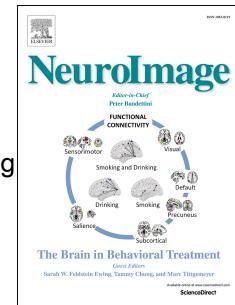


Accepted Manuscript

Resting state dynamics meets anatomical structure: Temporal multiple kernel learning (tMKL) model

Srinivas Govinda Surampudi, Joyneel Misra, Gustavo Deco, Raju Bapi Surampudi, Avinash Sharma, Dipanjan Roy



PII: S1053-8119(18)31859-7

DOI: [10.1016/j.neuroimage.2018.09.054](https://doi.org/10.1016/j.neuroimage.2018.09.054)

Reference: YNIMG 15290

To appear in: *NeuroImage*

Received Date: 3 September 2018

Accepted Date: 19 September 2018

Please cite this article as: Surampudi, S.G., Misra, J., Deco, G., Surampudi, R.B., Sharma, A., Roy, D., Resting state dynamics meets anatomical structure: Temporal multiple kernel learning (tMKL) model, *NeuroImage* (2018), doi: <https://doi.org/10.1016/j.neuroimage.2018.09.054>.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Resting State Dynamics Meets Anatomical Structure: Temporal Multiple Kernel Learning (tMKL) Model

Srinivas Govinda Surampudi^{1*}, Joyneel Misra^{1*},
 Gustavo Deco^{4,5}, Raju Bapi Surampudi^{2†}, Avinash Sharma^{1†}, Dipanjan Roy^{3†}

¹Center for Visual Information Technology, Kohli Center on Intelligent Systems, International Institute of Information Technology Hyderabad, Hyderabad, 500032, India

²School of Computer and Information Sciences, University of Hyderabad, Hyderabad - 500046, India

³Cognitive Brain Dynamics Lab, National Brain Research Centre, Manesar, Gurgaon, Haryana - 122051, India

⁴Center for Brain and Cognition, Dept. of Technology and Information, Universitat Pompeu Fabra, Carrer Tanger, 122-140, 08018, Barcelona, Spain

⁵Institució Catalana de la Recerca i Estudis Avançats, Universitat Barcelona, Passeig Lluís Companys 23, 08010 Barcelona, Spain

Abstract

Over the last decade there has been growing interest in understanding the brain activity, in the absence of any task or stimulus, captured by the resting-state functional magnetic resonance imaging (rsfMRI). The resting state patterns have been observed to be exhibiting complex spatio-temporal dynamics and substantial effort has been made to characterize the dynamic functional connectivity (dFC) configurations. However, the dynamics governing the state transitions that the brain undergoes and their relationship to stationary functional connectivity still remains an open problem.

One class of approaches attempts to characterize the dynamics in terms of finite number of latent brain states, however, such attempts are yet to amalgamate the underlying anatomical structural connectivity (SC) with the dynamics. Another class of methods links individual dynamic FCs with the underlying SC but does not characterize the temporal evolution of FC. Further, the latent states discovered by previous approaches could not be directly linked to the SC, thereby motivating us to discover the underlying lower-dimensional manifold that represents the temporal structure. In the proposed approach, the discovered manifold is further parameterized as a set of local density distributions, or latent transient states. We propose an innovative method that learns parameters specific to the latent states using a graph-theoretic model (temporal Multiple Kernel Learning, tMKL) that inherently links dynamics to the structure and finally predicts the grand average FC of the test subjects by leveraging a state transition Markov model.

The proposed solution does not make strong assumptions about the underlying data and is generally applicable to resting or task data for learning subject-specific state transitions and for successfully characterizing SC-dFC-FC relationship through a unifying framework. Training and testing were done using the rs-fMRI data of 46 healthy participants. tMKL model performs significantly

better than the existing models for predicting resting state functional connectivity based on whole-brain dynamic mean-field model (DMF), single diffusion kernel (SDK) model and multiple kernel learning (MKL) model. Further, the learned model was tested on an independent cohort of 100 young, healthy participants from the Human Connectome Project (HCP) and the results establish the generalizability of the proposed solution. More importantly, the model retains sensitivity toward subject-specific anatomy, a unique contribution towards a holistic approach for SC-FC characterization.

Keywords: rsfMRI, SC, dFC, FC, manifold, tMKL

1. Introduction

One of the major questions persistent in neuroscience community is to comprehend the influence of anatomical structure onto the rich repertoire of temporal fluctuations of neural processes [1]. Since its discovery over two decades ago, there has been a keen interest in investigating the spontaneous intrinsic activity of the human brain. This activity is measured via slow fluctuations in the functional magnetic resonance images (fMRI) when subjects are at rest and not engaged in any task [2]. These fluctuations are highly correlated and discovery of meaningful large-scale functional networks within these correlations led to the use of resting-state fMRI (rsfMRI) to discover human brain function(s) [3, 4]. The resulting matrix of pairwise correlations between regions of interest (ROIs) is termed the functional connectivity (FC) matrix. Complementing fMRI, Diffusion tensor imaging (DTI) captures the white matter streamlines that form the anatomical pathways along which neural activity spreads over the brain [5, 6]. The topography of the brain anatomy, called the structural connectivity (SC), is estimated by counting the number of streamlines connecting a pair of ROIs. Over the last decade, initial studies in understanding the link between anatomical topology and neural activity have provided evidence that the underlying structural topology largely explains the grand-average functional connectivity [7], the missing link being dynamics.

The last decade has witnessed several studies of rsfMRI alone revealing fluctuating spatial patterns which appear and dissolve with time, highlighting the spatiotemporal repertoire of spontaneous brain activity [8]. Temporal fluctuations have been discovered and analyzed, in an unsupervised manner, into finite latent states that appear and disappear with time. Inherent non-linearity of the neural fluctuations necessitates building complex computational models representing dynamics [9, 10]. Models discovering temporal dynamics of rsfMRI can be broadly categorized in the following terms: (i) dynamic functional connectivity (dFC) studies using sliding window approaches providing sequence of windowed FC (wFC) matrices that in turn identify stable transient patterns of functional connectivity fluctuations, called *latent states*, [11, 12, 13, 14], and (ii) Bayesian approaches applied directly on the BOLD timeseries [15, 16, 17, 18, 19], or on the independent components

¹Authors with * have equally contributed to this research work.

^{2†}Corresponding authors

(ICs) extracted from the BOLD timeseries [20, 11] which discover latent states in terms of multivariate Gaussian density distributions of the temporal signals.

How the relatively static SC sculpts the FC over the entire scan duration has been a challenging open research problem in the brain connectome research domain. In the last decade, computational models have been successful at linking SC to static FC. Whole brain computational models aid study and simulation of the temporal dynamics over the structure. Extant whole-brain models advancing our understanding of the SC-FC link can be broadly categorized as follows: (i) models incorporating non-linear dynamics [21, 22], (ii) graph theoretic models [23, 24, 25], (iii) models at the boundary of biophysics and graph-theoretic abstractions [26, 27, 28]. Abdelnour et al. [26] proposed a graph-diffusion framework relating linear diffusion equation of the neural activity over the structural topology to random walks of the activity over the structure. Becker et al. [25] mapped spectral signatures of the structural and functional topologies based on indirect structural walks of the neural activity. Surampudi et al. [27] proposed abstraction of non-linear diffusion equation into combinations of multi-scale diffusion to map subject specific SC-FC.

It has been highlighted that the ‘most interesting’ part of the task-specific dynamics can be discovered on a potentially lower-dimensional manifold. A general perspective is that the neural activity during a task, although being in a high-dimensional space, follows trajectories in a lower-dimensional task-specific manifold during the functional dynamics [29]. As the neural activity is highly coordinated resulting into distinct functional structures, represented as states, it may be appropriate to consider that the large-scale dynamics evinces an underlying lower-dimensional manifold. This sufficiently motivates the presence of a lower-dimensional manifold for rsfMRI as well. To the best of our knowledge, previous studies have not attempted at explicitly discovering an underlying manifold over which the entire dynamics resides. Manifold discovery is a critical step in the proposed model enabling succinct yet elaborate representation of dFC.

Moreover, the question of how a relatively fixed anatomical structure supports the rich spatiotemporal dynamics is still elusive. Abdelnour et al. [28] have extended their graph-diffusion framework for characterizing SC-dFC relationships. However, theoretical models incorporating principled amalgamation of structural topology and dynamics of rsfMRI are essential. Here, we propose an innovative solution for characterizing the SC-dFC-FC relationship. This is achieved by proposing three novel constructs: (i) discovery of an intrinsic lower-dimensional manifold that represents the latent structure of the resting-state temporal dynamics, (ii) temporal multiple kernel learning (*tMKL*) model that learns the SC-dFC mapping, and (iii) generation of latent time series for dFC-FC mapping. The proposed solution estimates grand average FC (gFC or FC) from SC by predicting dFC along with capturing the temporal evolution.

