

# Partitioning networks into clusters and residuals with average association

Martin Vejmelka<sup>a)</sup> and Milan Paluš

*Institute of Computer Science, Academy of Sciences of the Czech Republic,  
Pod Vodárenskou věží 2, 182 07 Praha, Czech Republic*

(Received 25 March 2010; accepted 10 June 2010; published online 23 July 2010)

We investigate the problem of detecting clusters exhibiting higher-than-average internal connectivity in networks of interacting systems. We show how the average association objective formulated in the context of spectral graph clustering leads naturally to a clustering strategy where each system is assigned to at most one cluster. A residual set is formed of the systems that are not members of any cluster. Maximization of the average association objective leads to a discrete optimization problem, which is difficult to solve, but a relaxed version can be solved using an eigendecomposition of the connectivity matrix. A simple approach to extracting clusters from a relaxed solution is described and developed by applying a variance maximizing solution to the relaxed solution, which leads to a method with increased accuracy and sensitivity. Numerical studies of theoretical connectivity models and of synchronization clusters in a lattice of coupled Lorenz oscillators are conducted to show the efficiency of the proposed approach. The method is applied to an experimentally obtained human resting state functional magnetic resonance imaging dataset and the results are discussed. © 2010 American Institute of Physics. [doi:10.1063/1.3460360]

**The proliferation of high-dimensional datasets in many areas of science has created a need for methods that help in the investigation of large sets of data. Relationships in such datasets are often modeled as networks of interacting or related elements. This paper proposes a method of finding cohesive functional units of elements in weighted undirected networks using a clustering approach. An investigated network is partitioned into a set of clusters representing these units and a residual set. The residual set contains elements which do not relate well to any of the identified clusters. This approach enables a concise characterization of functional units even in the presence of unrelated or confounding elements.**

## I. INTRODUCTION

The use of eigenvalues of the graph Laplacian in characterizing the connectedness of graphs can be traced back to Fiedler<sup>1,2</sup> and the use of eigenvectors in partitioning of graphs to Donath and Hoffman.<sup>3</sup> Since then, the field of spectral graph clustering has been steadily receiving increasing attention as graph partitioning methods have been shown to be effective in many settings.

A spectral graph clustering approach is typically based on the formulation of a discrete constraint optimization problem. Often a relaxed version of the problem can be solved using a spectral approach—by using the eigendecomposition of the connectivity matrix. The remaining problem is then to move from the solution of the relaxed problem (relaxed solution) back to the original discrete formulation. At this point, heuristics are usually applied and there is generally no guarantee of the optimality of the discrete solution thus ob-

tained. Nevertheless, spectral graph clustering has shown excellent results in many areas.<sup>4–8</sup>

Pothen *et al.*<sup>4</sup> used the eigenvectors of the graph Laplacian to find small vertex separators of large sparse graphs in the context of parallel sparse matrix factorization algorithms. Hagen and Kahng<sup>5</sup> applied the ratio cut criterion to partition netlists in various phases of very large systems integration circuit design. Shi and Malik<sup>6</sup> introduced a successful formulation of spectral graph partitioning using the normalized cut criterion for high-level image segmentation. The last work has sparked strong interest in spectral graph partitioning.

While the first papers were focused on investigating different criteria that were tractable by a relaxation approach, later papers have also explored the problem of obtaining cluster memberships from the relaxed solution. Notably, Yu and Shi<sup>9</sup> noticed that the solution of the relaxed multiway normalized cut optimization problem is not unique and proposed an iterative algorithm to find a new relaxed solution as close as possible to a solution of the original discrete normalized cut problem.

In the physics community, Newman<sup>8</sup> maximized his modularity criterion using an eigendecomposition of the modularity matrix. Methods related to spectral decomposition of the connectivity matrix have also been applied to cluster detection in multidimensional time series analysis. Bialonski and Lehnertz<sup>10</sup> used the empirically defined *participation index*<sup>11</sup> to find phase synchronization clusters in a lattice of Lorenz oscillators. Allefeld and Bialonski<sup>12</sup> proposed a method for detecting synchronization clusters using an approach based on Markov random walks. Azran and Ghaharmani<sup>13</sup> also based their method on Markov random walks and suggested a procedure to construct a hierarchical decomposition of a dataset using different path lengths of the random walk. Angelini *et al.*<sup>14</sup> constructed a method for the

<sup>a)</sup>Electronic mail: vejmelka@cs.cas.cz.

identification of modular structures based on the average association objective (also called the ratio association objective)<sup>6,15</sup> in the context of probabilistic autoencoders partitioning the network graph into modules.

In this work, we propose a method for identifying clusters (sets of elements with high mutual connectivity) in undirected networks based on maximization of the average association objective. Unlike existing methods, the proposed method is not constrained to cover the entire set of elements with clusters. Instead, each element is assigned to at most one cluster. The vertices not assigned to any cluster will be referred to as the *residual set*. The framework allows for an accurate identification of clusters even in the presence of elements that may be unrelated to the underlying structure of the investigated network.

Clustering using the average association objective entails the solution of a constrained discrete optimization problem. The standard strategy for finding a satisfactory solution is a relaxation of the discrete cluster memberships to continuous ones. This relaxation strategy has proved itself effective in several of the graph partitioning approaches recounted in Sec. I. The theory introduced in this paper shows how the participation index is naturally derived from a simple strategy mapping a relaxed solution to an admissible solution of the original discrete optimization problem. The support of an explicit mathematical framework also allows for the formulation of more effective methods of mapping a relaxed solution to a discrete solution. We suggest a method involving the variance maximizing rotation VARIMAX<sup>16</sup> and show its superior efficiency compared to the simple mapping strategy.

In Sec. II the theoretical framework of the clustering method based on the average association objective is presented. In Sec. III numerical studies illustrating the properties of the proposed method are presented as well as an example result from an experimentally obtained resting state functional magnetic resonance imaging dataset from a human subject. In Sec. IV some open questions are considered and the paper is summarized in Sec. V.

## II. METHODS

In this section the theoretical framework of the network partitioning method (or clustering method) based on the average association objective is formulated. Let  $V$ ,  $|V|=N$  be the set of objects  $p_i$  to be clustered. The connectivity (association, similarity) between each pair of objects  $p_i, p_j \in V$  is given by a non-negative number  $w_{i,j}$ . We require that the self-connectivity of an element is 0  $\forall p_i \in V: w_{i,i}=0$ . Single objects will thus not be able to form clusters that contribute to the objective. This scheme corresponds to a network with vertices corresponding to the objects and undirected weighted edges with weights given by the connectivity function.

Numerical studies in this work will deal with dynamical systems represented by their time series. To simplify later discussion, we will refer to the objects as processes and their connectivity will be denoted by  $w_{i,j}$ . The clustering method divides the set of processes  $V$  into  $K$  clusters or sets  $V_k \subset V$ ,  $k \in \{1, 2, \dots, K\}$  which are disjoint and their union

is a subset of  $V$ . The remaining set of objects  $V_o = V \setminus \bigcup_k V_k$  is the residual set. The sets  $V_1, V_2, \dots, V_K, V_o$  together form a partition  $\mathcal{V}$  of the set  $V$ . The number of processes in each cluster will be denoted by  $N_k = |V_k|$ . Given the number of clusters  $K$ , the clusters are found by maximizing the average association objective

$$J_K = \sum_{k=1}^K \sum_{p_i, p_j \in V_k} \frac{w_{i,j}}{N_k} \quad (1)$$

over all partitions  $\mathcal{V}$ . The objective can be interpreted as a sum of cluster strengths  $S_k$  for each cluster  $V_k$ ,

$$S_k = \sum_{p_i, p_j \in V_k} \frac{w_{i,j}}{N_k}. \quad (2)$$

The average association objective rates highly partition where each of the sets  $V_1, V_2, \dots, V_K$  contains processes which manifest high mutual connectivity. If a cluster  $V_k$  contains  $N_k$  coupled processes with mutual connectivity  $\rho$ , then the cluster strength of the cluster would be

$$S_k = \frac{\rho(N_k - 1)N_k}{N_k} = \rho(N_k - 1). \quad (3)$$

The cluster strength is linear with respect to the mutual connectivity if the number of processes in the cluster is held constant and increases linearly with the number of processes in the cluster if the connectivity is held constant. The cluster strength of a singleton cluster is zero.

We investigate the behavior of the average association objective on a very simple model of  $N$  processes with constant connectivity  $\rho$  among themselves. There is no structure in the model and thus no inherent clusters.

If a single cluster is sought by maximizing the objective  $J_1$ , then none of the processes should be split off into the residual set. Here the value of the average association objective only depends on the number of processes in the single cluster. The value of the objective  $J_1$  can be analytically calculated as

$$J_1 = S_1 = \sum_{p_i, p_j \in V_1} \frac{w_{i,j}}{N_1} = \frac{\rho(N_1 - 1)N_1}{N_1} = \rho(N_1 - 1). \quad (4)$$

The objective increases linearly with the number of processes in the cluster  $V_1$  and thus reaches its maximum value for  $|V_1|=|V|=N$ . The optimal solution is indeed a clustering with all the processes in one cluster and an empty residual set.

