc.). However, in a wide range of complex physical phenomena, non-linear and non-Gaussian processes are common, so that imposing such constraints can obscure the detection of the most interesting physical processes. This approach is particularly problematic in the many physical systems (e.g., the human brain, severe thunderstorms, etc) which are highly non-linear and possess strongly coupled dynamics.

The critical question, then, is whether there is a method by which the data itself can be used to inform the analysis procedure on the existence of compact regions of the parameter space that encapsulate the bulk of the information about the entire system. In other words, is there a way to ascertain if certain configurations of a system are more probable than others? In this paper we demonstrate such a method by incorporating coupling information directly from the data to generate the most probable configurations of a system's parameter space via the theory of *entropy spectrum pathways (ESP)* [32]. The peculiar localization phenomena of the ESP derived probabilities results in a limited spectrum of system configurations that can be quantified and ranked. This method, which we call the *entropy field decomposition (EFD)*, is non-linear and non-Gaussian by construction, and allows the efficient characterization and ranking of configurations or "modes" within complex multivariate spatio-temporal data.

As a demonstration of the theory, we present some initial results from the analysis of resting state functional magnetic resonance imaging (rsfMRI) data [33, 34, 35, 36], a technique in which the brain volume is observed over a period of time using MRI without the administration of cognitive task (as in standard fMRI experiments) and is thus observed in the so-called *resting state*. We demonstrate that the most modest and conservative declaration of prior information is sufficient to formulate the rsfMRI data question in a precise theoretical form that lends itself to an efficient and accurate computational method for assessing and categorizing brain activity.

For a second demonstration of our theory on a strikingly different type of data, we analyze mobile Doppler radar data collected during the formation of the Goshen County (Wyoming) tornado by the Doppler On Wheels (DOW [37]) team on June 5th, 2009 during the second Verification of the Origins of Rotation in Tornadoes Experiment (VORTEX2; [38, 39]). Our method is shown to provide a simple and efficient procedure to identify, quantitate, and track several important spatio-temporal features in a variety of parameters known to be pertinent for tornado development.

2. Probabilistic perspective and Information Field Theory

Consider the case where the data $\{d_{i,j}\}$ consist of n measurements in time t_i , $i = 1, \dots, n$ at N spatial locations x_j , $j = 1, \dots, N$ (or equivalently $\{d_l\}$, where ξ_l , $l = 1, \dots, nN$ defines a set of space-time locations). The spatial locations are assumed to be arranged on a Cartesian grid and the sampling times are assumed to be equally spaced. Neither of these is required in what follows, but merely simplify and clarify the analysis. For most applications we are interested in, the data d are assumed to be 4-dimensional, composed of temporal variations in the signal from a volumetric (three spatial dimensions) imaging experiment. Each data point is of the form

$$d_{i,j} = R s_{i,j} + e_{i,j} \quad (1)$$

where R is an operator that represents the response of the measurement system to a signal $s_{i,j}$ and $e_{i,j}$ is the noise with the covariance matrix $\Sigma_e = \langle e e^\dagger \rangle$ where $\dagger$ means the complex conjugate transpose.

From the Bayesian viewpoint, the goal is to estimate the unknown signal from the peak in the joint posterior probability, given the data d and any available prior information I . The posterior distribution can be written via Bayes theorem, as

$$\underbrace{p(s|d, I)}_{\text{Posterior}} = \frac{\overbrace{p(d, s|I)}^{\text{Joint probability}}}{\underbrace{p(d|I)}_{\text{Evidence}}} = \frac{\overbrace{p(d|s, I)}^{\text{Likelihood}} \overbrace{p(s|I)}^{\text{Prior}}}{p(d|I)} \quad (2)$$

For the case of a known model, the denominator is a constant and the posterior distribution is just the product of the likelihood and the prior distribution. With a non-informative, or “flat” prior, the posterior distribution is just the likelihood, and thus the peak in the posterior distribution is equivalent to the maximum of the likelihood [40, 30]. Thus maximum likelihood methods implicitly assume that 1) The model is correct and 2) There is no prior information.

Information field theory [31] re-expresses the probabilistic (or Bayesian) prescription in the language of field theory. The critical point to be stressed in our interpretation of Eqn 1 is that, although the data consist of discrete samples in both space and time, the underlying

signal s_I is assumed to be continuous in space-time, and thus characterized by a field $\psi(x, t) \equiv \psi(\xi)$ such that $s_I = \int \psi(\xi) \delta(\xi - \xi_I) d\xi$.

This characterization is particularly important in the present context because we seek to not only detect, but in effect, define quantitatively, what is meant by “modes” of space-time variations. As we will show, this can be accomplished well within the context of field theory because the general notion of spatio-temporal patterns can be codified as spatio-temporal modes of a self-interacting field.

The IFT formalism proceeds by identifying the terms in Eqn 2 with the corresponding structures of field theory:

$$H(d, \psi) = -\ln p(d, \psi | I) \quad \text{Hamiltonian} \quad (3a)$$

$$p(d | I) = \int d\psi e^{-H(d, \psi)} = Z(d) \quad \text{Partition function} \quad (3b)$$

so that the Bayes Theorem providing the posterior distribution Eqn 2 becomes

$$p(\psi | d, I) = \frac{e^{-H(d, \psi)}}{Z(d)} \quad (4a)$$

From the general results of [31] applied to a signal of the form Eqn 1, the information Hamiltonian can be written

$$H(d, \psi) = H_0 - j^\dagger \psi + \frac{1}{2} \psi^\dagger D^{-1} \psi + H_i(d, \psi) \quad (5)$$

where H_0 is essentially a normalizing constant that can be ignored, D is an information propagator, j is an information source, and H_i is an interaction term.

This formulation facilitates the identification of the standard statistical physics result [41] that the partition function (Eqn 3b) is the moment generating function from which the correlation functions (also called the connected components [42]) can be calculated

$$G_{ij \dots m}^c = \langle s_1 \dots s_n \rangle_c = \left. \frac{\partial^n \ln Z[j]}{\partial j_i \dots \partial j_n} \right|_{j=0} \quad (6)$$

with subscript c (for *connected*) a standard shorthand for the correlations, (e.g., $\langle a \rangle_c = \langle a \rangle$, $\langle ab \rangle_c = \langle ab \rangle - \langle a \rangle \langle b \rangle$ and so on for higher correlations, where brackets denote expectation values). The moments are calculated from an expression identical to Eqn 6 with $\ln Z$ replaced by Z .

If $H_i = 0$, Eqn 5 describes a *free* theory, whereas if $H_i \neq 0$, then it describes an *interacting* theory. The free theory provides only an initial step in the analysis of data that possesses spatio-temporal variations, for it implicitly assumes that the field components do not interact with one another. Yet this is the effectively the basis for the majority of analysis techniques. However, in most “real life” cases, such as in brain activity or in weather data, one would expect more complex spatio-temporal dynamics in which the modes interact with one another and thus can characterize more complex non-linear and non-Gaussian processes.

Interactions are incorporated into IFT by the inclusion of an interaction Hamiltonian [31]

$$H_i = \sum_{n=1}^{\infty} \frac{1}{n!} \int \cdots \int \Lambda_{\xi_1 \cdots \xi_n}^{(n)} \psi(\xi_1) \cdots \psi(\xi_n) d\xi_1 \cdots d\xi_n \quad (7)$$

We keep the terms with $n = 1$ and $n = 2$ assuming that they can be regarded as perturbative corrections to the source and the propagator terms. This interaction Hamiltonian includes anharmonic terms resulting from interactions between the eigenmodes of the free Hamiltonian and may be used to describe non-Gaussian signal or noise, a nonlinear response (i.e. mode-mode interaction) or a signal dependent noise (i.e. due to mode-noise interaction).