Temporal evolution is characterized by using a first-order Markov model between consecutive state transitions. This model is used for generating a generalized state sequence by unfolding the temporal evolution of Markov model. This state sequence is further replaced by sequence of corresponding state-specific FCs generated by the tMKL model. Finally, these state-specific FCs are factorized to recover a latent timeseries sequence. gFC is then computed on the reconstructed latent timeseries and compared with the empirical FC.

The proposed model recovers the FCs that are very close to empirical FCs as the state-specific FCs recovered with the tMKL model enable realization of

subject-specific functional dynamics. Further, various perturbation experiments demonstrate the robustness and validity of the proposed scheme.

The specific contributions of the work are the following:

1. Novel approach for learning the SC-FC mapping through characterizing the dynamic functional connectivity (dFC) over time windows.
2. Proposal of a novel multiple diffusion kernel model that learns to predict state-specific FCs from SC (tMKL model).
3. Estimating the latent fMRI time series by using the Markov transition probability matrix in conjunction with the tMKL model.

One of the benchmarks of the model is its reproducibility on a novel dataset. This validates that the model does not overfit on any particular dataset and may easily generalize over other unseen datasets. Moreover, variability of SC-dFC-FC across datasets challenges the applicability of the learnt model parameters over such unseen datasets. tMKL model is learnt over half of the typically developing cohort from the dataset acquired at the Berlin Center for Advanced Imaging, Charité University, Berlin, Germany [30] and is shown to be generalizable on a larger dataset - 100 unrelated subjects from the human connectome project (HCP) [31].

The rest of the paper is organized as follows. In the next section we present the details of the proposed solution. In the subsequent sections we present the details of the neuroimaging data set used, qualitative and quantitative evaluation results along with explanation for the choice of model parameters. Finally, we conclude by pointing out limitations and future research directions.

2. Methods

2.1. Proposed model

In this section we describe in detail the whole pipeline of the proposed parametric-model to map the relation between SC and FC using resting state fMRI timeseries data. The proposed model considers the importance of the underlying anatomical constraints to generate the temporal richness as well as to characterize whole-brain FC dynamics. Figure 1 shows a flowchart of the essential elements of the whole pipeline. Proposed model partitions aspects of the whole-brain dynamics essentially into two parts: (i) characterizing temporal dynamics through identification of latent transient states, and (ii) linking them to the underlying structural geometry. These two aspects are parameterized using a novel blend of unique methods. The model utilizes wFCs (steps 1. – 2.) for identifying states and from the resultant SC-wFC pairs, the relationship between the structure and functional dynamics is learned (steps 3. – 5.). Once these two parts successfully characterize the above mentioned aspects by tuning respective parameters, the model is tested for its generalizability using unseen test subjects (steps 6. – 9.).

Prior to identifying latent states within the dynamics, we discover an underlying globally non-linear manifold that spans all the wFCs (step 2.a), thus recovering the lower-dimensional space for expressive characterization of resting state dynamics. The technique used is called spectral embedding. Next, we employ a probabilistic framework, known as Gaussian mixture models (GMM), for estimating the number of states and the shape of each state in the lower-dimensional space, ensuring *soft assignments* of wFCs to its neighboring states

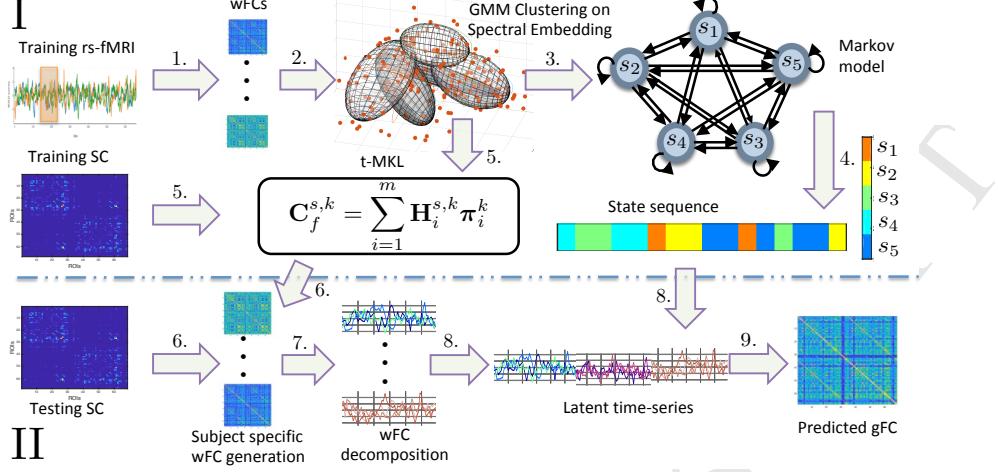


Figure 1: **Outline for temporal Multiple Kernel Learning (tMKL) model.** Figure shows the entire pipeline for predicting grand FC for a testing subject. The model incorporates subject specificity along with temporal variation characterization. Part (I) of the model, training phase, consists of three blocks. The first one, learns temporal variations in terms of distinct states via GMM clustering over the underlying manifold of wFCs (steps 1. and 2.). The second block utilizes the empirical transitions between these distinct states and captures dynamics in the first order Markov chain (steps 3. and 4.). The third block learns subject-specificity by modeling each state by its MKL model [27] (step 5.). Part (II) of the model validates its generalizability on unseen subjects. Importantly, only SC of a testing subject is required (step 6.). Each state for the testing subject is characterized in step 7. Each state-specific predicted FC is decomposed into a latent timeseries which are then concatenated using the steady state distribution of the Markov chain (steps 4. and 8.). Finally, grand average FC, static FC, is predicted for that subject (step 9.).

(step 2.b). These soft assignments are further used to estimate the transition dynamics between these states using a first-order state transition Markov model (step 3. – 4.). With respect to second part of the model, we adapt the multiple kernel learning (MKL) framework [27] for parameterizing the dependence of SC on wFCs for each state (step 5.). We observe that the parameters to be learned form a non-convex combination. Thus we formulate the learning objective into an optimization formulation.

The model predicts state-specific FCs (sFCs) for a test subject (step 6.). These sFCs are decomposed into a latent timeseries (step 7.) which are concatenated using the relative frequency of occurrence of states to generate a global timeseries for calculating the static FC of a subject (step 8.). This relative frequency sequence is also called the steady-state distribution of the Markov model. Thus, for a new subject, given the SC, static FC along with its state-specific FCs are predicted by the proposed model (step 9.).

In the subsequent subsections we elaborate each part of the proposed model. From now on, let $D = \{\mathbf{F}_w^1, \dots, \mathbf{F}_w^s, \dots, \mathbf{F}_w^p\}$ be the set of all wFC matrices obtained by sliding a window of fixed size ω over the n -dimensional fMRI time-series (n being the number of ROIs) belonging to all the p training subjects (step 1.).

2.1.1. Spectral Embedding

We propose to soft-cluster these wFC matrices into K states, first by vectorizing the lower triangular part of a wFC matrix into a column vector (of size $\frac{n(n-1)}{2} \times 1$). These wFCs may be sparsely spaced in a higher-dimensional space, but might originally lie on an intrinsic globally non-linear, lower-dimensional manifold [32] (step 2.a). Spectral embedding method is employed to reduce the dimensionality of the data, by finding a mapping to a lower dimensional manifold over which these wFCs reside [33]. Initially a similarity graph is constructed over the vectorized wFCs. This provides a discrete approximation of the continuous manifold over which these wFCs lie.

The principal eigenmaps (eigenvectors) of the Laplacian operator over this similarity graph form the natural mapping of these wFCs onto the manifold. One of the properties of the Laplacian matrix is that the increasing order of its eigenvalues corresponds to the decreasing order of the principal variance directions of the manifolds; i.e., the lowest K eigenvectors correspond to the first K principal variance directions [34]. The eigenvectors can be considered as principal directions that preserve the maximum variance of the elements on the manifold while also providing a mapping between the original space and the lower-dimensional manifold. Thus the Laplacian eigenmaps preserve the local structure, or the variance structure in the graph, thereby keeping the solution embedding robust to outliers and noise.

The spectral embedding method is applied as follows. Firstly, an adjacency matrix (of size $p \times p$) of the similarity graph is created by applying a radial basis function over the L1 distance between every pair of wFCs. This matrix captures pairwise relationship between wFCs in a relational graph. Next, we form the corresponding normalized graph Laplacian matrix and use the eigenvectors (of size $p \times 1$) corresponding to its lowest K eigenvalues to define the basis vectors of the embedding space [35, 36].