To verify if there is any preferred relative cluster size, the objective  $J_2$  is evaluated for the same model. Since there are no clusters in the model itself, any preferred relative size of the clusters would have to be a result of the structure of the objective function (1). The objective  $J_2$  again only depends on the sizes of the two clusters

$$J_2 = \sum_{k=1}^2 \sum_{p_i, p_j \in V_k} \frac{w_{i,j}}{N_k} = \rho(N_1 + N_2 - 2). \quad (5)$$

The objective  $J_2$  does not depend on the relative size of the clusters but only on the sum of their sizes. The optimal result in this case is that the union of the two clusters is the input

set  $V$ , while the division of processes into the clusters is arbitrary. There is no preferred relative size of clusters obtained by maximizing the average association objective on this model. This is not the case, for example, for the min-max-cut criterion,<sup>17</sup> which favors balanced clusters and would rate highly a division into (arbitrary) clusters of similar sizes.

The constant connectivity model can also be used to show why a self-connectivity of 0 is a reasonable requirement. Assuming that instead of zero, the self-connectivity would be equal to  $\rho_s > 0$ , we analyze again the constant connectivity model. By construction, a single cluster covering all the processes is a correct result, while a division of the set into multiple clusters covering the entire set is necessarily arbitrary. The objective value for a single cluster should thus be higher than that for two or more clusters to ensure that the former solution is preferred over the latter. If the self-connectivity is  $\rho_s$ , we can write the average association objective  $J_1$  for one cluster and  $J_K$  for  $K > 1$  clusters

$$J_1 = \frac{\rho(N-1)N + N\rho_s}{N} = \rho(N-1) + \rho_s, \quad (6)$$

$$J_K = \sum_{k=1}^K [\rho(N_k - 1) + \rho_s] = J_1 + (K-1)(\rho_s - \rho)$$

since  $\sum_{k=1}^K N_k = N$ . Our requirement that  $J_1 > J_K$  is equivalent to  $\rho_s < \rho$ . This should be true regardless of the value of  $\rho > 0$ , which uniquely determines the value of the self-connectivity as 0.

Further discussion is facilitated by reformulating the average association objective in matrix form. Let  $W = [w_{i,j}]_{N \times N}$  be the symmetric connectivity matrix and let the indicator vectors  $\mathbf{u}_k, k \in \{1, 2, \dots, K\}$  be defined by

$$[\mathbf{u}_k]_i = \begin{cases} 1 & \text{if } p_i \in V_k \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

The disjointness of the sets  $V_k$  requires that the vectors  $\mathbf{u}_i$  and  $\mathbf{u}_j$  are orthogonal for  $i \neq j$ . The objective  $J_K$  can be written in terms of the indicator vectors  $\mathbf{u}_k$  as

$$J_K = \sum_{k=1}^K \frac{\mathbf{u}_k^T W \mathbf{u}_k}{\mathbf{u}_k^T \mathbf{u}_k}. \quad (8)$$

The objective is to be maximized under the conditions  $\mathbf{u}_k \in \{0, 1\}^N$  and  $\mathbf{u}_i^T \mathbf{u}_j = 0$  if  $i \neq j$ . Equation (8) can be rewritten as a matrix trace by accumulating the vectors  $\mathbf{u}_k$  into a matrix  $U = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_K)$ . We can then write for the objective

$$J_K = \text{tr}\{(U^T U)^{-1} U^T W U\} = \text{tr}\{(U^T U)^{-1/2} U^T W U (U^T U)^{-1/2}\}, \quad (9)$$

where matrix  $U^T U$  is diagonal. The substitution  $Y = U(U^T U)^{-1/2}$  simplifies the optimization problem to  $J_K = \text{tr}\{Y^T W Y\}$ . The condition  $Y^T Y = I_K$  is automatically satisfied since

$$Y^T Y = (U^T U)^{-1/2} (U^T U) (U^T U)^{-1/2} = I_K. \quad (10)$$

The vectors  $\mathbf{y}_k$  thus have unit length and are orthogonal to each other. Their elements are

$$[\mathbf{y}_k]_i = \begin{cases} \frac{1}{\sqrt{N_k}} & \text{if } i \in V_k \\ 0 & \text{otherwise,} \end{cases} \quad (11)$$

where  $N_k = |V_k|$  is the size of the cluster  $V_k$ . This discrete optimization problem can be relaxed by dropping the requirement (11) and allowing the elements of  $Y$  to have continuous values while respecting the orthogonality and unit size constraints.

### A. Relaxed optimization problem

The relaxed optimization problem can be written in terms of the matrix  $Y$  as

$$\max_{Y^T Y = I} \text{tr}\{Y^T W Y\}. \quad (12)$$

The maximum for this problem is attained if  $\mathbf{y}_k$  are the eigenvectors of  $W$  which correspond to the  $K$  largest eigenvalues of the connectivity matrix  $W$ . The eigenvalues of the matrix  $W$  will be denoted as  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$  and the eigenvector corresponding to  $\lambda_k$  will be denoted as  $\mathbf{z}_k$ . Then the optimal solution of Eq. (12) is the matrix  $Z = (\mathbf{z}_1, \dots, \mathbf{z}_K)$ .

Each eigenvector  $\mathbf{z}_k$  represents a relaxed cluster in which element memberships are relaxed to continuous values. The cluster strength of such a relaxed cluster is equal to its corresponding eigenvalue

$$S_k = \frac{\mathbf{z}_k^T W \mathbf{z}_k}{\mathbf{z}_k^T \mathbf{z}_k} = \lambda \frac{\mathbf{z}_k^T \mathbf{z}_k}{\mathbf{z}_k^T \mathbf{z}_k} = \lambda_k. \quad (13)$$

The optimal value of the objective for the relaxed version of the problem is the sum of the cluster strengths

$$J_K = \sum_{k=1}^K \lambda_k. \quad (14)$$

Clearly, relaxed clusters with a negative cluster strength  $S_k = \lambda_k < 0$  would decrease the value of the criterion and cannot form a part of the relaxed solution. The contribution of each element to the average association inside cluster  $V_k$  can be written as

$$J_K = \sum_{k=1}^K \frac{\mathbf{z}_k^T W \mathbf{z}_k}{\mathbf{z}_k^T \mathbf{z}_k} = \sum_{k=1}^K \mathbf{z}_k^T W \mathbf{z}_k = \sum_{k=1}^K \lambda_k \mathbf{z}_k^T \mathbf{z}_k = \sum_{k=1}^K \sum_{i=1}^N \lambda_k z_{i,k}^2, \quad (15)$$

where  $z_{i,k}$  is the element at position  $(i, k)$  of the matrix  $Z$ . Note that the expression  $\lambda_k z_{i,k}^2$  is the participation index.<sup>11</sup> If the method was required to assign every process to exactly one cluster, the simplest way to maximize the objective (1) based on the obtained relaxed solution would be to assign each element  $p_i$  to the cluster  $V_k$  for which  $\lambda_k z_{i,k}^2$  is maximal over all  $k \in \{1, 2, \dots, K\}$ . The participation index  $\lambda_k z_{i,k}^2$  is the contribution of the element  $p_i$  to the final value of the objective  $J_K$  assuming that it is assigned to cluster  $V_k$ . In Ref. 10, processes were assigned to clusters in which their participation indices were maximal. Clearly, this is a reasonable straightforward way of maximizing the average association criterion based on the obtained relaxed solution assuming that each process must be assigned to exactly one cluster. In

the proposed method, not all processes are required to be parts of clusters, and thus the cluster assignment procedure is slightly more involved.

The above procedure of assigning elements to clusters is unable to cope with clusters of similar cluster strength. This problem has been noticed in numerical experiments<sup>10</sup> and the authors postulated that a “cluster strength” considering both connectivity and cluster size exists and clusters that have a similar strength will not be separated correctly in some configurations. We demonstrate in the experiments using theoretical connectivity models (cf. Sec. III A) that the cluster strength (2) has this property. If there are two or more clusters with similar cluster strengths, then cluster assignment based on the participation index will not be able to discriminate between them. However, the cluster assignment incorporating the VARIMAX rotation discriminates between clusters with equal or similar cluster strengths effectively.

## B. Cluster assignment

Cluster assignment is a procedure that maps an admissible solution of the relaxed problem (12) to an admissible solution of the original discrete clustering problem (1). For a problem with  $N$  processes and  $K$  clusters, cluster assignment may formally be specified as a function  $M_X: R^{N \times K} \rightarrow \{V_1, \dots, V_K, V_o\}$ , where the domain of the method is a set of matrices with orthonormal columns and the range is the set of partitioning of the set of processes  $V$ . Two methods of assigning clusters are presented here: the first will be referred to as simple mapping  $M_S$  and the second as VARIMAX mapping  $M_V$ .