The classical solution at the minimum of the Hamiltonian ($\delta H / \delta \psi = 0$) is

$$\psi(\xi) = D \left(j - \sum_{n=1}^{\infty} \frac{1}{n!} \cdot \int \cdots \int \Lambda_{\xi_1 \cdots \xi_n}^{(n+1)} \psi(\xi_1) \cdots \psi(\xi_n) d\xi_1 \cdots d\xi_n \right) \quad (8)$$

To make the connection with standard signal analysis methods, consider the special case where R in Eqn 1 represents signal model functions F and ψ are the model function amplitudes a and the signal is contaminated by zero mean Gaussian noise with variance $\Sigma_e = \sigma^2$. In the absence of interactions (i.e., the free theory) the expected value of the signal and its covariance (in the Gaussian model all higher order correlations vanish) then give, from Eqn 6, the estimates of the amplitudes and their covariances:

$$\langle a \rangle_c = D_j \quad (9)$$

$$\langle aa^\dagger \rangle_c = D \quad (10)$$

where the propagator is then just the noise-weighted covariance of the sampled model functions (sometimes called the “dirty beam” [43])

$$D = \sigma^2 \left(F^\dagger F \right)^{-1} \quad (11)$$

and the source is noise weighted projection of the signal onto the sampled model functions (sometimes called the “dirty map” [43]):

$$j = \sigma^{-2} F^\dagger d \quad (12)$$

and so the amplitude estimates are, from Eqn 9,

$$\langle a \rangle = F^+ d \quad (13)$$

$$F^+ = (F^\dagger F)^{-1} F^\dagger \quad \text{pseudo-inverse} \quad (14)$$

which is just the standard maximum *a posteriori* result that the amplitudes are found from the pseudo-inverse of the model functions times the data [30]. For example, if the signal model functions F are the Fourier basis functions, then the source is just the noise weighted Fourier transform of the data while the propagator is the covariance of the sampled Fourier model functions. The rationale for the names source and propagator becomes clear. The input data $d(t_i)$ projected along the k th component of the model function $F(t_i)$ provides the source of new information, which is then propagated by $(F^\dagger F)^{-1}_{kl}$ from which estimate $\hat{a}_k$ of the k th amplitude is derived.

The solution to Eqn 13 is found by computing the eigenvectors of the pseudo-inverse F^+ and projecting the data d along each of these eigenvectors. The relative contribution of these projections is determined by the eigenvalues associated with each eigenvector. The terminology used to describe this process is that data is being represented in terms of the *eigenmodes* of the pseudo-inverse.

The purpose of this example is to demonstrate how the general approach relates to standard statistical physics and reduces to well-known results from standard probabilistic methods in the case of very simple signal and noise models. However, we want to emphasize that our interest is in systems with much more complicated signal and noise characteristics. In particular, we are interested in FMRI signals from the resting state of the human brain and mobile Doppler radar data from tornadic storms, both of which involve highly non-linear and non-Gaussian processes. We will demonstrate that such systems can be characterized using the same logical framework, although the identification and description of simple source and propagator terms is no longer possible.

3. Description of Interaction Hamiltonian by Entropy Spectrum Pathways

The IFT approach outlined in the previous section can only be applied to a data analysis problem when (or if) an approximation that describes the nature of the interactions is already known and can be expressed in concise mathematical form. Practically it means that $\Lambda_{\xi_1 \dots \xi_n}^{(n)}$

terms in the interaction Hamiltonian Eqn 7 are known for at least one or several orders of interactions n . In this section we show how coupling information extracted from the data itself can be used to deduce or constrain the nonlinear anharmonic terms of the interaction Hamiltonian, thus providing an effective data analysis approach free of usual linearity and Gaussianity assumptions.

The idea that coupling information between different spatio-temporal points can provide powerful prior information has been formalized in the theory of entropy spectrum pathways (ESP), [32], which is based on extension of the maximum entropy random walk [44, 45, 46]. We will briefly summarize this approach and show how it can be used to obtain nonlinear an-harmonic terms of the interaction Hamiltonian (for details of ESP the reader is directed to [32]). The power of this concept is that one can generate this path entropy spectrum given any coupling information between neighboring locations, and thus this method can be used to turn coupling information into a quantitative measure of prior information about spatio-temporal configurations. And the concept of locations is a very general one. In the current problem, for example, we will be interested in the paths between two space-time locations $(x_a, t_a) \equiv \xi_a$ and $(x_b, t_b) \equiv \xi_b$ described by a continuous field $\psi(\xi_a)$ and $\psi(\xi_b)$.

The entropy field decomposition (EFD), which is the incorporation of ESP into IFT, is found to produce spatio-temporal modes of signal behavior that can be ranked according to their significance. The EFD can be summarized as follows: (i) Nearest-neighbor coupling information constructed from the data generates, via ESP, correlation structures ranked *a priori* by probability, (ii) non-Gaussian correlated structures of such shapes are expected in the signal, without the need for a detailed signal model, (iii) a phenomenological interaction information Hamiltonian is constructed from the ESP modes, and the coupling coefficients are computed up to an order determined from the significance of the different ESP modes, (iv) this then defines a maximum *a posteriori* (MAP) a signal estimator specifically constructed to recover the nonlinear coherent structures.

The ESP theory ranks the optimal paths within a disordered lattice according to their path entropy. This is accomplished by constructing a matrix that characterizes the interactions between locations i and j on the lattice called the *coupling matrix*:

$$Q_{ij} = e^{-\gamma_{ij}} \quad (15)$$

The γ_{ij} are Lagrange multipliers that define the interactions and can be seen as local potentials that depend on some function of the space-time locations ξ_i and ξ_j on the lattice. The eigenvector $\phi^{(k)}$ associated with the k 'th eigenvalue λ_k of Q

$$\sum_j Q_{ij} \phi_j^{(k)} = \lambda_k \phi_i^{(k)} \quad (16)$$

generates the transition probability from location j to location i of the k 'th path

$$p_{ijk} = \frac{Q_{ij} \phi_i^{(k)}}{\lambda_k \phi_j^{(k)}} \quad (17)$$

For each transition matrix Eqn 17 there is a unique stationary distribution associated with each path k :

$$\mu^{(k)} = [\phi^{(k)}]^2 \quad (18)$$

that satisfies

$$\mu_i^{(k)} = \sum_j \mu_j^{(k)} p_{ijk} \quad (19)$$

where $\mu^{(1)}$, associated with the largest eigenvalue λ_1 , corresponds to the maximum entropy stationary distribution [47, 48]. Note that Eqn 19 is written to emphasize that the squaring operation is performed on a pixel-wise basis. Considering only $\mu^{(1)}$, note that if the Lagrange multipliers take the form

$$\gamma_{ij} = \begin{cases} 0 & \text{connected} \\ \infty & \text{not connected} \end{cases} \Rightarrow Q_{ij} = e^{-\gamma_{ij}} = \begin{cases} 1 \\ 0 \end{cases} \quad (20)$$

then Q becomes simply an adjacency matrix A . The maximum entropy distribution constructed from this adjacency matrix is the maximum entropy random walk [44]. Thus it is the coupling matrix Q , rather than the adjacency matrix A , that is the fundamental quantity that encodes the coupling information [32]. Another major significant result of the ESP theory to the present problem is that it ranks multiple paths, and these paths can be constructed from arbitrary coupling schemes through Q_{ij} . The ESP prior can be incorporated into the estimation scheme by using the coupling matrix Q_{ij} Eqn 15 so that

$$p(s|d, I) = \frac{1}{|2\pi Q|^{1/2}} \exp \left(-\frac{1}{2} s_i^\dagger Q_{ij} s_j \right) \quad (21)$$