The K eigenvectors of p -dimensions each form a K -dimensional embedding space. When the s^{th} components of each of these K eigenvectors are aggregated into a K -dimensional vector, this vector then represents the transformation of \mathbf{F}_w^s into the lower-dimensional embedding space. Finally these K -dimensional wFCs are clustered using Gaussian Mixture Model (GMM), as explained in the next section.

2.1.2. GMM Clustering

Following the discovery of an approximation to the continuous lower-dimensional manifold, we now parameterize the local density distribution of wFCs over the manifold using a probabilistic framework, Gaussian mixture model (GMM) [37] (step 2.b). Gaussian mixture model is a factor analysis model that represents the probability density of a sample as a weighted combination of component Gaussians. Such a representation facilitates GMM to represent a large class of sample distributions. Specifically, distribution of wFCs over the manifold are modeled as a GMM.

Let the density of \mathbf{F}_w^s be a linear combination of K component Gaussian

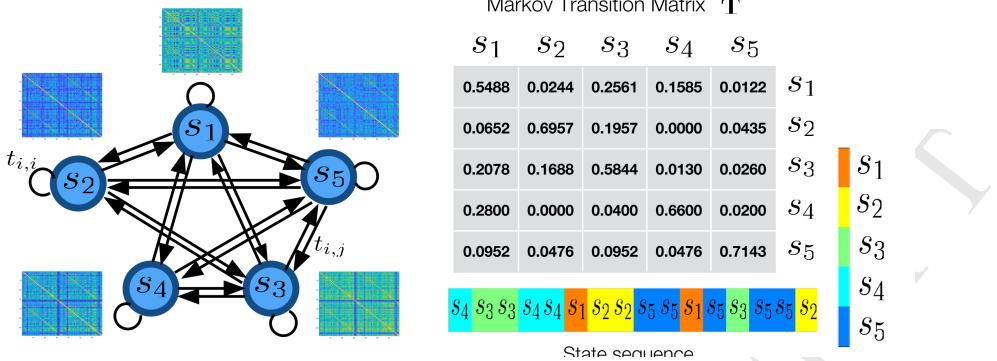


Figure 2: **Graphical depiction of proposed Markov state transition model.** An illustration of the first-order Markov chain used as a part of the proposed tMKL model. Each state has its unique distribution of FCs, represented as a Gaussian in the embedding space, from which subject-specific FCs can be sampled. The corresponding transition matrix (for $K=5$) and an example state sequence generated with a Markov random walk over the transition matrix is also depicted. Note that this Markov transition matrix is only for visualization purposes, and the identified number of states is $K = 12$.

densities, represented as follows:

$$\begin{aligned} \mathbf{P}(\mathbf{F}_w^s; \Theta) &= \sum_{k=1}^K \Psi^k(s) \mathcal{N}(\mathbf{F}_w^s; \mu^k, \Sigma^k) \\ \sum_{k=1}^K \Psi^k(s) &= 1, \forall s = 1, \dots, p \end{aligned} \quad (1)$$

where $\mathbf{P}(\cdot)$ denotes the probability density of a wFC. $\Psi^k(s)$ is the probability of \mathbf{F}_w^s belonging to the k^{th} component Gaussian. Each component Gaussian is a K -variate Gaussian probability density function of the form:

$$\mathcal{N}(\mathbf{F}_w^s; \mu^k, \Sigma^k) = \frac{1}{(2\pi)^{K/2} \det(\Sigma^k)^{1/2}} \exp \left\{ (\mathbf{F}_w^s - \mu^k)^\top \Sigma^{k-1} (\mathbf{F}_w^s - \mu^k) \right\},$$

where μ^k and Σ^k represent the mean vector and covariance matrix of the k^{th} Gaussian component.

GMM thus represents the manifold as a set of Gaussian densities and parameterizes it in terms of Θ :

$$\Theta = \{\Psi^k(\cdot), \mu^k, \Sigma^k\}, k = 1, \dots, K. \quad (2)$$

As the collection of these component Gaussians forms the manifold, the component Gaussians can be interpreted as a *latent transient state* visited by the brain. Each latent state is a component Gaussian whose location and shape is governed by μ^k and Σ^k , respectively in the manifold.

2.1.3. State Transition Markov Model

As described in the previous section, the wFCs are quantized into finite states $S = \{s_1, \dots, s_K\}$ by clustering all the wFCs using GMM. Each wFC sequence now corresponds to a cluster-label (state) sequence and transitions

between these states is representative of the dynamics in the BOLD rsfMRI time series (step 3.). We assume first-order dependence among these transitions and learn the Markov transition probability matrix, $\mathbf{T}_{K \times K}$ by estimating the state transitions within the training data.

Figure 2 shows an example of a Markov model for $K = 5$ and the corresponding transition probability matrix. Each edge $t_{i,j}$ captures the probability of transition from state i to state j . Similarly, self-loop edges $t_{i,i}$ depict the probability of remaining in the same state. For each state i we compute $t_{i,j}$ by counting the number of first-order transitions to state j in the state sequence. Finally, we normalize each row sum of \mathbf{T} to value 1 yielding a valid transition probability matrix. In the testing phase, the Markov matrix learned on training wFCs is used to generate a random state sequence, to eventually construct the latent timeseries for testing subjects.

As any Markov chain converges to its steady state distribution with time regardless of its initial state distribution, we find the steady state distribution over the transition matrix and use this distribution as frequency of occurrence of states over the time course. This gives us a state transition sequence for a test subject. Along with the state transition model which captures the dynamics of the latent states, a model that relates anatomical structure to these states is required. In the next section, we propose a temporal multiple kernel learning (tMKL) model that learns this mapping.

2.1.4. tMKL Model

The process by which randomly-moving objects wander away from where they started is usually known as a random walk. The set of inter-regional anatomical connections creates a network or a graph whose topology has been well studied [38]. Mean regional activities of all regions are assumed to be in a random walk over the SC graph. In case of SC-FC mapping, this phenomenon is represented by a linear differential equation whose analytical solution is the diffusion kernel over the graph defined by SC which is hypothesized to be representing FC [26]. [24] discovered that physical diffusion over such large scale graphs exhibits multi-scale relationships with FC, thus a linear combination of multiple diffusion kernels is considered a better representative of FC - this model is referred to as multiple diffusion kernel (MDK) model from now on. The linear combination coefficients are scalar values which equally weigh all regional activities at each diffusion-scale. But it may so happen that activities of non-physically connected regions may be modulated by other regions. This hypothesis was empirically validated by Surampudi et al. [27].

Adapting this line of models for every state, we represent the diffusion phenomenon and introduce the variables π_i 's of size $n \times n$, that capture the inter-regional co-activation patterns at diffusion-scale i , $\forall i = 1, \dots, m$, m being the number of diffusion-scales [27].

The model is shown in step 5. of Figure 1. Consider a single SC and a state, let a diffusion kernel defined at diffusion scale i be denoted by \mathbf{H}_i .

$$\mathbf{H}_i = e^{-\tau_i \mathbf{L}} \quad (3)$$

Here, τ_i is the spatio-temporal scale of heat diffusion and \mathbf{L} is the Laplacian matrix corresponding to the SC. We propose that a wFC matrix can be decomposed into a set of diffusion kernels multiplied with their co-activation patterns:

$$\mathbf{C}_f = \sum_{i=1}^m \mathbf{H}_i \boldsymbol{\pi}_i, \quad (4)$$

Here, \mathbf{C}_f denotes predicted wFC. We hypothesize that co-activation patterns are distinctly different for each state and hence we add a superscript index k ($k = 1 \cdots K$) to obtain $\boldsymbol{\pi}_i^k$. As the parameters $\boldsymbol{\pi}_i^k$'s are state dependent, state-specific predicted functional connectivity, $\mathbf{C}_f^{s,k}$, will be as follows:

$$\mathbf{C}_f^{s,k} = \sum_{i=1}^m \mathbf{H}_i^{s,k} \boldsymbol{\pi}_i^k = \sum_{i=1}^m e^{-\tau_i^k \mathbf{L}^s} \boldsymbol{\pi}_i^k \quad (5)$$

Here \mathbf{L}^s is the Laplacian matrix of the SC corresponding to the wFC, \mathbf{F}_w^s . This results in the following optimization problem for $\boldsymbol{\Pi}^k$ and $\boldsymbol{\tau}^k$:

$$\begin{aligned} & \text{minimize}_{\boldsymbol{\Pi}^k, \boldsymbol{\tau}^k} \quad \sum_{s=1}^p \left\| \Psi^k(s) \left(\mathbf{F}_w^s - \mathbf{C}_f^{s,k} \right) \right\|_F^2 + \lambda_1 \sum_{i=1}^m \|\boldsymbol{\pi}_i^k\|_1 + \lambda_2 \sum_{i=1}^m \|\boldsymbol{\pi}_i^k\|_2 \\ & \text{subject to} \quad \mathbf{C}_f^{s,k} = \sum_{i=1}^m e^{-\tau_i^k \mathbf{L}^s} \boldsymbol{\pi}_i^k \\ & \quad \boldsymbol{\pi}_i^k \in \mathcal{S}_+^n, i = 1, \dots, m, k = 1, \dots, K \\ & \quad \boldsymbol{\tau}^k \succeq \mathbf{0}. \end{aligned} \quad (6)$$

Here, \mathcal{S}_+^n is the convex set of positive semi-definite matrices. The objective function takes the form well known in regression analysis as *least absolute shrinkage and selection operator* (LASSO) that performs both variable selection and regularization. We arrived at the model parameters experimentally, for example, the number of scales m is empirically chosen (see Section 3.2).