### 1. Simple mapping

Simple mapping  $M_S$  has two phases. In the first phase, candidate elements are assigned to clusters. This is performed by assigning each element  $i$  to the cluster  $k$  such that

$$k = \arg \max_l \lambda_l z_{i,l}^2, \quad (16)$$

in other words by maximizing the participation index. This is motivated by the decomposition of the objective value of the relaxed problem (15), which is maximized by this strategy. The process  $i$  is assigned to the cluster  $V_k$ , where it will contribute most to the final value of the average association objective (1).

Using this assignment, a new matrix  $\tilde{Z} = [\tilde{z}_{i,j}]_{N \times K}$  is created as

$$\tilde{z}_{i,k} = \begin{cases} |z_{i,k}| & \text{if } p_i \in V_k \\ 0 & \text{otherwise.} \end{cases} \quad (17)$$

Each column  $\tilde{z}_l$  of  $\tilde{Z}$  is then rescaled to unit length. Finally, the best approximation in the least-squares sense of  $\tilde{z}_l$  by a discrete indicator vector of the form (11) is found. This can be done efficiently by sorting  $\tilde{z}_l$  in descending order and systematically computing the sum of squares distance to a vector  $\mathbf{d}(m) \in R^N$  of the form

$$[\mathbf{d}(m)]_i = \begin{cases} \frac{1}{\sqrt{m}} & \text{if } i \leq m \\ 0 & \text{if } i > m. \end{cases} \quad (18)$$

After sorting, the best matching indicator vector can be found in linear time. All but the  $m$  processes with the highest values of  $\tilde{z}_{i,k}$  are then removed from the cluster  $V_k$ . This procedure is repeated for each cluster  $V_k$ . The removed processes from each cluster constitute the residual set  $V_o$  of the clustering.

### 2. VARIMAX mapping

Yu and Shi<sup>9</sup> noticed that any orthogonally transformed solution of the relaxed multiway normalized cut problem is also a solution of the same problem. The insight is also applicable to the relaxed solution of the average association problem (12): given the relaxed solution  $Z$ , any matrix  $ZR$ , where  $R$  is an orthogonal matrix ( $R^T R = I$ ), is also a solution. This is easily seen as the objective value in Eq. (12) is in the form of a matrix trace, which is invariant under orthogonal transformation, and since  $(ZR)^T (ZR) = R^T Z^T Z R = R^T I R = I$ , the orthogonality constraint is automatically satisfied for  $ZR$ .

We suggest using the VARIMAX (Ref. 16) rotation which has been applied in factor analysis and principal component analysis. Application of the VARIMAX procedure results in an orthogonal matrix  $R_V \in \mathbb{R}^{K \times K}$  such that

$$R_V = \arg \max_R K \sum_{k=1}^K \left( \frac{1}{K} \sum_{i=1}^N ([ZR]_{i,k}^2 - \overline{[ZR]_{*,k}^2})^2 \right), \quad (19)$$

where  $[ZR]_{i,k}$  is the element at position  $(i,k)$  of the matrix  $ZR$  and  $[ZR]_{*,k}$  is the  $k$ th column vector of the same matrix. According to the criterion (19), the orthogonal matrix  $R_V$  is selected to maximize the sum of the variances of the squared elements in each column of  $ZR_V$ . The optimization algorithm for obtaining  $R_V$  is iterative, deterministic, and makes use of the optimal solution for the VARIMAX objective in two dimensions.<sup>16,18</sup>

The VARIMAX rotation attempts to find a suitable orthogonal transform  $R_V$  so that  $Z_V = ZR_V$  has a simple structure.<sup>19</sup> The principle of simple structure requires that each column vector of the matrix  $Z_V$  should have some elements with high absolute values (loadings) and all the other elements should have values close to zero. A relaxed solution with simple structure is more likely to correspond to the form of the indicator vector (11). In other words, such a relaxed solution may be closer to a solution of the discrete constrained optimization problem. Intuitively, the requirement of simple structure corresponds to a compact localization of a relaxed cluster on the input set  $V$ .

When the rotated matrix  $Z_V = [z_{i,j}]_{i,j}$  is obtained, candidate processes  $p_i$  are assigned to clusters  $V_k$  so that

$$k = \arg \max_{l \in \{1, \dots, K\}} [z_{i,l}]. \quad (20)$$

Each column of matrix  $Z_V$  is normalized to unit length and the discrete indicator vector template (18) is fit exactly as in simple mapping to obtain the final composition of each cluster.

### III. RESULTS

This section reviews numerical studies of theoretical connectivity models which illustrate selected aspects of the clustering method. The effectivity of the method is also tested using a partially coupled lattice of Lorenz oscillators. The models are identical to those of Bialonski and Lehnertz<sup>10</sup> and some results are also replicated for comparison purposes. Finally, the method is applied to a high-dimensional functional magnetic resonance imaging (fMRI) dataset and the results are discussed.

#### A. Theoretical connectivity models

We replicate the numerical experiments using a simulated connectivity matrix described in Ref. 10 and, in addition, we demonstrate how empirical observations therein can be understood within the context of the presented mathematical framework.

The connectivity matrices were constructed to simulate a population of 32 coupled systems in two clusters with constant connectivity levels inside each cluster and between clusters. Connectivity matrices with this type of simplified structure would be difficult to obtain from simulations involving coupled dynamical systems and were therefore constructed directly.

Each connectivity matrix contained samples from three random variables characterizing the connectivity in and between two clusters  $V_1 = \{p_1, p_2, \dots, p_r\}$  and  $V_2 = \{p_{r+1}, p_{r+2}, \dots, p_{32}\}$ . The connectivity inside cluster  $V_1$  was sampled from the normal distribution  $\mathcal{N}(\rho_1, \sigma_1)$  in cluster  $V_2$  from  $\mathcal{N}(\rho_2, \sigma_2)$  and between the two clusters from  $\mathcal{N}(\rho_{\text{int}}, \sigma_{\text{int}})$ . The standard deviation for each random variable was computed from the mean phase coherence population value  $\rho$  using the formula  $\sigma = (1 - \rho^2) / \sqrt{2n}$  with the sample size  $n = 200$ .

In Ref. 10, no cluster assignment was explicitly performed and the error of the clustering was objectively estimated using a function of the eigenvectors corresponding to the two largest eigenvalues. Let the functions  $\Delta_1$  and  $\Delta_2$  be given by

$$\begin{aligned} \Delta_1 &= \left( \sum_{i=1}^r \Theta(\xi_{i2}) \right)^{-1} \sum_{i=1}^r \xi_{i2} \Theta(\xi_{i2}), \\ \Delta_2 &= \left( \sum_{i=r+1}^{32} \Theta(\xi_{i1}) \right)^{-1} \sum_{i=r+1}^{32} \xi_{i1} \Theta(\xi_{i1}), \end{aligned} \quad (21)$$

where  $\Theta(\cdot)$  is the Heaviside function and the quantities  $\xi_{i1} = -\xi_{i2} = \lambda_1 z_{i1}^2 - \lambda_2 z_{i2}^2$  are the differences in participation indices corresponding to element  $p_i$  and cluster  $V_1$  and  $V_2$ , respectively. The function  $\Delta_1$  is the average difference of participation indices for elements that belonged to cluster  $V_1$  but were incorrectly assigned to  $V_2$  and the function  $\Delta_2$  is the average difference of participation indices for the converse situation. If the participation indices signify the correct cluster memberships, then  $\Delta_1 = \Delta_2 = 0$ . The total clustering error was quantified by

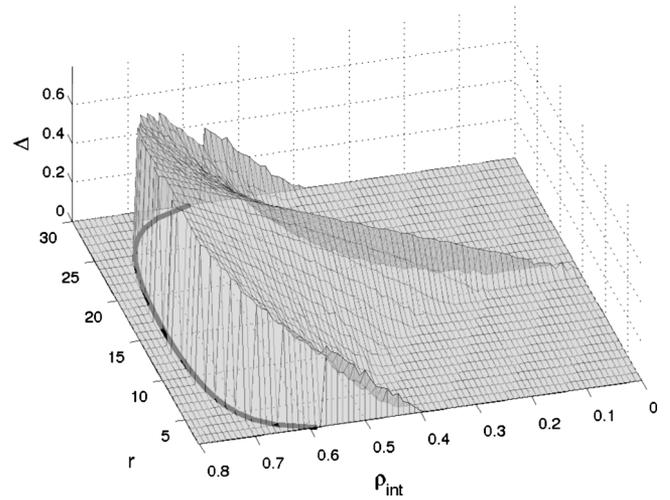


FIG. 1. The error  $\Delta$  as a function of the size  $r$  of cluster  $V_1$  and the mean connectivity between clusters  $\rho_{\text{int}}$ . The thick line is the theoretical prediction of the boundary between the region with two positive eigenvalues and the region with only one positive eigenvalue.

$$\Delta = \begin{cases} \Delta_1 + \Delta_2 & \text{if } N_\lambda = 2 \\ 0 & \text{otherwise,} \end{cases} \quad (22)$$

where  $N_\lambda$  was the number of eigenvalues of the connectivity matrix greater than 1. In the presented framework, the self-connectivity is required to be 0 instead of 1.<sup>10</sup> The eigenvectors do not depend on the self-connectivity but the eigenvalues are all decreased by 1 when the self-connectivity is 0. Therefore, in this experiment  $N_\lambda$  indicates the number of eigenvalues greater than 0.