Again, it is instructive to consider the simple case of Gaussian noise with this ESP prior where the propagator D in the information Hamiltonian Eqn 5 has the simple form

$$D = \left[Q + R^\dagger \sum_e^{-1} R \right]^{-1} \quad (22)$$

where Σ_e is defined in the paragraph following Eqn 1. Without interactions $H_I = 0$ and using linearly dependent on the data response-over-noise weighted information source

$$j = R^\dagger \Sigma_e^{-1} d \quad (23)$$

the propagator Eqn 22 is similar in form to Eqn 11 but now has recast the noise corrected propagator in the ESP basis in terms of an interaction free IFT model. The estimate of the signal is the (from either Eqn 6 and the resulting equivalent of Eqn 9 or from Eqn 8 with no interactions)

$$\psi = D_j = Q^+ d \quad (24)$$

$$\text{where } Q^+ = (Q + R^\dagger \Sigma_e^{-1} R)^{-1} R^\dagger \Sigma_e^{-1} \quad (25)$$

Thus in a similar fashion as discussed in the standard least squares Eqn 13, the signal is expressed in terms of the eigenmodes of an operator, but this time Q^+ rather than the pseudo-inverse of any model functions. (Q^+ is not actually a pseudo-inverse – we use the slight abuse of notation with a superscript $+$ to draw a similarity with Eqn 14).

Hence the ESP eigenmodes can be viewed also as free modes of IFT when the noise corrected coupling matrix Q (Eqn 22) is used as a propagator. However, in general, and in the specific applications examined below, these assumptions are violated and this simple description does not hold. The general EFD formalism, and the algorithm described below, does not depend on this description and no assumption of Gaussian noise is made.

It is important to emphasize a critical feature of the EFD construction at this point. The coupling matrix Q is not constructed from assumed model functions (as in the simple standard least squares example above) but is derived directly from the correlations in the data themselves. Moreover, it is not simply an adjacency matrix but can be constructed (by the user) from any desired coupling scheme consistent with the data and the application [32]. In the current paper, for example, we use nearest neighbor interactions, but more complicated interactions are possible as well. Thus, by construction, it may depend on the data in rather complex way, and hence the EFD model expressed by Eqns 5, 22 and 23, although remaining interaction free, does not possess the property that is a major limitation of many data analysis methods in areas ranging from brain imaging to weather related data processing to cosmic microwave background data assimilation – Gaussianity is not assumed in the EFD approach by its very construction.

An important practical implication of the EFD approach is that ESP ranking of eigenmodes allows reduction of the problem dimensions by writing a Fourier expansion using $\{\phi^{(k)}\}$ as the basis functions

$$\psi(\xi_l) = \sum_k^K \left[a_k \phi^{(k)}(\xi_l) + a_k^\dagger \phi^{\dagger, (k)}(\xi_l) \right] \quad (26)$$

and keeping number of modes K significantly smaller than the overall size of the problem nN by examining importance of the eigenvalues λ_k comparing them to the noise covariance $|\Sigma_e|$. Note that, as a consequence of Eqn 16, these basis functions are unique once the coupling matrix has been defined. Furthermore, the localization phenomena peculiar to the ESP eigenvectors distinguishes the eigenfunctions used in Eqn 26 from other harmonic bases.

The information Hamiltonian Eqn 5 can then be written in ESP basis Eqn 26 as

$$H(d, a_k) = -j_k^\dagger a_k + \frac{1}{2} a_k^\dagger \Lambda a_k + \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k_1}^K \cdots \sum_{k_n}^K \tilde{\Lambda}_{k_1 \cdots k_n}^{(n)} a_{k_1} \cdots a_{k_n} \quad (27)$$

where matrix Λ is the diagonal matrix $\text{Diag}\{\lambda_1, \cdots, \lambda_K\}$, composed of the eigenvalues of the noise corrected coupling matrix, and j_k is the amplitude of k th mode in the expansion of the source j

$$j_k = \int j \phi^{(k)}(\xi) d\xi \quad (28)$$

The expression for the classical solution Eqn 8 for the mode amplitudes a_k then becomes

$$\Lambda a_k = \left(j_k - \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k_1}^K \cdots \sum_{k_n}^K \tilde{\Lambda}_{k k_1 \cdots k_n}^{(n+1)} a_{k_1} \cdots a_{k_n} \right) \quad (29)$$

The new interaction terms $\tilde{\Lambda}^{(n)}$ are expressed through integrals over ESP eigenmodes

$$\tilde{\Lambda}_{k_1 \cdots k_n}^{(n)} = \int \cdots \int \Lambda_{\xi_1 \cdots \xi_n}^{(n)} \phi^{k_1}(\xi_1) \cdots \phi^{k_n}(\xi_n) d\xi_1 \cdots d\xi_n \quad (30)$$

The interaction terms $\Lambda^{(n)}$ should be specified in order to be able to estimate amplitudes a_k of the self-interacting modes. The simplest way to take into account the interactions would be an assumption of local-only interactions. This can be easily accomplished by factorization $\Lambda^{(n)}$ in a product of delta functions $\alpha^{(n)} \delta(\xi_1 - \xi_2) \cdots \delta(\xi_1 - \xi_n)$, here $\alpha^{(n)} < 1$ are constants. This results in simple but not particularly useful expression for $\tilde{\Lambda}_{k_1 \cdots k_n}^{(n)}$

$$\tilde{\Lambda}_{k_1 \dots k_n}^{(n)} = \alpha^{(n)} \int \phi^{k_1}(\xi) \dots \phi^{k_n}(\xi) d\xi \quad (31)$$

which after substituting it into e.g. Eqn 29 just provides the expression for the classical local-only interacting field [31] recast in the reduced dimensions ESP eigenmodes basis.

To obtain more interesting (and more practically useful) expression for estimating amplitudes of interacting modes we may assume that the nonlinear interactions between different modes will reflect the coupling. A natural way to take coupling into account would be through factorization of $\Lambda^{(n)}$ in powers of the coupling matrix, i.e. we can assume that

$$\Lambda_{\xi_1 \dots \xi_n}^{(n)} = \frac{\alpha^{(n)}}{n} \sum_{p=1}^n \prod_{\substack{m=1 \\ m \neq p}}^n Q_{\xi_p \xi_m} \quad (32)$$

which results in

$$\tilde{\Lambda}_{k_1 \dots k_n}^{(n)} = \frac{\alpha^{(n)}}{n} \sum_{p=1}^n \left(\frac{1}{\lambda_{k_p}} \prod_{m=1}^n \lambda_{k_m} \right) \int \left(\prod_{r=1}^n \phi^{k_r}(\xi) \right) d\xi \quad (33)$$

Here values of the coefficients $\alpha^{(n)}$ should be chosen small enough to ensure the convergence of the classical solution Eqn 29. From a practical standpoint values of $\alpha^{(n)} \lesssim 1/\max(j_k^n/\lambda_k)$ provide good starting estimate for further adjustments.

This expression is correct up to the third ($n = 3$) order but discards various chain-like factorizations (e.g. $Q_{\xi_1 \xi_2} Q_{\xi_2 \xi_3} \dots Q_{\xi_{n-1} \xi_n}$) for higher ($n > 3$) orders. These chain-like terms may be included as well by re-expanding required nonlinear combinations of ESP basis functions through the same basis. We would like to emphasize, that this task is not impracticable as in many “real life” applications the ESP eigenmodes are expected to be compactly localized because of the unique localization properties of the ESP eigenvectors. Therefore, nonlinear expressions that involve various powers of ESP eigenmodes can be expected to decay significantly faster than nonlinear terms expressed either through the whole domain integration or with the traditional trigonometric functions or polynomials used in the whole domain Fourier-like expansions. Nevertheless, this fact was neither explored nor used to obtain the results presented in the following sections.