Finally, the model consists of m distinct $\boldsymbol{\pi}_i^k$'s which are learned for each of the K states.

2.1.5. Generation of latent timeseries for testing subjects

As described in the previous section, for a subject, we predict the state-specific FC matrix for each of the states using the input SC matrix of the testing subject and the learned tMKL model (step 6.). Based on the learned Markov chain state transition matrix, a sequence of states is generated by randomly sampling from the steady state distribution of the transition matrix (step 4.). Each of the state-specific FCs in the resulting sequence is factorized into state-specific latent timeseries and concatenated, using the generated state sequence, to obtain the latent timeseries for the testing subject.

In the training phase, wFCs are obtained by computing Pearson correlation coefficients of the windowed rsfMRI timeseries over various regions. We know that Pearson correlation between two timeseries A, B is $\rho(A, B) = \frac{\text{cov}(A, B)}{\sigma_A \sigma_B}$. Hence the state-specific wFC matrix works out to be the covariance of its state-specific latent timeseries $\hat{Z}_{n \times \omega}$. The eigenvalues of this covariance matrix represent the variance of the data along the principle component basis and are therefore non-negative, thus making the predicted sFC matrix a positive semi-definite matrix.

Thus we can factorize a state-specific wFC as follows:

$$\begin{aligned}\mathbf{C}_f^{s,k} &= \mathbf{U}_{n \times n} \boldsymbol{\Lambda}_{n \times n} \mathbf{U}_{n \times n}^\top \\ &= (\sqrt{\boldsymbol{\Lambda}} \mathbf{U}^\top)^\top (\sqrt{\boldsymbol{\Lambda}} \mathbf{U}^\top) \\ \hat{\mathbf{Z}} &= \sqrt{\boldsymbol{\Lambda}} \mathbf{U}^\top.\end{aligned}\tag{7}$$

Thus, using Eq. 7, we recover latent timeseries matrix $\hat{\mathbf{Z}}$ that can be taken as approximated timeseries used for obtaining wFC (step 7.). For a testing subject, each cluster-specific wFC is decomposed into latent timeseries and these are concatenated into a grand timeseries (step 8.). The latent time series are concatenated by considering the steady state distribution of the Markov chain. Steady state distribution is the probability of being in a state which remains the same throughout transitions. Every random walk over the transition matrix approximates this distribution after infinitely long time. Finally, as Pearson correlation is order-agnostic, calculating Pearson correlation matrix of the grand timeseries generates the predicted grand-average FC (gFC) for the testing subject (step 9.).

3. Experiments & Results

Performance of the proposed model was evaluated in the following setup. We used Pearson correlation coefficient between empirical and predicted functional connectivities (FC) as the measure of model performance in order to keep the measure of model performance consistent with the extant literature.

We will show all our major results on the dataset acquired at the Charité University (we will be calling it as Charité University Berlin, CUB, dataset) and show the reproducibility of the model on the human connectome project dataset (called as HCP from now on) [39, 40, 41, 42]. The following simulations are performed on the Berlin dataset: A randomly chosen set of half of the cohort (23 participants) was used for training and the other half (23 participants) for testing. We first compare the performance of the proposed model against several extant methods that provide SC-FC mapping followed by explaining the rationale behind the choice of optimal model parameters. We also conduct k-fold cross validation results and perturbation experiments, the results of which support generalizability of our model to other data splits.

The proposed model predicts state-specific FCs which are thereby used to produce the gFC. The quality of the gFC prediction is highly dependent upon the reproducibility of states and their transition patterns across multiple train-test splits. Obtaining different set of states in different splits shall attest the robustness of the proposed model at question. We analyze the states discovered from our model by observing the state-specificity property of the model and compare it with the states learned using k-means algorithm in Allen et al. [11]. Finally, to validate the generalizability and reproducibility of the model, we test the learnt model on the hcp dataset.

3.1. Grand average FC (gFC) prediction

We compare the performance of the proposed model with several existing approaches: single diffusion kernel (SDK) model [26], the non-linear dynamic mean field (DMF) model [22] and multiple kernel learning (MKL) model [27].

To our knowledge, ours is the only model that incorporates structural information along with temporal dynamics for predicting grand average FC. DMF and SDK models do not incorporate learning in their formulation and tune the parameters for each subject separately. DMF model inherently captures non-stationarity, therefore it is directly used for gFC prediction without computing wFCs. We estimated the optimal parameters of the DMF and SDK models from the training wFCs and predicted the gFCs of testing subjects using these optimal parameters. The mode of the performance distribution histogram for the training set was used to select the optimal model parameters.

Figure 3 shows that tMKL has superior performance compared to the others. Figure 4 shows the performance of tMKL model on an unseen dataset, validating models' reproducibility. As MKL model is shown to perform better than the other two models, we compared MKL and tMKL only on the HCP dataset. The plot suggests that even on a new dataset tMKL model performs at par or better than the MKL model.

To further validate the generalizability of the tMKL model over unseen testing data, we performed k-fold cross-validation experiment whose results are listed in Table 1. These results suggest that performance of our solution is consistent across various splits, hence supporting our claim of generalizability of our model on unseen data.

k	fold-1	fold-2	fold-3	fold-4	fold-5	mean
2	0.757	0.732	-	-	-	0.745
3	0.771	0.811	0.778	-	-	0.787
5	0.785	0.809	0.813	0.809	0.808	0.805

Table 1: **Cross-validation experiments suggesting generalizability of tMKL model.** Mean k -fold cross-validation results for $k = 2, 3, 5$ are shown in the corresponding rows for k -values. As the number of training samples increases with the number of folds, the mean performance also increases suggesting that the model is learning well with increased samples and is able to replicate the same for testing subjects.

Now, the choice of various model parameters is explained in the next subsection.

3.2. Parameter Selection

1. **Choice of size of sliding-window, ω :** Within the extant literature, the choice of a suitable sliding window size is an open problem with respect to the analysis of temporal dynamics in rs-fMRI [10]. The sliding window size should be small enough so as not to miss the state transitions and should be large enough to capture the state transitions reliably. Keeping this in mind, we followed Allen et al. [11] by using a sliding window of $\omega = 22$ TRs. The window was tapered at the ends by convolving it with a Gaussian of $\sigma = 3$ TRs and was slid with a stride of 5 TRs to create wFCs.
2. **Choice of GMM parameters:** Each *latent transient state* in which the wFCs lie is represented using a component Gaussian of the GMM. In order to choose the optimal number of these states, K , we selected the GMM model corresponding to a minimum BIC score. Bayesian information criterion (BIC) is a statistical measure based on the log-likelihood function



Figure 3: **Model performance comparison between tMKL and existing models.** Pearson correlation between the empirical and predicted gFCs for all the testing subjects is shown for all models. As can be seen, MKL model outperforms other two models, and tMKL model is at par or better than MKL for all but one testing subjects. Even though there is marginal gain in the overall prediction quality, tMKL provides rich insights into the temporal dynamics thus gaining its superiority over extant models.