#### 1. Influence of intercluster connectivity on clustering solution

The first experiment examined a transition from a two cluster scenario to a single cluster. The mean intracluster connectivities were set to  $\rho_C = \rho_1 = \rho_2 = 0.8$  and the mean intercluster connectivity  $\rho_{\text{int}}$  was increased in steps of 0.01 from 0.0 to 0.8. For  $\rho_{\text{int}} = 0$ , there are two clusters  $V_1$  and  $V_2$  with internal connectivity  $\rho_C = 0.8$ , and for  $\rho_{\text{int}} = 0.8$ , there is only one cluster  $V_1 \cup V_2$  covering the entire set of processes. Clearly, there must be a boundary in the space parametrized by  $r = |V_1|$  and  $\rho_{\text{int}}$  separating the region where one cluster exists from the region where two clusters exist.

The purpose of this experiment was twofold: to show the effectivity of the VARIMAX rotation in determining the correct cluster assignments and to show that the boundary between one- and two-cluster regions can be analytically calculated from the proposed theory. The boundary location agrees with the spectral properties of the model connectivity matrix.

The error surface  $\Delta$  (22) for the given parameter space is shown in Fig. 1. Although the self-connectivity was equal to 1 in the numerical studies of Bialonski and Lehnertz,<sup>10</sup> the error surface is practically unchanged when the self-connectivity is defined according to the proposed theory as 0. For similar cluster sizes  $r \approx 15$ , cluster assignment based on the participation index was unable to consistently separate the two clusters even if their internal connectivity was much

higher than the connectivity between them. The error surface sharply drops to zero for higher inter-cluster connectivities (region on the left of Fig. 1) because the number of positive eigenvalues of the connectivity matrix changes from two to one and the error function  $\Delta$  is zero by definition. Additionally, there were errors for higher intercluster connectivities when the cluster sizes were unbalanced.

These results have been compared to those computed from eigenvectors rotated using an orthogonal matrix obtained from the VARIMAX method. The full VARIMAX mapping  $M_V$  was not applied as here only the effect of the VARIMAX rotation is appraised. The error of the solution was quantified differently because the cluster assignment is computed without the involvement of the eigenvalues (cf. Sec. II B 2). To evaluate the cluster assignment error that would be made based on the rotated eigenvectors, the error contributions were defined as

$$\tilde{\xi}_{i1} = -\tilde{\xi}_{i2} = [ZR_V]_{i1}^2 - [ZR_V]_{i2}^2, \quad (23)$$

where  $R_V$  is the rotation computed by the VARIMAX method and  $Z$  is the matrix with columns containing the first two eigenvectors of the connectivity matrix. We denote the modified error function as  $\Delta_V$ . The modification preserves the spirit of the error function  $\Delta$  for the rotated eigenvectors. The error function  $\Delta_V$  is zero everywhere in the parameter space of the experiment (figure not shown). Hence, assigning cluster memberships based on the rotated eigenvectors would result in correct assignments in the entire parameter space.

Using this idealized example, it is possible to analytically compute the expected value  $E[J_1]$  for the situation where all the elements are in one cluster and  $E[J_2]$  for the two model clusters, both with empty residual sets,

$$E[J_1] = (N-1)\rho_C - \frac{2r(N-r)(\rho_C - \rho_{\text{int}})}{N}, \quad (24)$$

$$E[J_2] = \rho_C(r-1) + \rho_C(N-r-1) = (N-2)\rho_C.$$

For each cluster size parametrized by  $r$ , a critical intercluster connectivity  $\rho_{\text{int}}^{\text{crit}}(r)$  may be computed from the condition that the expected values of the criteria  $J_1$  and  $J_2$  are equal, which results in the equation

$$\rho_{\text{int}}^{\text{crit}}(r) = \rho_C \left[ 1 - \frac{N}{2r(N-r)} \right]. \quad (25)$$

The critical intercluster connectivity  $\rho_{\text{int}}^{\text{crit}}(r)$  separates two regions where the optimal clustering differs. In the region where  $\rho_{\text{int}} > \rho_{\text{int}}^{\text{crit}}(r)$ , the theoretically optimal clustering is a single cluster covering the entire set of elements, with an empty residual set. In the region where  $\rho_{\text{int}} < \rho_{\text{int}}^{\text{crit}}(r)$ , the optimal split is into two clusters  $V_1$  and  $V_2$  again with an empty residual set.

The curve  $\rho_{\text{int}}^{\text{crit}}(r)$  is indicated by a thick black line in Fig. 1 and coincides exactly with the boundary separating the region where the model connectivity matrix has one positive eigenvalue from that with two positive eigenvalues. This boundary is formed by the locations where the  $\Delta$  error func-

tion (22) sharply drops to zero, as by definition the error function  $\Delta$  is only nonzero if the connectivity matrix has exactly two positive eigenvalues.

We have shown that the optimal number of clusters according to the proposed theoretical framework is equal to the number of positive eigenvalues of the model connectivity matrix. Thus, the curve  $\rho_{\text{int}}^{\text{crit}}(r)$  is not an arbitrary division of the parameter space into a one- and two-cluster regions but is in agreement with the spectral properties of the connectivity matrix.

Finally, we note that the clustering solutions actually obtained by running the clustering method with the full VARIMAX mapping (cf. Sec. II B 2) result in an accurate clustering on the entire parameter space, with an empty residual set everywhere and a correct assignment into one or two clusters depending on the theoretically optimal clustering according to the expected values of the objectives  $J_1$  and  $J_2$ .

## 2. Influence of relative cluster strength on clustering solution

In the second theoretical experiment, the cluster memberships were again determined by a single parameter  $r$  as in the previous experiment and the connectivities were set  $\rho_1 = 0.8$ ,  $\rho_{\text{int}} = 0.2$ , and  $\rho_2$  was decreased in steps of 0.01 from 0.8 to 0.2. This experiment simulated the transition from a two-cluster configuration to a single cluster configuration by decreasing the connectivity inside the second cluster.

The purpose of the second numerical study is to show that the theoretical cluster strengths (2) of the model clusters predict where cluster assignment based on the participation index will fail and that the VARIMAX rotation prevents this cluster assignment failure effectively.

Let the expected cluster strength be quantified according to the formula (2) derived in the average association framework as

$$E[S_1] = \rho_1(r-1), \quad (26)$$

$$E[S_2] = \rho_2(N-r-1).$$

We can now predict where the participation index will confuse the two model clusters. The region where the participation index method produces assignment errors should be near the curve where the expected cluster strengths  $E[S_1]$  and  $E[S_2]$  are equal. This curve, parametrized by the size of the cluster  $V_1$ , is given by the equation

$$\rho_2^{\text{crit}}(r) = \rho_1 \frac{r-1}{N-r-1}. \quad (27)$$

The  $\Delta$  error surface parametrized by the size  $r$  of cluster  $V_1$  and the connectivity  $\rho_2$  of the cluster  $V_2$  is shown in Fig. 2 and is in agreement with the previously reported results.<sup>10</sup> The curve  $\rho_2^{\text{crit}}(r)$  is indicated by a thick black line in the figure. The largest error  $\Delta$  is directly above the theoretical curve. Thus, for parameters  $r$ ,  $\rho_2$  which produce clusters of similar cluster strengths  $S_k$  developed in the proposed mathematical framework, the clustering method based on the participation index fails to correctly discriminate between the clusters.

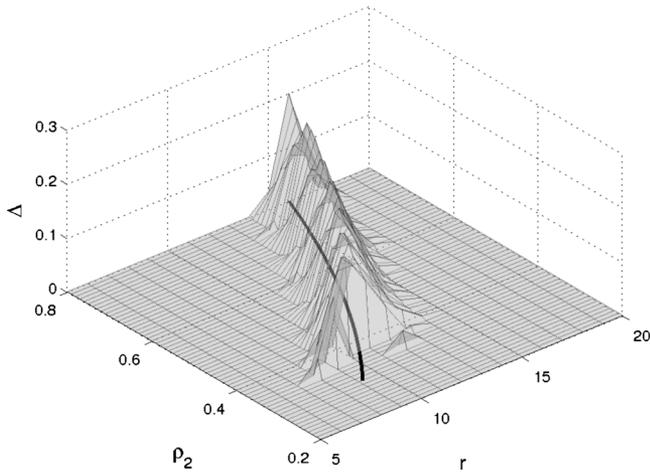


FIG. 2. The  $\Delta$  error surface for the participation index as a function of  $r$ , the size of the model cluster  $V_1$ , and of  $\rho_2$ , the internal connectivity of model cluster  $V_2$ . The thick black curve indicates the theoretical prediction of the region where the participation index method assigns cluster memberships incorrectly.