Of course, to be completely fair, the traditional trigonometric functions and polynomials (e.g. Legendre or Chebyshev) have an important advantage when used as basis functions, especially for deriving analytical relations – their nonlinear forms can easily be expressed through the linear forms by using simple recurrence formulas, i.e. as simple as frequency scaling for the exponentials. Besides, in many practical applications temporal variations are well-characterized by frequency modes and thus instead of using the ESP expansion in the whole space-time domain Eqn 26, it may often be beneficial to use ESP basis only for

spatial coordinates $\{x_j\}$ while keeping the traditional Fourier polynomial expansion in the temporal $\{t_j\}$ domain

$$\psi(x_i, t_j) = \sum_{k,l} \left[a_{k,l} e^{i\omega_l t_j} \phi^{(k,l)} + a_{k,l}^\dagger e^{-i\omega_l t_k} \phi^{\dagger,(k,l)} \right] \quad (34)$$

Note that the coupling matrix Q is now expressed in the frequency domain rather than the spatial domain and thus is now different at different frequencies ω_k , hence, the spatial ESP basis functions $\phi^{(k,l)}$ depend on frequency as well. Except for the appearance of the second index in this form of expansion, the rest of the approach, including the information Hamiltonian Eqn 27 and the form of the interaction terms Eqn 32 can easily be recast using this new basis.

We would like to stress once more that this non-Gaussian and non-linear EFD approach represents a natural special case of the general IFT for this particular type of prior information and can produce solutions using all the useful techniques, including Feynman diagrams, that were shown in [31], or just by using any suitable iterative method for the classical solution of Eqn 29. In the next section (and in the Supplementary Material) we illustrate the EFD method using several simple models of spatially non-overlapping and overlapping time periodic and non-periodic sources, and show that using simple anharmonic terms Eqn 32 in the interaction Hamiltonian Eqn 27 allows reliable and natural identification and separation of spatially overlapping non time periodic modes, a task which is important in many (unrelated) areas that require analysis of spatio-temporal data.

4. Implementation

The general EFD formalism is very flexible and allows for multiple spatial and temporal correlation orders to be incorporated, and can include a wide range of prior information, such as more realistic models for the relationship between the brain blood flow and metabolism and the resulting fMRI signal, which is known to be quite complicated and non-linear [49, 50, 51, 52]. However, for this initial paper we limit our implementation to nearest neighbor interactions (in both space and time) and a Gaussian noise model. This is a reasonable first approximation for the data to be analyzed. The rsfMRI data we will analyze, for example, has had the non-linear physiological noise measured and removed before our analysis. Nevertheless, this rather straightforward implementation is sufficient to demonstrate the power and utility of the method. Two slightly different implementations were used, the first was using a complete spatial-temporal ESP basis for signal expansion Eqn 26, and the second was based on spatial ESP but employed Fourier expansion in the temporal domain Eqn 34. The details of these implementations are described below. All the algorithms implemented in this paper were written in standard ANSI C/C++. The spatio-temporal EFD procedure used in this paper for estimating the signal modes consisted of the following steps:

- i. Generate coupling matrix Eqn 15 using simple nearest neighbor coupling $Q(\xi_i, \xi_j) = d(\xi_i) d(\xi_j) A_{ij}$, where A_{ij} is the space-time adjacency matrix, i.e.

A_{ij} equals 1 if i and j are nearest neighbors in space or time domains, and 0 otherwise.

- ii. Find ESP eigenvalues λ_k and eigenvectors $\phi^{(k)}$ for the coupling matrix $Q(\xi_i, \xi_j)$ solving the eigenvalue problem Eqn 16.
- iii. Use ESP eigenvalues λ_k and eigenvectors $\phi^{(k)}$ to construct the information Hamiltonian Eqn 27, where Λ is simply the diagonal matrix $\text{Diag}\{\lambda_1, \dots, \lambda_K\}$ of ESP eigenvalues, and the interaction terms $\Lambda^{(n)}$ are constructed from ESP eigenvalues and eigenvectors with the help of Eqn 33.
- iv. Finally, the amplitudes a_k that describe both spatially and temporally interacting modes of the information Hamiltonian Eqn 27 are found from the nonlinear expression for the classical solution Eqn 29.

The alternative implementation (corresponding to Eqn 34) although follows the above estimation steps nevertheless has several important differences worth to be mentioned. First of all, instead of generating nearest neighbor coupling both in space and time domains it employs frequency dependent spatial coupling matrix $Q(x_i, x_j, \omega)$ taking nearest neighbor coupling only in spatial domain (here x_i and x_j are spatial coordinates and ω is a frequency). Second, the strength of coupling for each frequency depends on the temporal pair correlation function. There are different ways to introduce this temporal correlation dependence. We used the following form of coupling matrix

$$Q(x_i, x_j, \omega_0) = \mathcal{R}_{ij}^m d(x_i, \omega_0) d(x_j, \omega_0), \quad (35)$$

$$Q(x_i, x_j, \omega_l) = \mathcal{R}_{ij}^m (\phi^{(1)}(x_i, \omega_0) d(x_j, \omega_l) + \phi^{(1)}(x_i, \omega_0) d(x_i, \omega_l)), \quad (36)$$

here $\mathcal{R}_{ij}$ is either the mean

$$\mathcal{R}_{ij} = \frac{1}{T} \int_0^T \int_0^T d(x_i, t - \tau) d(x_j, \tau) d\tau dt \quad (37)$$

or the maximum

$$\mathcal{R}_{ij} = \max_t \int d(x_i, t - \tau) d(x_j, \tau) d\tau \quad (38)$$

of the temporal pair correlation function computed for spatial nearest neighbors i and j , $\phi^{(1)}(x_j, \omega_0)$ is the eigenmodes that corresponds to the largest eigenvalue of $Q(x_i, x_j, \omega_0)$, and the exponent $m = 0$ is used to attenuate the importance of correlations.

The additional implementation steps can be summarized then as follows:

- i. Generate the temporal pair correlation functions $\mathcal{R}_{ij}$ for every i and j pair of spatial nearest neighbors.
- ii. Compute the mean (or largest) correlation value for each pair and use this as a coupling coefficient from the propagator to find the spatial eigenmodes $\phi(x_i, \omega_0)$ with zero frequency ω_0 (that is, the mean field) by solving the eigenvalue problem Eqn 16 for the coupling matrix $Q(x_i, x_j, \omega_0)$ from Eqn 35.
- iii. Use the zero frequency spatial eigenmodes $\phi(x_i, \omega_0)$ to construct coupling matrices $Q(x_i, x_j, \omega)$ in Eqn 36 and solve the eigenvalue problem Eqn 16 for each ω_l frequency.
- iv. Generate the information Hamiltonian Eqn 27 by summation of input from interaction terms Eqn 33 and solve for the mode amplitudes a_k in a way similar to the last two items of the spatio-temporal approach.

The first three steps determine the values of mean field at every spatial position and then determine the spatio-temporal eigenmodes in spatial-frequency (i.e. Fourier) space assuming non-interacting fields. The last step determines the interactions between these eigenmodes. The final results are space/time localization patterns that are our *definition* of the “modes” of the data.

5. Results

To illustrate the capabilities of our method we apply the method to two toy test cases and then to two real data sets.

The two toy examples are inspired by FMRI which comes in two “flavors” that are both easily amenable to idealized simulations and serve as good paradigms to test and validate the EFD method and compare it to existing methods. The second of these examples is also a reasonable model for the situation faced in the analysis of mobile Doppler radar data. We conclude the section with a demonstration of the EFD method on two real data sets that are part of ongoing research in our lab: resting state FMRI data and mobile Doppler radar data from a tornadic supercell thunderstorm.