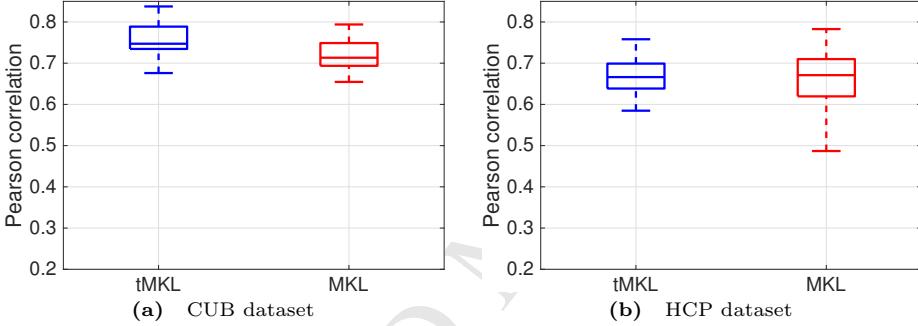


Figure 4: **Model performance between tMKL and MKL over a novel dataset.** In order to validate that the model is reproducible on a new dataset, we tested tMKL and MKL models, learned over the 23 training subjects of the CUB dataset, on the HCP dataset. Pearson correlation between the empirical and predicted gFCs for (a) the 23 testing subjects of the CUB dataset as well as for (b) the 100 unrelated subjects of the HCP dataset. It can be seen in (a) that the median performance of tMKL is better than MKL for the CUB dataset. While the medians for the HCP dataset are comparable (0.66 for tMKL and 0.67 for MKL). However tMKL exhibits lower variance in performance across the HCP subjects in comparison to MKL. Therefore this plot suggests that tMKL performs at par or better than MKL for unseen data.

used for selecting a model amongst a finite set of alternatives, where the model corresponding to the lowest BIC score is chosen. The plot in Figure 5 shows BIC scores for the models obtained by fitting GMM for a large range of K (2 to 19), where the minimum value was obtained for $K = 12$. For each K , we ran GMM 100 times and noted the minimum BIC score, these BIC scores were used in the figure. To retain generality of the component Gaussians, we ran our experiments by considering a unique full covariance matrix for each component Gaussian.

3. Choice of number of diffusion scales for tMKL, m : The scale values were sorted in ascending order, where lower values correspond to local diffusion phenomenon and higher values correspond to global diffusion phenomenon. Scale values lying in-between correspond to intermediate diffusion phenomena. We ran several experiments by varying m at powers

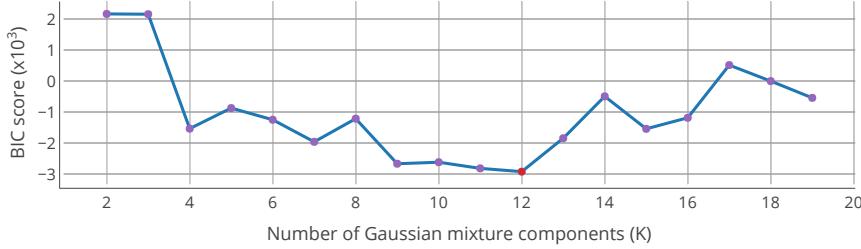


Figure 5: Bayesian information criterion (BIC) score for selecting the number of components in Gaussian mixture model. The GMM is fit over the training wFCs lying in the lower dimensional manifold. The BIC score is reported by varying the number of component Gaussians (K) from 2 till 19. The Gaussian mixture model corresponding to $K = 12$ (shown in red) has the lowest BIC score among others and is therefore preferred. The plot shows local minima at $K = 4, 7$, and 15 which may mislead the user while selecting the optimum model. This local minima suggests the choice of number of components in Allen et al. [11].

of 2 from 2 to 32. While the performance for all the scales was reasonable, however in order to carry out comparative analysis with the MKL model, we chose the number of scales as $m = 16$ for all the experiments.

3.3. Robustness of the model

In order to validate the robustness of our model we performed various experiments to assess whether our solution overfits the training data and also whether the prediction of the grand average FC is agnostic to the particular SC matrix.

1. **Reproducibility of states :** As mentioned in Section 2.1.2, GMM yields K soft assignment vectors for the training wFCs. We validated reproducibility of this clustering by ensuring replication of the same for wFCs of the testing subjects. We generated wFCs for all the testing subjects using the sliding window approach. Soft assignment vectors were generated for these testing wFCs using the GMM employed on the training data, which is then used to compute the Markov transition matrix and the corresponding steady state distribution. Figure 6 shows an example of the steady state distribution for our proposed method. We evaluated the similarity between the Markov transition matrix and steady state distributions of the training and testing wFCs by finding the Pearson correlation coefficients. Table 2 shows that the states are highly replicable for multiple train-test splits of the data.
2. **Perturbation experiments :** Each testing subject SC was perturbed $N = 150$ times and using the learned model we predict the grand average FC. We perturbed every SC by randomly generating it from the power law distribution followed by its elements. The generated state-specific wFCs may have non-positive eigenvalues. Here we considered only the real part of the generated time series in order to estimate (predict) grand average FC. Figure 7 shows this observation over all the 23 testing subjects. Box plots for each subject depict the range of correlation values for random SCs. Here we observe less correlations between empirical and predicted gFCs using the perturbed SC, validating that our model respects

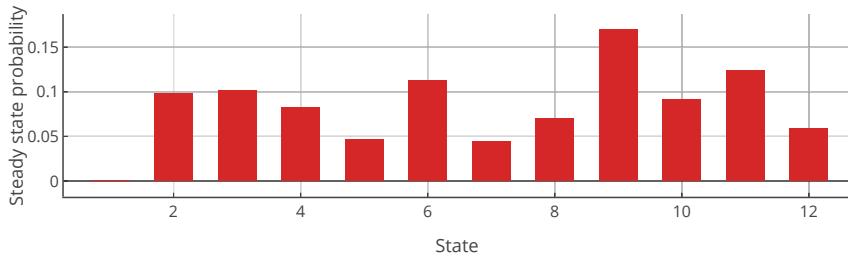


Figure 6: **Steady state distribution of the Markov chain.** After the states are retrieved using GMM, the Markov chain transition matrix was learned over the resulting state-sequences of the wFCs of the training subjects. The figure shows the steady state distribution of the transition matrix, which represents the probability distribution of occurrence of a state after infinite amount of time.

Run Index	ρ_{TM}	ρ_{SSD}	e_{TM}	e_{SSD}
1	0.9947	0.9509	0.1337	0.0564
2	0.8683	0.8546	0.7379	0.2703
3	0.9440	0.8839	0.5120	0.1433
4	0.9035	0.9809	0.7154	0.1004
5	0.8624	0.9604	0.7094	0.1332
6	0.9665	0.8337	0.3824	0.1119
7	0.9131	0.8563	0.6263	0.1107
8	0.9746	0.6824	0.3381	0.1521
9	0.9275	0.8691	0.6671	0.0950
10	0.8623	0.9599	0.7299	0.1608
11	0.9777	0.9596	0.3250	0.0482
mean	0.9301	0.8692	0.5068	0.1358
stdev	0.0501	0.1155	0.2093	0.0636

Table 2: **Comparison of Markov chain transition matrix (TM) and its steady state distribution (SSD) between training and testing subjects.** Comparison is done computing the Pearson correlation coefficient (ρ) and the L2 distance (e) between the training-TM, testing-TM and training-SSD, testing-SSD respectively. This experiment is repeated for 11 train-test splits of the data. Consistent high values of ρ and low values of e across multiple splits show similarity of the states and their transition behavior across train-test subjects, therefore establishing the reproducibility of states.

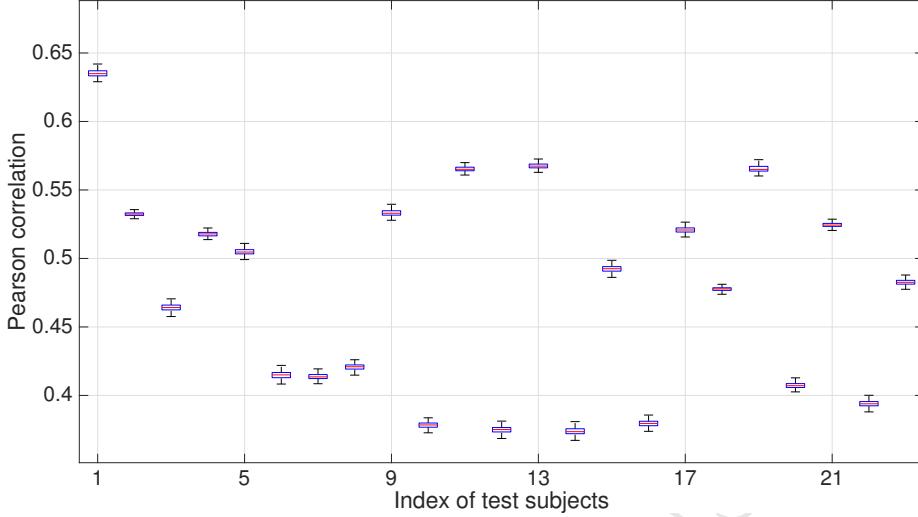


Figure 7: **Effect on the performance of tMKL model due to perturbation of SC matrices of test participants.** Shown here are the box plots (blue) of Pearson correlation between empirical and predicted grand FCs when SCs are perturbed for all the test participants.

the topology of input SC. This suggests that the model is not overfitting the data and is sensitive to perturbation in SC.