We wish to make a clear distinction between the cluster strengths of the underlying clusters, which may be estimated analytically from the model, and the cluster strengths of the relaxed clusters, which are equal to the corresponding eigenvalues of the connectivity matrix. In general, it is not possible to infer if the underlying clusters of similar strength are present in the investigated problem from the eigenvalue spectrum computed from a given connectivity matrix. For example, in the current problem, the largest eigenvalue was approximately in the range  $\langle 6.7, 21.1 \rangle$  and the second largest belonged to the interval  $\langle 1.6, 8.7 \rangle$ . The differences between the two largest eigenvalues over the parameter space never dropped below 4.9 (cf. Fig. 3) even when the model cluster strengths were almost equal. Hence, it is important to resolve the problem of discriminating clusters with similar strength at the method level as it may be difficult to find out whether or not the obtained solution is perturbed by the existence of such clusters.

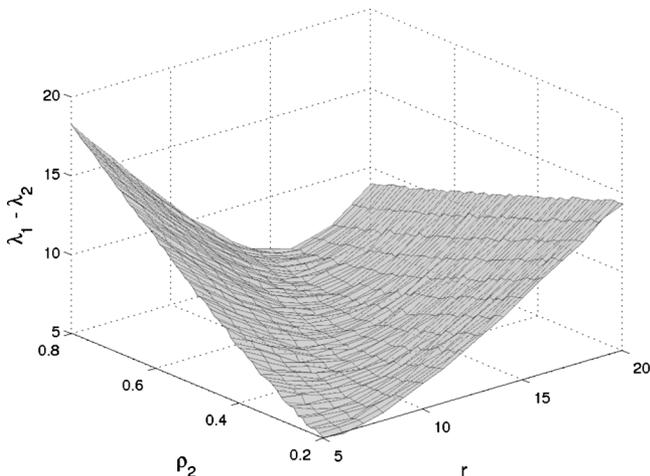


FIG. 3. The difference between the first and the second largest eigenvalues of the connectivity matrix as a function of  $r$ , the size of the model cluster  $V_1$ , and of  $\rho_2$ , the internal connectivity of model cluster  $V_2$ .

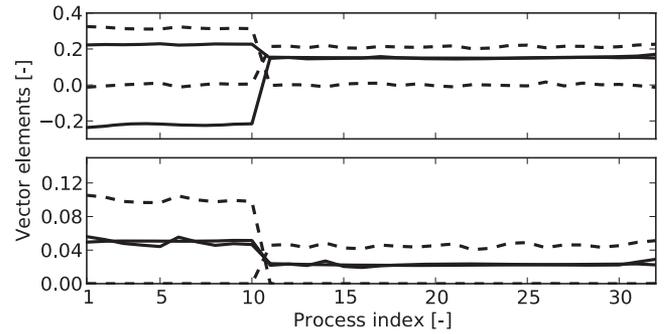


FIG. 4. Top: the first two eigenvectors (full line) and the same eigenvectors rotated by the VARIMAX method (dashed line). Bottom: squared elements of the same vectors. The parameters for the experiment were  $r=10$  and  $\rho_2=0.34$ . The VARIMAX-rotated eigenvectors are close to zero for processes outside their cluster.

The VARIMAX method separates clusters with similar cluster strength effectively. It rotates the eigenvectors using an orthogonal matrix  $R_V$  so that the sum of the variance of the squared entries of the rotated vectors is maximal. An example of the solution for  $r=10$  and  $\rho_2^{\text{crit}}=0.34$ , which produces model clusters with expected strengths  $S_1=7.2$  and  $S_2=7.14$ , is shown in Fig. 4. The eigenvectors obtained from the decomposition of the matrix do not have any elements close to zero, although the structure indicating the two clusters is clearly visible. On the other hand, the rotated eigenvectors clearly indicate that the two clusters and additionally their elements belonging to processes outside the indicated cluster are close to zero.

The error surface  $\Delta_V$ , which quantifies the assignment error based on the VARIMAX-rotated eigenvectors, is zero everywhere in the parameter space. Hence, it was demonstrated that use of the VARIMAX rotation resolves the problem of identifying clusters of similar strengths and results in a correct cluster assignment. The formal cluster strength (2), which was introduced within the average association framework, was shown to predict the regions where cluster assignment using the participation index is unable to separate the two model clusters.

**B. Lorenz lattice model**

This section details results from experiments on a lattice of coupled Lorenz oscillators.

**1. Error quantification**

To objectively measure the quality of the clustering from the average association objective, we employed the F-measure, recall and precision,<sup>20</sup> well-known functions in the domain of information retrieval. The functions allow for the quantitative comparison of a clustering result to the correct solution which is known for the Lorenz lattice model analyzed here. Given a set indicating a model or true cluster  $A$  and an identified cluster  $B$ , we can define the precision  $P$  and recall  $R$  as

$$P(A,B) = \frac{|A \cap B|}{|B|}, \quad R(A,B) = \frac{|A \cap B|}{|A|}. \quad (28)$$

Recall and precision are values with range  $\langle 0, 1 \rangle$  with 0 indicating that the model cluster and the identified cluster are disjoint. A precision of 1 indicates that the identified cluster is a subset of the model cluster. A recall of 1 indicates that the model cluster is a subset of the identified cluster. If both recall and precision are equal to 1, then the clusters  $A$  and  $B$  are identical.

The F-measure is a harmonic mean of the precision and the recall between a model and identified cluster. Using recall and precision, the F-measure can be written as

$$F(A,B) = \frac{2P(A,B) \times R(A,B)}{P(A,B) + R(A,B)}. \quad (29)$$

If sets  $A$  and  $B$  are equal, then their F-measure is equal to 1 and if their intersection is empty then their F-measure is 0. The F-measure is used to match each model cluster to one of the identified clusters. A model cluster is matched to the identified cluster for which the F-measure to the model cluster is maximal. The matching is not exclusive: one identified cluster may be matched to zero or more model clusters. The residual set is excluded from the matching. The F-measure is used for single cluster matching because it effectively combines precision and recall into a single variable. The curves reported in the experiment are the average precision and average recall over all matched pairs.

If three of the identified clusters  $V_1, \dots, V_K$  were exactly equal to the model clusters  $C_1, C_2, C_3$ , then the average precision and recall would both be equal to 1. If two model clusters  $C_1, C_2$  were part of one identified cluster  $V_1$ , then both pairs  $C_1, V_1$  and  $C_2, V_1$  would have a recall of 1 but low precision (which would reduce the average precision). On the other hand, if model cluster  $C_1$  was split into two identified clusters  $V_1$  and  $V_2$ , then the precision of the pairs  $C_1, V_1$  and  $C_1, V_2$  would be 1 but the recall would be affected (reduced to 0.5 if the  $C_1$  was split evenly into  $V_1$  and  $V_2$ ). The average precision and recall thus indicate what type of errors the various clustering strategies are producing.

## 2. Lorenz lattice model

A lattice of 32 identical Lorenz oscillators with three clusters of six coupled oscillators identical to that used by Bialonski and Lehnertz<sup>10</sup> in simulated experiments will be studied as a numerical example. Each system  $i$  is defined by the equations

$$\begin{aligned} \dot{x}_i &= -8/3x_i + y_i z_i + \epsilon_i(x_{D_{1,2,3}} - x_i), \\ \dot{y}_i &= 28z_i - y_i - x_i z_i, \\ \dot{z}_i &= 10(y_i - z_i), \end{aligned} \quad (30)$$

where  $\epsilon_i$  is the coupling strength between the system  $i$  and one of three selected driving systems  $D_{1,2,3}$  (cf. Fig. 5). For all systems belonging to one of the clusters,  $\epsilon_i = \epsilon$  so that a single parameter controls the simulation. The remaining oscillators are uncoupled with  $\epsilon_i = 0$ . The Lorenz oscillators were integrated by the Runge–Kutta {4,5} scheme as imple-

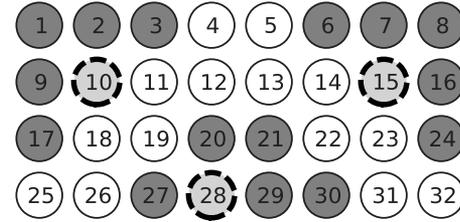


FIG. 5. Lattice of 32 Lorenz systems with three clusters indicated by neighboring gray-shaded circles. Light gray circles with dashed boundary indicate driving systems  $D_1=10$ ,  $D_2=15$ , and  $D_3=28$ , which drive the rest of the cluster to which they belong.

mented in MATLAB (function ode45) with a time-step  $dt=0.01$ . Starting the integration from random initial conditions, the first  $10^4$  samples were discarded to remove transient effects and time series of  $N=5 \times 10^5$  samples of the  $x$  variable were used for further analysis. The coupling parameter  $\epsilon$  was varied from 0 to 1.4 in steps of 0.05.