5.1. Demonstration on simulated 2D data and comparison with state-of-the-art methods

In “traditional” FMRI, a subject is presented with a well defined input stimulus, such as a visual stimulation consisting of a rapidly flashing (e.g. 8Hz) checkerboard that is presented for a short period (e.g., 10sec), turned off for the same period, and this pattern of presentation is repeated several times, resulting in a so-called *block stimulus design*. This is an example of *task-based* FMRI, so-called because the input task (the stimulus) is known. While the relationship between the input stimulus and the FMRI signal is actually quite complicated [49, 50, 51, 52], it is often quite close to the stimulus. Thus simple correlation of the input stimulus (perhaps convolved with a neuronal response function) with the signal is a useful and established analysis method [53], as long as the signals are not spatially or temporally overlapping. If they are, traditional correlation analysis methods fail and most sophisticated techniques such as ICA have been employed, though are known to be

insufficient even in relatively simple cases [22]. Our first example is one such simple case in which the state-of-the-art ICA method fails whereas EFD is able to recover the correct signals and thus provides a rather simple but powerful demonstration of EFD capabilities that can be directly compared with ICA results.

While a full comparison of the EFD method with existing traditional methods is beyond the scope of this paper, a direct comparison with ICA in a very simple idealized numerical model of brain activation will serve to illustrate the essential features of our method. Consider two spatially overlapping ellipsoidal regions in which every voxel within the first region has the same square wave activation pattern and every voxel within the second region also has the same activation waveform, though different than the waveform in the first region. The entire image is contaminated by Gaussian noise. The amplitudes of the activations are such that the signal-to-noise is low. The square wave pattern is an idealization of the simplest fMRI experiment (often called “task based” fMRI) in which a subject is presented with a stimulus that is turned on and off at regular intervals and the brain regions activate in concert with the stimulus. This particular example would thus represent a highly idealized brain with distinct ellipsoidal regions that each exhibit activity correlated with a different one of the two stimuli. The signals from the two regions are assumed to be additive. The brain activation patterns from a true rsfMRI experiment are much more complicated than this example, being non-linear, coupled and in three spatial dimensions, so this test should represent a simple benchmark for the efficacy of any proposed rsfMRI method. The simulated data are shown in the top row of Figure 1. The EFD analysis are shown in the middle row. Only two modes are detected and these are seen to correspond to the correct spatial regions with the correct temporal profiles. EFD has thus identified the correct space-time regions of the signals.

On the other hand, the ICA results shown in the bottom row of Figure 1 are erroneous as has been previously pointed out [22]. The components are a mixture in both space and time of the two true modes. While the algorithm undoubtedly constructs two “independent” components, this example clearly illustrates that this is a poor model for even this simple brain activation model, and thus most likely for actual brain activation data. Indeed, the signal modes are not independent in that they share at least some portion of the same space-time region. Requiring them to be maximally independent is thus forcing on them a property they do not intrinsically have. The EFD procedure, on the other hand, simply constructs the most probable pathways in space-time based on the *measured* correlations in the data. Because of the localization properties of ESP (first observed in the maximum entropy random walk [44]), there are in fact very few space-time parameter configurations consistent with the prior coupling information. The “modes” thus represent the configurations that are consistent with the data and the most probable. We would like to reiterate that the simplicity of this example was for demonstrative purposes but emphasize that the EFD methods does not assume Gaussian noise, simple additivity of the signal, or linearity of the signal.

5.2. Demonstration on simulated 3D non-periodic, overlapping data

The second toy example is an idealization of the more practical situation faced in many scientific applications, including our own particular cases of fMRI and mobile Doppler

radar data and consists of mixing different time varying signals inside several three-dimensional spatial domains. This is a model for the second “flavor” of fMRI is the acquisition data while the subject is not presented with any stimulus and is simply “resting”. This is called *resting state* fMRI, or rsfMRI, and the analysis of the detected spatio-temporal fluctuations presents a tremendous challenge because they are characterized by being non-linear, non-periodic, and spatially and temporally overlapping and there is no known input stimulus with which to compare these fluctuations, as they are thought to be due to “intrinsic” modes of brain activity.

This example is also a reasonable idealized model for the problem faced in mobile Doppler radar data from severe thunderstorms where complex spatio-temporal variations in the detected reflectivity and wind speeds are driven by complex storm dynamics characterized by coherent variations in dynamical parameters such as vertical vorticity and vorticity stretching.

The simulation is shown in Figure 2, which consists of a central sphere (white) located at the origin ($\{x, y, z\} = \{0, 0, 0\}$) oscillating at a single, periodic frequency, surrounded by six spherical or ellipsoidal regions along the principal axes $\{x_1, 0, 0\}$, $\{-x_2, 0, 0\}$, $\{0, y_1, 0\}$, $\{0, -y_2, 0\}$, $\{0, 0, z_1\}$, $\{0, 0, -z_2\}$. The signals are the same throughout the volume of any particular domain. In addition, Gaussian noise has been added. Three spheres (red, green, blue), spatially separated from the central sphere, each oscillate at a single, distinct frequency, though at different maximum amplitudes. Three ellipsoids (magenta, yellow and cyan) overlap the central sphere and have non-periodic time courses (again with different maximum amplitudes) created by filtering a sinusoidal amplitude with a Fermi filter, which turns the signal on and off smoothly in time. Both the periodic and non-periodic objects overlap spatially, such that in the center area of the white sphere, signals from four different objects are mixed (one periodic signal from the white sphere itself and three different nonperiodic signals from magenta, yellow and cyan ellipsoids). This example illustrates the important fact that extracted non-linear EFD modes need not be orthogonal. This is intuitively clear from the fact that the interaction terms involve products of the ESP eigenfunctions. While the individual ESP eigenmodes are mutually orthogonal, products of these functions are not. This is crucial in many, if not most, applications, such as in the case of rs-fMRI data below, where one would not expect the data modes to be orthogonal.

Three panels of Figure 3 show spatial and temporal patterns for extracted non-periodic ellipsoids overlapped with the central sphere. The temporal profiles for all spheres correspond to SNR from 6.8 to 7 (we used Gaussian $\sigma=0.1$ in this example). Figure 4 shows the signal extracted from the area of the central sphere where four different signals are mixed.

5.3. Results for practical applications

To emphasize a rather generic nature of presented theory we provide quick results for data analysis of two datasets from completely unrelated areas, both from physical and informational sense. The first example is based on biological data – human resting state functional magnetic resonance imaging (rs-fMRI) data. Atmospheric data from a mobile Doppler radar system was used for the second example.

The resting state fMRI data used in this study were from a single subject from a previously published study [54]. Results of the EFD analysis is shown in Figure 5 where contours of power in two (arbitrarily chosen) modes of brain activation in a normal human subject from the resting state data described in [54] are shown. Distinct spatio-temporal patterns of activation are evident in each mode. It is evident that different modes can have spatially overlapping regions because they do not have spatio-temporal overlapping regions. Aside from the initial data preparation described in [54] (described in the Appendix), no processing other than the EFD algorithm described above was used. In particular, no additional noise filtering of any kind was employed.

The data analyzed in the second example used for this study were from the 5 June 2009 tornadic supercell in Goshen County, Wyoming and collected using the Doppler On Wheels (DOW [37]) mobile Doppler radar system during the second Verification of the Origins of Rotation in Tornadoes Experiment (VORTEX2; [38, 39]). Figure 6 shows the distinct core of low-level rotation consistently detected by EFD in one of the major modes and appears to be consistent with recent tornadogenesis theories [55, 56, 57, 39].