3.4. State-specificity of the tMKL model

In the previous section, we successfully investigated whether the estimated state transition matrix is general enough in the sense of being reproducible with several train-test splits of the data (refer Table. 2). Other critical questions are whether different model components such as the Π^k 's as well as the predicted FCs are distinct for different states or not. If they are not distinct, the resulting MKL models for different states become redundant. In order to verify the state-specificity of the model, we performed three simulation experiments: i) to show that the learned state-specific model parameters on the training data are distinct for different states, ii) to show that the predicted state-specific FCs during testing phase are also distinct from one another and, iii) to evaluate the accuracy of state-specific assignments of the model prediction using precision and recall measures.

As summarized in Section 2.1.4, the full model consists of estimating $m = 16$ π_i^k 's for all the $K = 12$ states (i ranging from 1 to 16 and k ranging from 1 to 12). We perform a comparison experiment to see whether, for a fixed i , π_i^k 's are dissimilar from each other. The results of the first experiment are depicted as $m = 16$ similarity matrices in Figure 8. It appears that the learned π_i^k 's are indeed different for different states, especially in the similarity matrices for global scales (see the top row of Figure 8).

In the second experiment, we verified whether the predicted state-specific FCs during the testing phase are distinct from one another. For a test subject, there are K state-specific FCs predicted based on the SC of the subject (see step 6. of Figure 1). In the previous experiment, as the Π^k 's have been demonstrated to be distinct from each other, given a fixed test SC, the state-specific

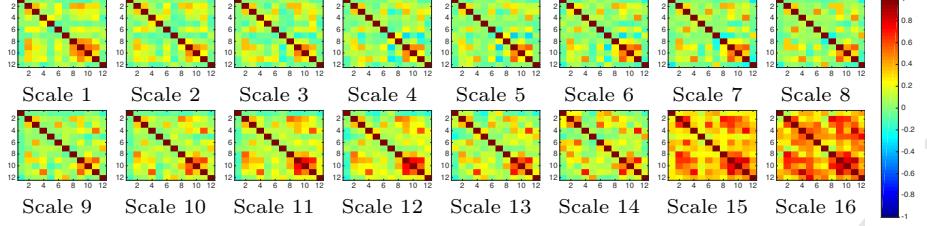


Figure 8: **Distinctness of π_i^k 's.** After state-specific MKL models are learned, we check the distinctness of π_i^k 's for every scale value ranging from $i = 1, \dots, m (= 16)$ using Pearson correlation coefficient between every pair of states. Each of these m matrices is a $K (= 12) \times K$ similarity matrix. Distinctness of π_i^k 's would ideally result in a $K \times K$ identity matrix and any deviation would indicate lack of distinctness of the learned π_i^k 's. As observed in most of these m similarity matrices, majority of the off-diagonal entries in these pairwise correlation matrices are zero, indicating the distinctness π_i^k 's. They are significantly distinct for global scales (scale indices $i = 1, \dots, 8$ in the top row) in comparison to local scales (scale indices $i = 9, \dots, 16$ in the bottom row), where they appear to be similar.

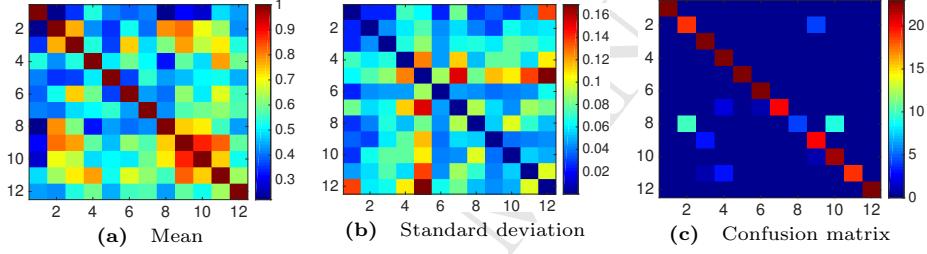


Figure 9: **Quantitative state specificity of the model.** Pearson correlation between all the possible pairs of the K state-specific FCs for a subject was calculated and stored in a $K \times K$ matrix. Element-wise mean and standard deviation across all the subject-specific matrices is shown in (a) and (b) respectively. The confusion matrix in (c) can be used for measuring the overall accuracy of state-specific predictions, and precision and recall for each state. For each testing subject there are K number of state-specific FCs. The state label against which a state-specific FC is predicted from tMKL model serves as the ground truth for this experiment. Empirically, each state-specific FC must be nearer to the training wFCs belonging to that state than from the training wFCs belonging to other states, thus attesting the accuracy of model prediction. As the manifold was constructed based on L_1 similarity, we found the neighbors of the predicted FCs in the original space (of size $n(n-1)/2$). For this purpose, we searched for 25 nearest neighbors for each state-specific prediction and voted for the empirical state-belongingness. Rows (columns) in the confusion matrix depict the actual (predicted) states. Overall accuracy of tMKL model prediction for all the test subjects is 87.68%. It can be seen that non-zero off-diagonal entries result in reduced accuracy. To get a subject-specific measure of the state-specificity, we ran the same experiment for all the testing subjects independently. Noticeably, mean matrix is similar to the confusion matrix with very less standard deviation.

predictions are also expected to be distinct. Consequently, we computed pairwise correlations between the K predicted FCs leading to a $K \times K$ similarity matrix for each test subject. We then calculated the element-wise mean (Figure 9(a)) and standard deviation (Figure 9(b)) across the 23 similarity matrices. As shown in the figure, the dominant identity matrix pattern observed in the mean matrix combined with low values in the standard deviation matrix, verifies that the predictions are indeed distinct from one another.

In the third experiment, we evaluated the accuracy of the predicted state-specific assignments of the proposed model. A wFC in the training phase is assigned to a state which it belongs to with the maximum probability of belongingness as described in 1 in section 2.1.2. The cluster assignments of wFCs to states should obey the principle of maximum intra-state similarity as well as maximum inter-state dissimilarity. Therefore the predicted FC for a state in the testing phase should have maximum similarity with the training wFCs belonging to the same state and also minimum similarity with training wFCs belonging to other states. For each predicted state-specific FC, a set of training wFCs is computed which lie in its proximity in the original space (of size $n(n-1)/2$). The mode of the wFC state-labels of this set of neighbouring training wFCs would indicate the estimated state label for the predicted FC. Recall from step 6. of Figure 1 that the tMKL model implicitly assigns a state-label to the predicted FCs. In this experiment, our aim is to compare the implicit label with the estimated label. The concurrence is measured through a confusion matrix aggregated over all the test subjects (see Figure 9(c)). The accuracy of the predicted state-specific assignment measures the number of instances for which this estimated state label matches the implicit state label. The confusion matrix (Figure 9(c)) has a dominant main diagonal and low off-diagonal elements, indicating that the implicit assignments seem to be valid. Overall accuracy of the state-specific assignments for the test subjects works out to be 87.68%.

In summary, the above-mentioned experiments establish the ‘distinctness’ of states and their corresponding predictions.

4. Discussion

How the relatively static SC sculpts the FC over the entire scan duration has been a challenging open research problem in the brain connectome research domain. Besides understanding the relationship between the anatomical architecture and the functional dependencies, over the last decade, characterization of the temporal richness of the resting state functional MRI signal also has been a major trend in the field of Cognitive Neuroscience. Several approaches have been proposed to understand the inherent richness observed in the spontaneous spatio-temporal BOLD activity. Operator-based formulations of neural dynamics [21, 22] propose a generative model to predict functional connectivity from the structural connectivity by incorporating temporal dynamics into the model. Another class of techniques pioneered the use of spectral graph theoretic methods [26, 24, 28, 27] primarily focusing on mapping the eigen-spectrum of SC and FC of individual subjects, but with minimal focus on the temporal richness.