The connectivity between the systems was quantified by the mean phase coherence<sup>21</sup> (MPC) estimated by

$$\rho_{ij} = \left| \frac{1}{L} \sum_{l=1}^L e^{i(\phi_i(l) - \phi_j(l))} \right|, \quad (31)$$

where  $\phi_i(\cdot)$  is the phase of Lorenz system  $i$ . Mean phase coherence has the range  $\langle 0, 1 \rangle$  with 1 indicating perfectly phase synchronized systems. Phase was extracted using the analytical signal method by computing the imaginary part of the signal  $x_i(t)$  using the Hilbert transform defined as

$$(Hx_i)(t) = \frac{1}{\pi} \mathcal{P} \int_{-\infty}^{\infty} \frac{x_i(\tau)}{t - \tau} d\tau, \quad (32)$$

where  $\mathcal{P}$  indicates the Cauchy principal value of the integral. We have applied a more efficient method of computing the imaginary part given by

$$(Hx_i)(t) = \mathcal{F}^{-1}[(\mathcal{F}x_i)(\omega) \text{sgn}(\omega)](t), \quad (33)$$

where  $\mathcal{F}$  and  $\mathcal{F}^{-1}$  indicate the Fourier transform and its inverse, respectively. The average of each time series  $x_i(t)$  was removed, then the first and last 10% of the time series were tapered by a cosine half-wave. After computing the imaginary part of the analytical signal using Eq. (33), phase was extracted from the entire time series and the parts tapered by the cosine half-wave were discarded.

A key indicator of connectivity in the model systems is the mean phase coherence inside and outside the model clusters. The average MPC in the clusters and between the residuals as a function of  $\epsilon$  is shown in the left frame of Fig. 6. The average MPC inside the clusters grows rapidly for  $\epsilon > 0.4$ , while for weaker couplings, there is a smaller difference between the average MPC inside clusters and a set of six oscillators randomly selected from the residual set. The right frame of Fig. 6 shows the evolution of the five largest eigenvalues against the coupling parameter  $\epsilon$ . The top three eigenvalues show a very similar progression with respect to the coupling parameter.

We reiterate that the cluster eigenvalues are equal to the strength of the relaxed clusters (13) and that the cluster

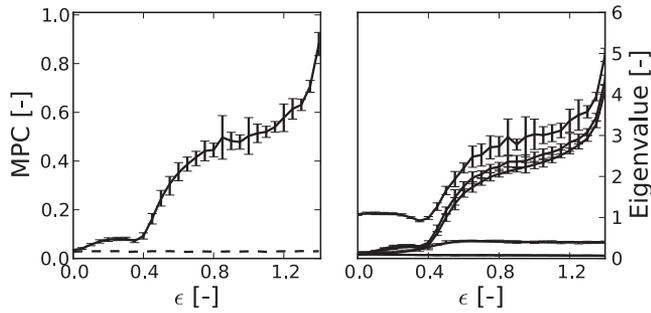


FIG. 6. Left: average mean phase coherence in the clusters (full line) and between uncoupled systems (dashed line) as a function of the coupling strength  $\epsilon$ . Error bars are standard deviations estimated from 20 realizations of the coupled system for each coupling strength  $\epsilon$ . The error bars of the connectivity of the uncoupled systems are too small to show. Only one curve is displayed for the three clusters as the coupling  $\epsilon$  is duplicated across clusters. Right: five largest eigenvalues vs coupling strength  $\epsilon$ . The fifth eigenvalue is very close to zero for all coupling strengths.

strength grows with the connectivity inside the cluster. Mean phase coherence has a nonlinear relationship with the coupling, as indicated by the left frame in Fig. 6. The evolution of the average MPC connectivity resembles that of the three largest eigenvalues closely with the exception of the top eigenvalue in the coupling range  $\epsilon < 0.4$ . The results of the simple mapping  $M_S$  (as discussed below) detect a single cluster covering the entire set of oscillators for this coupling range. This is in agreement with the other two eigenvalues being close to zero (cf. Fig. 6, right). Thus, the top eigenvalue reflects the strength of the single relaxed cluster covering the entire set. Although there are three model clusters, the eigenspectrum of the connectivity matrix does not reflect this configuration in the weak coupling range. In Sec. III it will be shown that nevertheless the three model clusters can be recovered even in the weak coupling range with use of the VARIMAX mapping.

### 3. Results

The experiment was in effect run twice for two different strategies of obtaining clusters. The first strategy identified clusters by maximizing the average association objective over cluster count  $K \in \{1, \dots, K_{\max}\}$  with  $K_{\max} = 10$ , which will be denoted as the max-K strategy. The cluster membership and the number of clusters were identified at the same time. In the second strategy the resulting clustering was obtained by maximizing  $J_3$  (the max-3 strategy), hence the number of clusters was constrained to at most three.

Twenty realizations of the time series from each oscillator were generated using the Lorenz lattice model for each connectivity strength  $\epsilon$  with random initial conditions. Each realization was clustered using four methods: simple mapping  $M_S$  and VARIMAX mapping  $M_V$ , both combined with the max-K and the max-3 strategy.

Figure 7 summarizes the results of the 580 runs: the left column contains average precision and recall curves for clustering results obtained using the max-K strategy and the right column contains results from the max-3 strategy. The top row shows results obtained by using the VARIMAX mapping and the center row by the simple mapping. The bottom

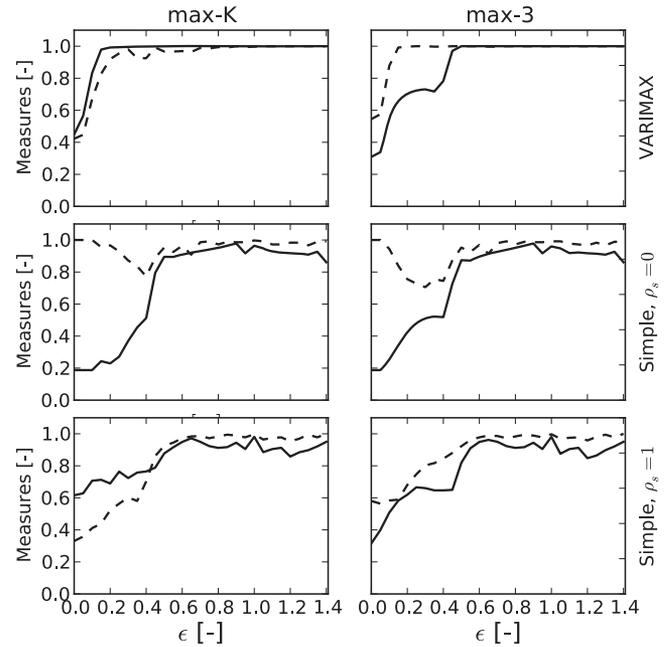


FIG. 7. Cluster matching results. The curves represent the average precision (full curve) and average recall (dashed curve). The results shown are averages over 20 realizations for each coupling strength  $\epsilon$ . Left column contains results for maximal objective clustering (max-K strategy); right column shows results for clustering using the max-3 strategy. The top row displays VARIMAX mapping (Sec. II B 2) results, center row simple mapping (Sec. II B 1) results, and the bottom row simple mapping with self-connectivity equal to 1 for comparison purposes.

row is shown for comparison purposes and applies simple mapping with self-connectivity equal to 1 (as in the previous study<sup>10</sup>).

The results for VARIMAX mapping are shown in the top row of Fig. 7: the results were markedly better than those obtained using simple mapping  $M_S$  (center and bottom rows). Of interest is the result in top left of Fig. 7: the model clusters were detected accurately already for very weak couplings  $\epsilon \geq 0.2$ , where the connectivity inside the clusters was quite low (cf. Fig. 6). Using the max-K strategy with the VARIMAX mapping, about six clusters were consistently obtained regardless of the value of the coupling term  $\epsilon$ . The residual set remained empty.

Surprisingly, even though that by construction the model has exactly three clusters, neither mapping combined with the max-3 strategy (curves in right column of Fig. 7) was able to identify the model clusters for weak couplings. For stronger couplings  $\epsilon \geq 0.5$ , the VARIMAX mapping (with the max-3 strategy) identified the model clusters accurately and all the uncoupled systems were assigned to the residual set. It seems that although it is possible to recover model clusters even for very weak couplings, more than three eigenvectors are required to do so. Examination of the clustering results indicated that if the number of clusters was constrained, oscillators from the residual set were affixed to the model clusters by the VARIMAX mapping, which reduced the precision of the method but did not adversely affect the recall. In the max-K strategy, the uncoupled oscillators formed very weak spurious clusters, which may have facilitated the precise identification of the model clusters.