To date, the primary method of analysis in Doppler radar studies of tornadogenesis has been confined to laborious and primarily qualitative analyses based on traditional visualization methods such as contour diagrams, isosurfaces, and streamline generation (e.g. [56, 57, 39]). While these methods are straightforward (though burdensome) to implement, they are limited in their ability to capture and quantify spatio-temporal correlations patterns, or “modes”, that typically characterize complex systems such as tornadic storms.

6. Discussion & Conclusion

The problem of detecting non-linear and non-Gaussian signals in multivariate data is ubiquitous in the sciences and of ever increasing importance as instrument technologies (medical imaging scanners, space telescopes, etc) continue to advance. In fact, the ability of modern instruments to collect huge amounts of multivariate data now results in many situations in which the data are of too great a dimensionality to employ traditional, straightforward analysis methods. This is particularly true in situations in which little is known about the form of the expected signal, in which case the problem is too poorly posed to be tractable. The exemplary cases of resting state fMRI data and mobile Doppler radar data were chosen for this paper not only because of their significant scientific interest but also because they embody the essential difficulties facing data analysts in a wide range of fields. Both types of data consist of large volumes (i.e., three spatial dimensions) of densely sampled, relatively low signal-to-noise, voxels acquired at many time points. And within these data the quantities of interest are consistent but non-linear and non-Gaussian spatio-temporal patterns, or “modes”, of parameters that characterize the physical systems. In fMRI, these modes represent patterns of coherent brain activity, whereas in mobile Doppler radar, the modes characterize the development and interplay of severe thunderstorm parameters such maximum radar reflectivity, vertical vorticity, tilting and stretching of the vertical vorticity. The dimensionality of the possible characterizations of these systems is huge since spatio-temporal variations can occur on a multitude of spatial and temporal scales, and potentially interact in an exceedingly large number of combinations. Thus the

definition of a “mode” is itself an open question, and without a clear definition the question of the proper method of data analysis is moot. This problem can only be made tractable by using prior information to reduce the possible reasonable sets of parameters.

While the use of probability theory in the analysis of scientific data is always the correct approach, its necessity often only becomes evident when the explicit incorporation of prior information is required since, in the case of non-informative priors, probability theory reduces to standard techniques. Moreover, in many cases, of which the rs-fMRI and mobile Doppler radar are just two examples, the underlying system that one seeks to describe from the discrete data is continuous, and thus the spatio-temporal variations are most appropriately represented as a field. These two considerations lead to the adoption of information field theory (IFT, [31]) which is a reformulation of (Bayesian) probability theory in the language of field theory. The ability to incorporate prior information is critical to constructing a tractable formulation of many problem such as that posed by rsfMRI and mobile Doppler radar data. In particular, it facilitates the integration of our recent theory of entropy spectrum pathways (ESP, [32]) which uses local coupling information to construct and rank optimal pathways (and thus patterns) in a problems parameter space.

We would like to emphasize that this unique synthesis of ESP and IFT produces a method that has demonstrable advantages over a plethora of methods developed for detecting non-linear and non-Gaussian spatial correlations in multidimensional time series, generally referred to as multivariate analysis. Most of these multivariate analysis techniques (e.g. EOF, SSA, M-SSA) [11, 58] as well as various constrained data assimilation models [59] simply compute the spatial modes of correlations. Moreover, these methods are typically faced with the difficulty of searching for solutions within a huge parameter space. However, the main purpose of our EFD approach is to detect and separate spatio-temporal modes that are inherently non-Gaussian and non-linear in both space and time and are produced as a result of complex spatio-temporal interacting pathways or conditions. Additionally, the probable parameter configurations have been drastically reduced by employing the ESP theory in conjunction with the correlations within the data.

The current state-of-the-art in rsfMRI data analysis is *independent components analysis* (ICA) which purportedly has the ability to detect spatial and temporal characteristics of multiple non-linear signals with unknown parameters. ICA is predicated on a number of assumptions, such as that the activated regions of the brain are spatially both sparse and independent [60, 9, 21]. This “classical” ICA is implicitly based on the assumption that the signal has no noise and thus is not a proper probabilistic formulation for which significance of the estimates can be obtained. To address this deficiency, probabilistic ICA (PICA) was developed and involves estimation of independent components in a subspace that contains the signal and the noise orthogonal to the signal, then estimating the statistical significance of the detected sources [24, 61]. This has become the most widely used implementation of ICA for rsfMRI analysis [1]. Yet PICA is no less an *ad-hoc* procedure that involves numerous assumptions, such as additive Gaussian noise, maximal non-Gaussianity of the sources, non-degenerate mixing matrix, and the efficacy of subspace and model order determination by probabilistic principle components analysis (PPCA [62]), to name but a

few [24]. Machine learning ICA methods have also been used which are essentially mixed-effect group models that also employ linear models [63].

The theory has been applied to the analysis of data collected from two distinctly different physical processes. In the first example, resting state functional magnetic resonance imaging (rs-fMRI) data of human brain activity was analyzed and some initial results are presented. We showed that the most modest and conservative declaration of prior information and model assumptions is sufficient to formulate the rs-fMRI data question in a precise theoretical form that lends itself to an efficient and accurate computational method for assessing and categorizing modes of brain activity. For the second example, we applied the technique to mobile Doppler radar data collected in a tornadic supercell and demonstrated the identification and tracking of several important spatio-temporal parameter features pertinent to tornado development. The complete description of both of these applications will be presented in separate publications.

Acknowledgments

The authors thank Dr Alec Wong and Dr Tom Liu at the UCSD CFMRI for providing the resting state data. We gratefully acknowledge helpful review of the manuscript with a number of important comments and suggestions from anonymous reviewers. LRF and VLG were supported by NSF grants ACI-1440412, DBI-1143389, DBI-1147260, EF-0850369, PHY-1201238, ACI-1550405 and NIH grant R01 MH096100. Doppler On Wheels (DOW) tornado data courtesy of Dr Joshua Wurman, Center for Severe Weather Research (CSWR). DOWs and DOW data collection supported by National Science Foundation (NSF) grants 1211132, 1361237 and 1447268.

Appendix A: FMRI data

The resting state FMRI data used in this study were from a single subject from a previously published study [54]. All data shown were collected post-administration of 200 mg of caffeine. Blood oxygenation level dependent (BOLD) imaging data were acquired on a 3T GE Discovery MR750 whole body system using an eight channel receiver coil. High resolution anatomical data were collected using a magnetization prepared 3D fast spoiled gradient (FSPGR) sequence (TI = 600 ms, TE = 3.1 ms, flip angle = 8, slice thickness = 1 mm, FOV = 25.6 cm, matrix size = 256 256 176). Whole brain BOLD resting-state data were acquired over thirty axial slices using an echo planar imaging (EPI) sequence (flip angle = 70, slice thickness=4mm, slice gap=1mm, FOV=24cm, TE= 30 ms, TR = 1.8 s, matrix size = 64 64 30). Field maps were acquired using a gradient recalled acquisition in steady state (GRASS) sequence (TE1 = 6.5 ms, TE2 = 8.5 ms), with the same in-plane parameters and slice coverage as the BOLD resting- state scans. The phase difference between the two echoes was then used for magnetic field inhomogeneity correction of the BOLD data. Cardiac pulse and respiratory effect data were monitored using a pulse oximeter and a respiratory effort transducer, respectively. The pulse oximeter was placed on each subject's right index finger while the respiratory effort belt was placed around each subject's abdomen. Physiological data were sampled at 40 Hz using a multi-channel data acquisition board.