Here, we have proposed an innovative method which combines both anatomical constraints as well as incorporating temporal richness present in the endogenous activity. More specifically, our proposed model learns parameters specific to these latent states using temporal Multiple Kernel Learning (tMKL) and finally predicts the grand average functional connectivity (gFC) of the test subjects by employing a state transition Markov model. The model links the underlying biophysics to spectral graph theoretic tools abstracting out the problem into graph-theoretic/mathematical concepts. One of the interesting proposals in the framework is that tMKL learns a mapping between the underlying anatomical network and the temporal structure present in the empirical data to quantify gFC. It may seem that such a amalgamation of SC, dFC and FC

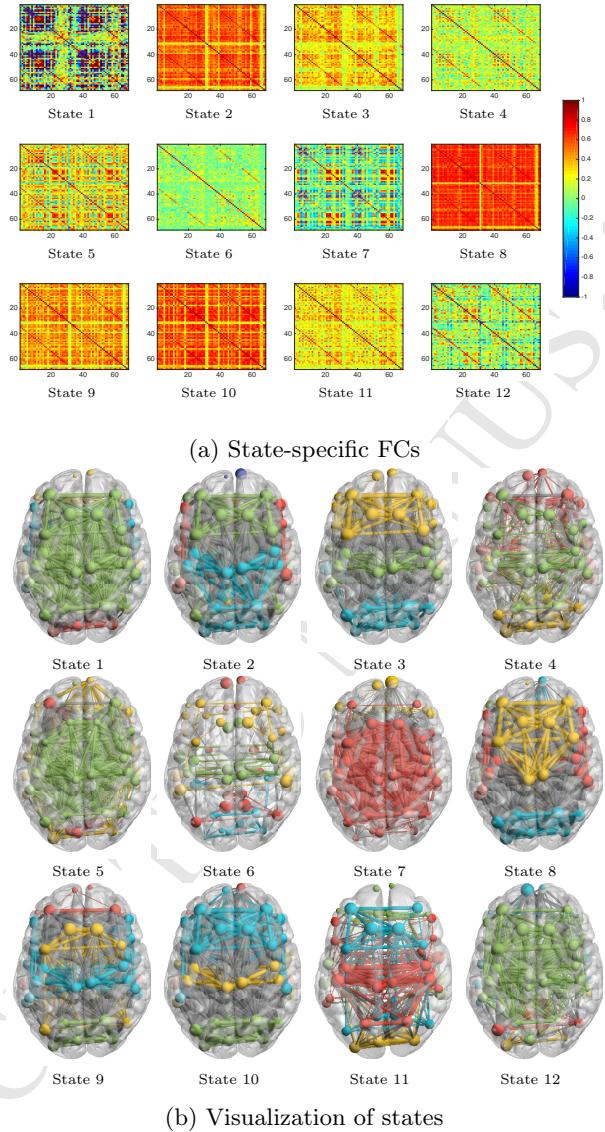


Figure 10: **Qualitative state specificity** State specific FCs predicted for every subject are averaged across all testing subjects. (a) Visually distinct FC matrices are shown for all the 12 states. (b) Communities are identified for these mean FCs using Louvian algorithm available in brain-connectivity-toolbox [43] and Brain-net-viewer [44] was used for visualization of these communities. The distinct community structures clearly suggest that transient states are modeling different brain dynamics.

requires a complex paradigm, but it is a unique combination of simple methods to link anatomical and temporal structures. Interestingly proposed model succinctly juxtaposes SC and rsfMRI timeseries, bridges the gap between them and further links them to gFC.

Further, we have introduced a learning framework to find model-specific parameters via state-specific optimization formulations and yet the model performs at par or better than state-of-the-art models for predicting the gFC. Our proposed model optimizes state-specific windowed FCs (wFCs) individually without incorporating gFC in the optimization formulation. But we still find that the overall performance of the model is superior to that of the state-of-the-art models. Moreover, our proposed model shows sensitivity toward individual subject's SC as demonstrated with the perturbation experiments. However, before we can clearly appreciate the novelty of the proposed techniques we need to understand the existing methods of relating underlying SC with windowed FC (wFC).

4.1. Relating underlying structural connectivity to Windowed FC

A different line of study by Allen et al. [11] and subsequent works by them and other authors have focused on the temporal structure of the windowed FCs (wFCs) and were able to successfully characterize state transitions [45, 46, 47, 10]. Allen et al. collected all the wFCs from the training subjects and learned a k-means clustering model to discover distinct latent states (common to the cohort) that are visited by the brain. Recently, Taghia et al. [19] present a novel Bayesian unsupervised learning model - Bayesian switching dynamical systems (BSDS) - to discover the spatio-temporal evolution of latent brain states that obviates the need of tuning parameters essential to sliding window approaches such as window size, stride, optimal number of clusters etc. BSDS is an extension of Bayesian switching factor analysis approach [17], which employ HMM on latent space representations of BOLD timeseries to identify brain states and their temporal evolution. Such models have been successful at identifying meaningful states in resting state and task conditions and relating them to cognition, however these models do not specifically relate the temporal dynamics to the underlying structure. In this perspective, the proposed model indeed can be considered a stepping stone towards linking SC specifically to timeseries and finally to cognition.

Our preliminary attempts at fusing SC using MKL model [27] with the temporal structure learned with k-means did not give satisfactory results (summarized in Figure S1 in Supplementary materials and the description therein). As summarized in Figure 1, the proposed pipeline partitions the BOLD time series of a subject into windows yielding wFCs. The underlying structure of the wFCs was learned via a manifold whose structure was further parameterized using gaussian mixture model (GMM). The GMM components were hypothesized to be the states whose temporal evolution was succinctly captured in a Markov chain transition matrix. The proposed temporal multiple kernel learning (tMKL) model in this paper belongs to the class of spectral graph theoretic methods [26, 24, 28, 27]. The model attempts both to characterize the temporal richness of the signal and improving upon the quality of SC-FC mapping while incorporating the structural information in a principled way.

The proposed temporal multiple kernel learning (tMKL) model is an attempt toward generating BOLD time series of a subject using only the SC. For each state, MKL model was learned to capture the SC-dFC relationship. The learned

model was utilized to predict state-specific FCs for a test subject. These predicted FCs were further factored into latent time series and concatenated using the steady-state properties of the transition matrix. Pearson correlation of this final time series generates the predicted FC for a subject.

4.2. Rationale for discovering latent temporal structure

In the following section, we will explain the rationale for various steps in the proposed pipeline. It appears that while modeling dFC using unsupervised techniques for clustering wFCs into states, one faces the curse of dimensionality problem head-on. During clustering, wFCs ought to be assigned to the same state as that of their neighbors because they are temporally contiguous and might share similarities. As we can see, wFCs lie in a high dimensional space, but based on their similarity with respect to their neighbors, they may lie on an intrinsic lower-dimensional manifold. This lower dimensional manifold becomes the space over which temporal structure could be efficiently identified. Spectral embedding techniques utilize the similarity between the neighboring wFCs to discover the underlying manifold. After representing the temporal structure as a manifold, the next task is to parameterize the lower-dimensional structure. Once we obtain a lower-dimensional embedding, we need to cluster the wFCs to discover the discrete state space. Unsupervised approaches such as k-means clustering would yield spherical clusters, limiting the shape and size of states, whereas GMM clustering is a generalized clustering scheme. We parameterize the local density-distribution of wFCs over the manifold to a factor analysis model that further represents the manifold as a set of component Gaussians at various locations whose shape, orientation, and size depend on the local densities of the wFCs.

We call these component Gaussians as latent states. After discovery of the states we learn MKL model for each state and call the set of state-specific MKL models as the tMKL model. For each state, the MKL parameters are independently learned by minimizing the reconstruction loss between the predicted and empirical wFC matrices native to the state, without resorting to global optimization of gFC. Finally, we observe that by combining all the state-specific MKL models with the steady state distribution of the Markov chain, gFC could be predicted well. The proposed model is cohort-based and hence the underlying assumption is that the model is generalizable to unseen test subjects. We have learned the Markov transition probability matrix on training wFCs and used this to generate long sequences of time series for test subjects eventually yielding a good approximation of the grand average FC with a maximum correlation of 0.8 (see Figure 3).

4.3. Reproducibility

After presenting the rationale behind designing the proposed model that is shown to be successful at mapping SC-dFC-FC tripartite relationship, several expectations arise such as reproducibility of discovered states and their corresponding predicted FCs, sensitivity of the model to the underlying anatomical structure, state-specificity of the tMKL model and importantly verifying that the model does not overfit the training data. Several experiments were conducted in order to verify that the model satisfies these claims and the results presented in Section 3 point to the robustness of the performance of the model.

After learning the model parameters over half of the CUB dataset, we tested is on the HCP dataset. This implies that the transitions are generalizable over another cohort and are not restricted only to the subjects used for learning. Thus, the temporal dynamics of fMRI time series are well represented by the learned Markov model. This clearly indicates that the model is sensitive to the individual SC differences as well as does not overfit to any particular dataset.

Further, in order to verify whether the state-specific FCs predicted for a subject are distinct, we performed community detection over the mean state-specific FCs of the test cohort (see Figure 10). As can be clearly seen in the figure, regions in each state show distinct interaction patterns among themselves. The states seem to characterize the transient relationship among the ROIs which appear and disappear across the duration of the resting state scan. The markov chain state transition model further allows the characterization of the temporal fluctuations of the states that approximates latent temporal structure.