When clusters were identified using the simple mapping involving the participation index (cf. Fig. 7, center row), the accuracy of the clustering suffered in two respects. For low couplings, the precision of the method was low as often one cluster contained most of the oscillators, while other clusters contained parts of the model clusters. This finding is consistent with the one dominant eigenvalue in this coupling range (cf. Fig. 6) as the relative size of the eigenvalues influences the participation indices. For very low couplings  $\epsilon < 0.2$ , all the oscillators were in one cluster, hence the average recall was close to 1. For slightly higher couplings  $0.2 \leq \epsilon \leq 0.5$ , the single cluster began splitting but the identified clusters did not match the model clusters and recall was negatively affected. For intermediate couplings, the clustering results were typically close to the model clusters but the accuracy did not match that of the VARIMAX mapping. For high couplings, the model clusters were often identified correctly but in some cases model clusters were lumped together in one identified cluster, this is visible as a slight decrease in precision for  $\epsilon > 1$ . An inspection of the results revealed that this is a manifestation of the inability to discriminate clusters with similar cluster strengths illustrated in theoretical connectivity models (Sec. III A). For such high couplings, the MPC between coupled systems is very close to 1 and all the model clusters have almost the same exact cluster strength (2). Thus, the simple mapping strategy was adversely affected for both low and high values of the coupling term  $\epsilon$ .

For comparison purposes, the results using an eigendecomposition of the MPC connectivity matrix with a unit diagonal and using simple mapping closely following previous experiments<sup>10</sup> are also shown. This would be equivalent to setting the self-connectivity to 1 in the presented framework. In this case, the maximum objective clustering resulted typically in ten or more clusters, some of which are composed of only two elements. By comparing the center and bottom row of Fig. 7 some differences can be seen. For low couplings  $\epsilon < 0.5$ , the precision is much higher for the unit-diagonal decomposition but the recall has dropped strongly, indicating that the dataset is split into multiple clusters which do not match the model clusters. For high couplings, the precision and recall are comparable with the results of the simple mapping strategy with zero self-connectivity (center row of Fig. 7).

The most important results of this numerical study are that the VARIMAX mapping resulted in clusters which match the model clusters accurately. Especially together with the max-K strategy, the model clusters were recovered even for weak coupling strengths. The VARIMAX mapping also accurately discriminates between clusters with similar cluster strength.

### C. Results from experimental fMRI data

The results obtained from the application of the proposed method to fMRI data measured from a resting human subject (eyes open fixation on a point) are reported. The measurements reflect the spontaneous blood oxygen level dependent activity in the subject.<sup>22</sup> The objective of the experiment is to identify functional units of the brain, which are commonly termed RSNs or intrinsic connectivity networks.

The measurement device was a 1.5 T MRI scanner. A time series was measured from each voxel ( $3 \times 3 \times 3$  mm<sup>3</sup> cube) inside the measurement volume with a repetition time TR=2 s. The analyzed segment contained 300 time points or 10 min of activity in 27 644 voxels, which were located in the gray matter of the subjects' brain. Before analysis, the dataset was subjected to standard preprocessing procedures. The data were aligned in the temporal and spatial domains and registered to the Talairach atlas.<sup>23</sup> The dataset was then spatially smoothed with a 6 mm full width half maximum Gaussian filter and bandpassed in the temporal domain in the 0.009–0.08 Hz frequency band. The time series of each voxel was orthogonalized with respect to selected sources of spurious variance: six parameters obtained by rigid body correction of head motion, whole-brain signal, signal from a ventricular region, and signal from a white matter region.

The connectivity function is based on the correlation coefficient between the time series of each voxel and reflects *a priori* knowledge about the problem. Resting state networks are rather loosely defined as regions exhibiting higher-than-average functional connectivity in the resting state. Some RSNs are recognizable as functionally specific regions of the brain which have been found using task-related studies.<sup>24</sup> Such studies have proceeded by first specifying a seed region in the brain and identifying voxels the time series of which were positively correlated with the time series of the seed region. This approach results in brain regions composed of voxels with mutually positively correlated time series.

Independent component analysis (ICA) has also been applied to detect RSNs.<sup>25,26</sup> ICA decomposition results in *maps*, some of which correspond to resting state networks and indicate their spatial location in the brain. Each element of the map quantifies how much the corresponding voxel is loaded by the RSN. Loadings may be positive or negative and RSNs are typically identified by selecting those voxels which have a positive loading larger than some threshold. This practice implies that the correlations between the projections of the time series corresponding to the RSN in each selected voxel are positive.

From the above, it seems that voxels inside one resting state network should have a pairwise positively correlated time series. Regions that correlate negatively to a given RSN may constitute another anticorrelated RSN.<sup>24</sup> Thus, in our analysis we have set any negative correlations to zero to prevent negatively correlated regions from falling into the same cluster.

The processing of the fMRI dataset required the eigendecomposition of a  $27\,644 \times 27\,644$  dense symmetric matrix which was performed on a distributed memory parallel computer using the message-passing linear algebra library ScaLAPACK.<sup>27</sup> Since the computation of eigenvectors and eigenvalues is time-consuming, only 50 largest eigenvalues and eigenvectors were computed. The value of the objective has not peaked even for 40–50 clusters and thus it was not possible to follow the max-K strategy for this dataset. Twenty-five clusters were extracted as a manageable number that could be visually scanned and labeled. Approximately 1200 voxels were assigned to the residual set.

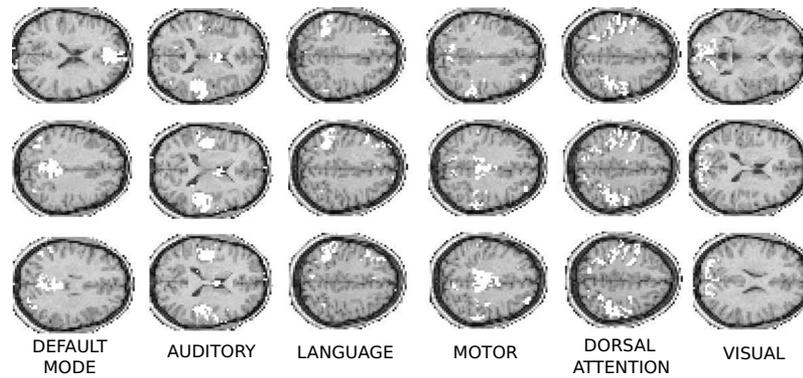


FIG. 8. Selected axial slices of the brain showing RSNs detected in the example fMRI dataset, one RSN in each row. Voxels that are members of the respective networks are white. Each RSN corresponds to one identified cluster except in the last row (default mode network), where the left slice is from a different cluster than the other two slices. The names of the respective resting state networks are displayed on each row.

An inspection of the results revealed that the localization of some clusters matched the previously known resting state networks.<sup>28</sup> Other clusters were clearly of artifactual nature and contained voxels at the edge of the gray matter or partially overlapping regions containing cerebrospinal fluid. Clusters corresponding to well-known resting state networks are shown on selected axial slices in Fig. 8. From top to bottom these are visual network, dorsal attention network, motor network, language network, auditory network, and default mode network.

Interestingly, the default mode network was split into an anterior part (left slice at the bottom row) and a posterior part (center and right image at the bottom row) and has thus been mapped into two clusters. When the number of clusters was reduced, the default mode network emerged as one cluster but some of the other resting state networks were combined into single clusters as well. We conclude that clusters identified by average association clustering of the fMRI data reflect the functional structure of the spontaneous activity of the human brain. Commonly reported RSNs can be matched to clusters at different levels of detail which may be obtained by changing the number of requested clusters  $K$ .

## IV. DISCUSSION

Here we discuss some open questions related to aspects of the proposed theoretical framework.

### A. Cluster assignment

One important open question is that of the optimality of the cluster assignment procedure. In spectral graph clustering approaches, the procedure is typically a heuristic and there is no guarantee that a good relaxed solution will be mapped to a good discrete solution (discrete cluster memberships). An important step in finding better discrete solutions was the realization that the relaxed solution of the multiway normalized cut objective was not unique<sup>9</sup> which also holds for the framework of average association clustering.

In this work, we have proposed using the VARIMAX rotation<sup>16</sup> which has been applied in psychology and social sciences to discover concise factors influencing observed variables. The VARIMAX objective (19) rates highly relaxed

solutions that have simple structure,<sup>19</sup> resulting in indicator vectors close to the form of the normalized discrete indicator vectors (18): some elements close to zero and some as large as possible while respecting the unit length constraint on the vector. This approach has been shown to be effective but improved approaches may be possible.

More generally, if the framework is further relaxed and the discrete solution is said to lie in the linear subspace spanned by the  $K$  largest eigenvectors, then it becomes possible to seek relaxed clusters which are not orthogonal to each other. This has also been explored in factor analysis and principal component analysis resulting in new methods and criteria such as OBLIMIN.<sup>29</sup> New cluster assignment procedures can be formulated, which exploit these possibilities. An open question is whether cluster assignments can be found which would provide tight bounds on the difference of the objective of the relaxed and discrete solutions.

### B. Relevant clusters

The proposed method accepts a connectivity matrix as input, a number of requested clusters  $K$ , and produces  $K$  clusters and a residual set (which may be empty). The clustering procedure finds the most prominent or salient subsets of the clustered data which have a sufficiently high internal connectivity. This raises two important issues: how to find clusters in problems where the number of clusters is unknown and what do the discovered clusters represent.