Appendix B: Tornado data

The data analyzed in the second example used for this study were from the 5 June 2009 tornadic supercell in Goshen County, Wyoming and collected using the Doppler On Wheels (DOW [37]) mobile Doppler radar system during the second Verification of the Origins of Rotation in Tornadoes Experiment (VORTEX2; [38, 39]). Dual Doppler data from DOW6 (Lat/Lon = $-104.34732, 41.49556$) and DOW7 (Lat/Lon = $-104.25203, 41.61437$) were combined in an objective analysis at 17 time points from 2142–2214 UTC equally spaced by 2 minute intervals on a Cartesian grid (centered at DOW6) of dimensions $\{nx, ny, nz\} = \{301, 301, 41\}$. The field of view in each dimensions were $\{D_x, D_y, D_z\} = \{100m, 100m, 100m\}$. The elevation angles (in degrees) used in the objective analysis for the two volumes were DOW6: $\{0.5, 1, 2, 3, 4, 5, 6, 8, 10, 12, 14, 16\}$, DOW7: $\{1, 2, 3, 4, 5, 6, 0.5, 8, 10, 12, 14, 16\}$. Barnes analysis was used with $\kappa = 0.216m^2$ (horizontal = vertical) [64]. For the dual analysis the minimum beam angle was $\phi_{min} = 30^\circ$. A 3-step Leise smoothing filter [65] is applied to the vorticity, tilting, and stretching vectors and a one-step Leise filter to the velocity components (u, v, w). The mesocyclone movement was subtracted from the velocities (u, v).

References

- Smith SM, Beckmann CF, Andersson J, Auerbach EJ, Bijsterbosch J, Douaud G, Duff E, Feinberg DA, Griffanti L, Harms MP, Kelly M, Laumann T, Miller KL, Moeller S, Petersen S, Power J, Salimi-Khorshidi G, Snyder AZ, Vu AT, Woolrich MW, Xu J, Yacoub E, Urbil K, Van Essen DC, Glasser MF, Consortium, for the WU-Minn HCP. Neuroimage. 2013; 80:144–168. [PubMed: 23702415]
- Heine L, Soddu A, Gómez F, Vanhaudenhuyse A, Tshibanda L, Thonnard M, Charland-Verville V, Kirsch M, Laureys S, Demertzi A. Frontiers in Psychology. 2012; 3:1–12. [PubMed: 22279440]
- Wurman JM, Kosiba KA, Markowski P, Richardson Y, Dowell D, Robinson P. Mon Wea Rev. 2010; 138:4439–4455.
- Marquis J, Richardson Y, Markowski P, Dowell D, Wurman J. Mon Wea Rev. 2012; 140:3–27.
- Kosiba K, Wurman J. Weather Forecast. 2013; 28:1552–1561.
- Christakos G. Systems, Man and Cybernetics, IEEE Transactions on. 1991; 21:861–875.
- Broomhead D, King G. Physica D. 1987; 20:217–236.
- Spencer K, Dien J, Donchin E. Psychophysiology. 2001; 38:343–358. [PubMed: 11347879]
- Bell A, Sejnowski T. Neural Comput. 1995; 7:1129–1159. [PubMed: 7584893]
- Hyvarinen A, Oja E. Neural Networks. 2000; 13:411–430. [PubMed: 10946390]
- Ghil M, Allen MR, Dettinger MD, Ide K, Kondrashov D, Mann ME, Robertson AW, Saunders A, Tian Y, Varadi F, Yiou P. Reviews of Geophysics. 2002; 40
- Wallace JM, Smith C, Bretherton CS. Journal of Climate. 1992; 5:561–576.
- Plaut G, Vautard R. J Atmos Sci. 1994; 51:210–236.
- Dixon M, Wiener G. Journal of Atmospheric and Oceanic Technology. 1993; 10:785–797.
- Min W, Wynter L. Transportation Research Part C-Emerging Technologies. 2011; 19:606–616.
- Hill MJ, Donald GE. Remote Sensing of Environment. 2003; 84:367–384.
- Gallez D, Babloyantz A. Biological Cybernetics. 1991; 64:381–391. [PubMed: 2049414]
- Bijma F, de Munck JC, Heethaar RM. NeuroImage. 2005; 27:402–415. [PubMed: 16019231]
- Plis S, George J, Jun S, Paré-Blagoev J, Ranken D, Wood C, Schmidt D. Physical Review E. 2007; 75:011928.
- Lamus C, Haemaelaeninen MS, Temereanca S, Brown EN, Purdon PL. NeuroImage. 2012; 63:894–909. [PubMed: 22155043]

21. McKeown MJ, Makeig S, Brown GG, Jung TP, Kindermann SS, Bell AJ, Sejnowski TJ. *Science*. 1998; 6:160–188.
22. Calhoun VD, Adali T, Pearlson GD, Pekar JJ. *Human Brain Mapping*. 2001; 13:43–53. [PubMed: 11284046]
23. Kiviniemi V, Kantola JH, Jauhainen J, Hyvarinen A, Tervonen O. *Neuroimage*. 2003; 19:253–260. [PubMed: 12814576]
24. Beckmann CF, Smith SM. *IEEE Trans Med Imaging*. 2004; 23:137–152. [PubMed: 14964560]
25. Tian L, Kong Y, Ren J, Varoquaux G, Zang Y, Smith SM. *PLoS ONE*. 2013; 8:1–12.
26. Krioukov D, Kitsak M, Sinkovits RS, Rideout D, Meyer D, Boguñá M. *Scientific Reports*. 2012; 2:1–6.
27. Bourdieu, P. *Distinctions A Social Critique of the Judgment of Taste*. Harvard University Press; 1984.
28. Uhl C, Friedrich R, Haken H. *Zeitschrift für Physik B Condensed Matter*. 1993; 92:211–219.
29. Uhl C, Friedrich R, Haken H. *Physical Review E*. 1995; 51:3890–3900.
30. Jaynes, E. *Probability Theory: The Logic of Science*. New York: Cambridge University Press; 2003.
31. Enßlin TA, Frommert M, Kitaura FS. *Phys Rev D*. 2009; 80:105005.
32. Frank LR, Galinsky VL. *Phys Rev E*. 2014; 89(3):032142. URL <http://link.aps.org/doi/10.1103/PhysRevE.89.032142>.
33. Smith SM, Beckmann CF, Andersson J, Auerbach EJ, Bijsterbosch J, Douaud G, Duff E, Feinberg D A, Griffanti L, Harms MP, Kelly M, Laumann T, Miller KL, Moeller S, Petersen S, Power J, Salimi-Khorshidi G, Snyder AZ, Vu AT, Woolrich MW, Xu J, Yacoub E, Uurbil K, Essen DCV, Glasser MF. *NeuroImage*. 2013; 80:144–168. mapping the Connectome URL <http://www.sciencedirect.com/science/article/pii/S1053811913005338>. [PubMed: 23702415]
34. Haimovici A, Tagliazucchi E, Balenzuela P, Chialvo DR. *Phys Rev Lett*. 2013; 110(17):178101. URL <http://link.aps.org/doi/10.1103/PhysRevLett.110.178101>. [PubMed: 23679783]
35. Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, Filippini N, Watkins KE, Toro R, Laird AR, Beckmann CF. *Proceedings of the National Academy of Sciences*. 2009; 106:13040–13045.
36. Zalesky A, Fornito A, Cocchi L, Gollo LL, Breakspear M. *Proceedings of the National Academy of Sciences*. 2014; 111:10341–10346. URL <http://www.pnas.org/content/111/28/10341.abstract>.
37. Wurman J, Straka J, Rasmussen E, Randall M, Zahrai A. *J Atmos Oceanic Technol*. 1997; 14:1502–1512.
38. Wurman J, Dowell D, Richardson Y, Markowski P, Rasmussen E, Burgess D, Wicker L, Bluestein HB. *Bull Amer Meteor Soc*. 2012; 93:11471170. URL <http://dx.doi.org/10.1175/BAMS-D-11-00010.1>.
39. Kosiba K, Wurman J, Richardson Y, Markowski P, Robinson P, Marquis J. *Mon Wea Rev*. 2013; 141:1157–1181.
40. Jaynes, E. *Probability Theory with Applications in Science and Engineering*. Washington University; 1974. Fragmentary ed edition
41. Chaichian, M.; Demichev, A. *Path Integrals in Physics: Volume II Quantum Field Theory, Statistical Physics and other Modern Applications*. Taylor & Francis; 2001. Institute of physics series in mathematical and computational physics URL <https://books.google.com/books?id=Kai-yszYbhAC>
42. Ryder, L. *Quantum Field Theory*. 1st. Cambridge University Press; 1985.
43. Tan SM. *Monthly Notices of the Royal Astronomical Society*. 1986; 220:971–1001.
44. Burda Z, Duda J, Luck J, Waclaw B. *Phys Rev Lett*. 2009; 102:160602. [PubMed: 19518691]
45. Burda, Z.; Duda, J.; Luck, JM.; Waclaw, B. *The Various Facets of Random Walk Entropy Acta Physica Polonica B. Jagellonian Univ, Marian Smoluchowski Inst Phys; PL-30059 Krakow, Poland: 2010. p. 949-987.*
46. Burda Z, Duda J, Luck J, Waclaw B. *Arxiv preprint arXiv:1004.3667*. 2010
47. Jaynes E. *Physical Review*. 1957; 106:620–630.

48. Jaynes E. Physical Review. 1957; 108:171.
49. Buxton R, Frank L. J Cerebr Blood F Met. 1997; 17:64–72.
50. Buxton RB, Wong E, Frank L. Magn Reson Med. 1998; 39:855–864. [PubMed: 9621908]
51. Buxton R, Uluda K, Dubowitz D, Liu T. Neuroimage. 2004; 23:S220–S233. [PubMed: 15501093]
52. Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. Nature. 2001; 412:150–157. [PubMed: 11449264]
53. Bullmore E, Brammer M, Williams SCR, Rabe-Hesketh S, Janot N, David A, Mellers J, Howard R, Sham P. Magn Reson Med. 1996; 35:261–277. URL <http://dx.doi.org/10.1002/mrm.1910350219>. [PubMed: 8622592]
54. Wong CW, Olafsson V, Tal O, Liu TT. Neuroimage. 2013; 83:983–990. [PubMed: 23899724]
55. Markowski PM, Richardson YP. Atmos Res. 2009; 93:3–10.
56. Markowski P, Richardson Y, Marquis J, Wurman J, Kosiba K, Robinson P, Dowell D, Rasmussen E, Davies-Jones R. Mon Wea Rev. 2012; 140:2887–2915.
57. Markowski P, Richardson Y, Marquis J, Davies-Jones R, Wurman J, Kosiba K, Robinson P, Rasmussen E, Dowell D. Mon Wea Rev. 2012; 140:2916–2938.
58. von Storch, H.; Zwiers, F. Statistical Analysis in Climate Research. Cambridge University Press; 2001. URL http://books.google.com/books?id=_VHxE26QvXgC
59. Jazwinski, A. Stochastic Processes and Filtering Theory. Elsevier Science; 1970. Mathematics in Science and Engineering URL <http://books.google.com/books?id=nGlSNvKyY2MC>
60. Comon P. Signal Process. 1994; 36:287–314.
61. Beckmann CF, DeLuca M, Devlin JT, Smith SM. Philos T Roy Soc B. 2005; 360:1001–1013.
62. Tipping ME, Bishop CM. Neural Computation. 1999; 11:443–482. [PubMed: 9950739]
63. Varoquaux G, Sadaghiani S, Pinel P, Kleinschmidt A, Poline JB, Thirion B. Neuroimage. 2010; 51:288–299. [PubMed: 20153834]
64. Barnes SL. J Appl Meteorol. 1964; 3:396–409.
65. Leise JA. NOAA Tech Memo ERL WPL-82. 1982

Figure 1.

Comparison of EFD with ICA in task-based fMRI simulation. (Top) Simulated signals with additive Gaussian noise so that the signal-to-noise is $SNR = 2.5$. The spatial dimensions are $(64 \times 64 \text{ voxels})$ and there are 160 time points. (Middle) Estimated modes using EFD. The cutoff for defining “relevant” modes was determined by the ratio of the mode powers and was set to -30 dB signal attenuation. (Bottom) Estimated modes using ICA. All voxels in region *A* have the same time course (4 cycles) and all voxels in region *B* have the same time course (5 cycles).

Figure 2.

Toy example with seven non-periodic and spatially overlapped regions. Red, green, blue, and white spheres are oscillating at a single distinct frequency, whereas magenta, yellow and cyan ellipsoids are oscillating with non-periodic time courses created by filtering the sinusoidal amplitude variations with a Fermi filter that smoothly but rapidly turns the activation on, then off. The width of the Fermi filter is 30% of the length of the time series with a transition width of 2 time points. The filter begins 30% of the way into the time series.

Figure 3.

Spatial (top) and temporal (bottom) pattern of extracted modes for temporally non-periodic ellipsoids that overlap with each other and with a periodically oscillating central sphere. The time sequences on the bottom panel show the original signals in the overlap region of each ellipsoid (solid line) and the extracted signal (black dashed line). The original signal from the isolated regions (without mixing with the signal from the neighboring overlapping areas) are also included with dotted lines. As the whole volume is contaminated by Gaussian noise with $\sigma = 0.1$ the signal for all spheres corresponds to a signal-to-noise of $SNR \approx 6.9$.

Figure 4.

Spatial (a) and temporal (b) pattern of the extracted mode for the central overlapping sphere. This sphere overlaps with three neighboring ellipsoids (cyan, magenta and yellow), all four different signals contribute to the overall signal at the center area of the white sphere. The time sequence on panel (b) shows the original signals at the center of the white sphere (solid line) and the extracted signal (dashed line). The original signal from the isolated white sphere (without mixing with the signal from all four neighboring overlapping regions) is also shown by the dotted lines. In panel (c) restored signals for all four modes that give input to the original signal at the central area are plotted (signals from magenta, yellow, cyan ellipsoids, and gray sphere). Panel (d) shows this original noisy signal (solid line) and the sum of all four restored signals (dashed line).

Figure 5.

EFD results on a normal human subject from the resting state data described in [54]. The power in a single (arbitrarily chosen) mode (the 5th) is shown in the blue-green contours on the bottom, along with the first eigenmode (red contours) of the functional tractography (top). The functional tractography was seeded by the high probability regions of the power. In this dataset, 23 significant modes were detected. These modes have significant overlap in both space and time with one another. However, they are distinct modes because they do not have overlapping spatio-temporal regions. Aside from the initial data preparation described in [54], no processing other than the EFD algorithm described above was used. In particular, no additional noise filtering of any kind was employed. Movie collage of several significant modes is available in Videos 11 and 12 of the Supplement.

Figure 6.

EFD analysis of mobile Doppler radar data for the second Verification of the Origins of Rotation in Tornadoes Experiment (VORTEX2; [38, 39]). Modes of reflectivity (pink/green/blue contours, R), vertical vorticity (purple, ζ), stretching of vertical vorticity (yellow, σ) from a single time step are shown along with vorticity tracts (red, V) generated by functional tractography. The generation and intensification of low-level rotation is clearly detected in the major modes detected by EFD, and appears to be consistent with recent theories focusing on the role of the descending reflectivity core (*DRC*) [55, 56, 57, 39]. Aside from the initial data preparation described in [38, 39], no processing other than the EFD algorithm described above was used. In particular, no noise filtering of any kind was employed. Two movies of the results for all time frames are available in Videos 13 and 14 of the Supplement.