In the numerical study of the Lorenz lattice, spurious clusters not related to genuine coupling were returned by the cluster method when the max- $K$  strategy was applied (cf. Sec. III B). Another problem in practical applications, for example, within the context of computational neuroscience, is that some clusters may be generated by mechanisms the investigator is not interested in. We shall refer to these clusters as artifactual clusters. Ideally, artifactual clusters should not form part of a clustering result. It should be stressed that artifactual clusters are an entirely different concept from spurious clusters. Spurious clusters arise by chance when the estimated connectivity in an unrelated set of systems is higher than the average connectivity level among other unrelated systems. On the other hand, artifactual clusters are

induced by connectivity patterns which do not reflect the underlying mechanisms the investigator is interested in.

These connectivity patterns may result in structures that obscure clusters generated by the physical mechanism that is the focus of the conducted research. An important example in neuroimaging studies is given by motion artifacts, where the time series corresponding to sets of voxels on the border of gray matter are simultaneously affected by the motion of the subjects head. This introduces commonalities into the time series and their connectivity is thus artificially increased. If the effect is sufficiently strong, then this group of voxels will be assigned to their own cluster, instead of participating in a cluster to which they functionally belong. Additionally, the affected (true) cluster is weaker because it has diminished in size.

The problem of spurious clusters may be resolved by testing each identified cluster against a cluster model, assuming such a model is available. However to mitigate the problem of artifactual clusters, it would be necessary to have a model that would be able to discriminate between different types of underlying coupling resulting in the observed connectivity. Alternatively, prior information related to the coupling mechanisms could result in a modification of the connectivity function which would then ideally be insensitive to artifactual connectivity.

## V. CONCLUSION

In this work a mathematical framework based on spectral graph clustering approaches was described. The method identifies clusters of coupled systems by maximizing the average association objective while allowing for a residual set. The residual set is composed of systems that do not belong to any identified cluster. The formulation of the approach allows for effective means of introducing prior information into the clustering procedure via the connectivity function. By formally relaxing the discrete problem into a continuous domain, it could be recognized that the nonuniqueness of the relaxed solution of the multiway normalized cut objective<sup>9</sup> also holds for average association objective clustering, and classical methods<sup>16,19</sup> for factor rotation could be applied to significantly improve the performance of the clustering method especially for weakly coupled systems. The theoretical investigation yielded a relationship between the average association objective and some spectral properties of the connectivity matrix.

The effectivity of the proposed solution was demonstrated using numerical studies of theoretical connectivity models and a Lorenz lattice of 32 oscillators.<sup>10</sup> The numerical studies where the model clusters were known have enabled the comparison of two different cluster assignment strategies. The simple mapping cluster assignment (based on the participation index) was not able to detect weak clusters consistently and failed to recover well-separated clusters of similar cluster strengths (the latter issue was also noted previously<sup>10</sup>). The VARIMAX mapping has demonstrated a marked improvement over simple mapping in recovering weak model clusters and is able to separate clusters of equal or similar strengths effectively.

An example functional magnetic resonance imaging dataset assessing resting state activity in a human subject was analyzed using the proposed method with VARIMAX mapping and the resulting clusters corresponded to several previously described resting state networks.<sup>28</sup> The example asserted that the method can be applied in situations where the data have very high dimensionality and short time series.

Some open questions pertaining to the use of clustering methods for exploratory data analysis were discussed and potential ways of addressing the problems were suggested. We conclude that average association clustering is a flexible and useful approach to discovering communities in networks.

## ACKNOWLEDGMENTS

This work has been supported by the EC FP7 project BrainSync [Contract Nos. HEALTH-F2-2008-200728 (EC) and MSM/7E08027 (CR)] and in part by the Institutional Research Plan (Contract No. AV0Z10300504). The authors would like to thank Chris Lewis for providing the sample neuroimaging dataset and Maurizio Corbetta for a careful examination of the clustering results. The authors would like to thank Jaroslav Hlinka and Martin Holeňa for interesting discussions. We also gratefully acknowledge donation of computational time on the VIC3 supercomputer from the Laboratorium voor Neuro-en Psychofysiologie of the Katholieke Universiteit Leuven. The authors would also like to thank the anonymous reviewers for their careful review and constructive criticism.

<sup>1</sup>M. Fiedler, Czech. Math. J. **23**, 298 (1973).

<sup>2</sup>M. Fiedler, Czech. Math. J. **25**, 619 (1975).

<sup>3</sup>W. E. Donath and A. J. Hoffman, *IBM J. Res. Dev.* **17**, 420 (1973).

<sup>4</sup>A. Pothén, H. D. Simon, and K. P. Liou, *SIAM J. Matrix Anal. Appl.* **11**, 430 (1990).

<sup>5</sup>L. Hagen and A. B. Kahng, *IEEE Trans. Comput.-Aided Des.* **11**, 1074 (1992).

<sup>6</sup>J. Shi and J. Malik, *IEEE Trans. Pattern Anal. Mach. Intell.* **22**, 888 (2000).

<sup>7</sup>A. Y. Ng, M. I. Jordan, and Y. Weiss, *Advances in Neural Information Processing Systems* (MIT Press, Cambridge, MA, 2001), Vol. 14, pp. 849–856.

<sup>8</sup>M. E. J. Newman, *Phys. Rev. E* **74**, 036104 (2006).

<sup>9</sup>S. Yu and J. Shi, “Multiclass spectral clustering,” *IEEE International Conference on Computer Vision*, 2003, pp. 313–319.

<sup>10</sup>S. Bialonski and K. Lehnertz, *Phys. Rev. E* **74**, 051909 (2006).

<sup>11</sup>C. Allefeld, M. Müller, and J. Kurths, *Int. J. Bifurcation Chaos Appl. Sci. Eng.* **17**, 3493 (2007).

<sup>12</sup>C. Allefeld and S. Bialonski, *Phys. Rev. E* **76**, 066207 (2007).

<sup>13</sup>A. Azran and Z. Ghaharmani, “Spectral methods for automatic multiscale data clustering,” *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2006, Vol. I, pp. 190–197.

<sup>14</sup>L. Angelini, S. Boccaletti, D. Marinazzo, M. Pellicoro, and S. Stramaglia, *Chaos* **17**, 023114 (2007).

<sup>15</sup>I. S. Dhillon, Y. Guan, and B. Kulis, *IEEE Trans. Pattern Anal. Mach. Intell.* **29**, 1944 (2007).

<sup>16</sup>H. F. Kaiser, *Psychometrika* **23**, 187 (1958).

<sup>17</sup>C. H. Q. Ding, X. F. He, H. Y. Zha, M. Gu, and H. D. Simon, “A min-max cut algorithm for graph partitioning and data clustering,” in *IEEE Conference on Data Mining*, edited by N. Cercone, T. Y. Lin, and X. D. Wu (IEEE Computer Society, Los Alamitos, CA, 2001), pp. 107–114.

<sup>18</sup>K. Nevels, *Psychometrika* **51**, 327 (1986).

<sup>19</sup>L. L. Thurstone, *The Vectors of Mind: Multiple-Factor Analysis for the Isolation of Primary Traits* (University of Chicago Press, Chicago, IL, 1935).

<sup>20</sup>C. van Rijsbergen, *Information Retrieval* (Butterworths, London, 1979).

- <sup>21</sup>M. Hoke, K. Lehnertz, C. Pantev, and B. Lütkenhöner, in *Dynamics of Cognitive and Sensory Processing in the Brain*, edited by E. Başar and Th. Bullock, (Springer, Berlin, 1988), pp. 84–105.
- <sup>22</sup>S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 9868 (1990).
- <sup>23</sup>J. Talairach and P. Tournoux, *Co-Planar Stereotaxic Atlas of the Human Brain: 3-Dimensional Proportional System—An Approach to Cerebral Imaging* (Thieme, New York, 1988).
- <sup>24</sup>M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen, and M. E. Raichle, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 9673 (2005).
- <sup>25</sup>V. D. Calhoun, T. Adali, G. D. Pearlson, and J. J. Pekar, *Hum. Brain Mapp* **14**, 140 (2001).
- <sup>26</sup>M. J. McKeown, T. P. Jung, S. Makeig, G. Brown, S. S. Kindermann, T. W. Lee, and T. J. Sejnowski, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 803 (1998).
- <sup>27</sup>J. Y. Choi, J. J. Dongarra, R. Pozo, and D. W. Walker, “Scalapack—A scalable linear algebra library for distributed-memory concurrent computers,” in *Fourth Symposium on the Frontiers of Massively Parallel Computation*, edited by H. J. Siegel (IEEE Computer Society, McLean, VA, 1992), pp. 120–127.
- <sup>28</sup>J. S. Damoiseaux, S. A. Rombouts, F. Barkhof, P. Scheltens, C. J. Stam, S. M. Smith, and C. F. Beckmann, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 13848 (2006).
- <sup>29</sup>J. B. Carroll, *Science* **126**, 1114 (1957).