4.4. Future directions

As part of future work, it will be interesting to explore the biophysical meaning of the model parameters. Other direction could be to characterize the dynamics better by predicting the time series itself rather than working with correlation matrices. An immediate investigation would be to explore the relationship between the latent timeseries and the actual BOLD timeseries. Another line of work would be to apply the proposed model to characterize dFC in various conditions such as neurodegenerative and psychiatric diseases, lesions, healthy and pathological aging etc.

5. Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request. The codes for the tMKL model are available at <https://github.com/SriniwasGovindaSurampudi/tMKL>.

6. Acknowledgements

The authors would like to thank Arpan Banerjee for his insightful feedback and proofreading the manuscript. DR is supported by the Ramalingaswami Fellowship (BT/RLF/Re-entry/07/2014) from Department of Biotechnology (DBT), Ministry of Science & Technology, Government of India.

Data were provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

References

- [1] O. Sporns, G. Tononi, R. Kötter, The human connectome: a structural description of the human brain, *PLoS computational biology* 1 (4) (2005) e42.

- [2] B. Biswal, F. Z. Yetkin, V. M. Haughton, J. S. Hyde, Functional connectivity in the motor cortex of resting human brain using echo-planar MRI, *Magnetic Resonance in Medicine* 34 (4) (1995) 537–541.
- [3] B. B. Biswal, M. Mennes, X.-N. Zuo, S. Gohel, C. Kelly, S. M. Smith, C. F. Beckmann, J. S. Adelstein, R. L. Buckner, S. Colcombe, et al., Toward discovery science of human brain function, *Proceedings of the National Academy of Sciences* 107 (10) (2010) 4734–4739.
- [4] M. P. Van Den Heuvel, H. E. H. Pol, Exploring the brain network: a review on resting-state fmri functional connectivity, *European neuropsychopharmacology* 20 (8) (2010) 519–534.
- [5] S. Mori, P. van Zijl, Fiber tracking: principles and strategies—a technical review, *NMR in Biomedicine* 15 (7-8) (2002) 468–480.
- [6] G. Gong, Y. He, L. Concha, C. Lebel, D. W. Gross, A. C. Evans, C. Beaulieu, Mapping anatomical connectivity patterns of human cerebral cortex using in vivo diffusion tensor imaging tractography, *Cerebral cortex* 19 (3) (2008) 524–536.
- [7] J. L. Vincent, G. H. Patel, M. D. Fox, A. Z. Snyder, J. T. Baker, D. C. Van Essen, J. M. Zempel, L. H. Snyder, M. Corbetta, M. E. Raichle, Intrinsic functional architecture in the anaesthetized monkey brain, *Nature* 447 (7140) (2007) 83.
- [8] C. Chang, G. H. Glover, Time-frequency dynamics of resting-state brain connectivity measured with fmri, *Neuroimage* 50 (1) (2010) 81–98.
- [9] I. Cribben, R. Haraldsdottir, L. Y. Atlas, T. D. Wager, M. A. Lindquist, Dynamic connectivity regression: determining state-related changes in brain connectivity, *Neuroimage* 61 (4) (2012) 907–920.
- [10] M. G. Preti, T. A. Bolton, D. Van De Ville, The dynamic functional connectome: State-of-the-art and perspectives, *NeuroImage* 160 (2017) 41–54.
- [11] E. A. Allen, E. Damaraju, S. M. Plis, E. B. Erhardt, T. Eichele, V. D. Calhoun, Tracking whole-brain connectivity dynamics in the resting state, *Cerebral cortex* 24 (3) (2014) 663–676.
- [12] V. Kiviniemi, T. Vire, J. Remes, A. A. Elseoud, T. Starck, O. Tervonen, J. Nikkinen, A sliding time-window ica reveals spatial variability of the default mode network in time, *Brain connectivity* 1 (4) (2011) 339–347.
- [13] G. Deco, M. L. Krriegelbach, V. K. Jirsa, P. Ritter, The dynamics of resting fluctuations in the brain: metastability and its dynamical cortical core, *Scientific Reports* 7 (3095).
- [14] K. Vemuri, B. R. Surampudi, An exploratory investigation of functional network connectivity of empathy and default mode networks in a free-viewing task, *Brain Connectivity* 5 (6) (2015) 384–400.
- [15] A. P. Baker, M. J. Brookes, I. A. Rezek, S. M. Smith, T. Behrens, P. J. P. Smith, M. Woolrich, Fast transient networks in spontaneous human brain activity, *Elife* 3.

- [16] S. Ryali, K. Supekar, T. Chen, J. Kochalka, W. Cai, J. Nicholas, A. Padmanabhan, V. Menon, Temporal dynamics and developmental maturation of salience, default and central-executive network interactions revealed by variational bayes hidden markov modeling, *PLoS Computational Biology* 12 (12) (2016) e1005138.
- [17] J. Taghia, S. Ryali, T. Chen, K. Supekar, W. Cai, V. Menon, Bayesian switching factor analysis for estimating time-varying functional connectivity in fmri, *NeuroImage* 155 (2017) 271–290.
- [18] D. Vidaurre, S. M. Smith, M. W. Woolrich, Brain network dynamics are hierarchically organized in time, *Proceedings of the National Academy of Sciences* 114 (48) (2017) 12827–12832.
- [19] J. Taghia, W. Cai, S. Ryali, J. Kochalka, J. Nicholas, T. Chen, V. Menon, Uncovering hidden brain state dynamics that regulate performance and decision-making during cognition, *Nature communications* 9 (1) (2018) 2505.
- [20] S. M. Smith, K. L. Miller, S. Moeller, J. Xu, E. J. Auerbach, M. W. Woolrich, C. F. Beckmann, M. Jenkinson, J. Andersson, M. F. Glasser, et al., Temporally-independent functional modes of spontaneous brain activity, *Proceedings of the National Academy of Sciences* 109 (8) (2012) 3131–3136.
- [21] P. Robinson, Interrelating anatomical, effective, and functional brain connectivity using propagators and neural field theory, *Physical Review E* 85 (1) (2012) 011912.
- [22] G. Deco, A. Ponce-Alvarez, D. Mantini, G. L. Romani, P. Hagmann, M. Corbetta, Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations, *Journal of Neuroscience* 33 (27) (2013) 11239–11252.
- [23] V. Pernice, B. Staude, S. Cardanobile, S. Rotter, How structure determines correlations in neuronal networks, *PLoS computational biology* 7 (5) (2011) e1002059.
- [24] S. G. Surampudi, S. Naik, A. Shrama, R. S. Bapi, D. Roy, Combining multiscale diffusion kernels for learning the structural and functional brain connectivity, *bioRxiv* (2016) 078766.
- [25] C. O. Becker, S. Pequito, G. J. Pappas, M. B. Miller, S. T. Grafton, D. S. Bassett, V. M. Preciado, Spectral mapping of brain functional connectivity from diffusion imaging, *Scientific reports* 8 (1) (2018) 1411.
- [26] F. Abdehnour, H. U. Voss, A. Raj, Network diffusion accurately models the relationship between structural and functional brain connectivity networks, *Neuroimage* 90 (2014) 335–347.
- [27] S. G. Surampudi, S. Naik, R. B. Surampudi, V. K. Jirsa, A. Sharma, D. Roy, Multiple kernel learning model for relating structural and functional connectivity in the brain, *Scientific reports* 8 (1) (2018) 3265.

- [28] F. Abdelnour, M. Dayan, O. Devinsky, T. Thesen, A. Raj, Functional brain connectivity is predictable from anatomic network's laplacian eigenstructure, *NeuroImage* 172 (2018) 728–739.
- [29] A. S. Pillai, V. K. Jirsa, Symmetry breaking in space-time hierarchies shapes brain dynamics and behavior, *Neuron* 94 (5) (2017) 1010–1026.
- [30] M. Schirner, S. Rothmeier, V. K. Jirsa, A. R. McIntosh, P. Ritter, An automated pipeline for constructing personalized virtual brains from multimodal neuroimaging data, *Neuroimage* 117 (2015) 343–357.
- [31] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, K. Ugurbil, W.-M. H. Consortium, et al., The wu-minn human connectome project: an overview, *Neuroimage* 80 (2013) 62–79.
- [32] M. Belkin, P. Niyogi, Laplacian eigenmaps and spectral techniques for embedding and clustering, in: *Advances in neural information processing systems*, 2002, pp. 585–591.
- [33] M. Belkin, P. Niyogi, Laplacian eigenmaps for dimensionality reduction and data representation, *Neural computation* 15 (6) (2003) 1373–1396.
- [34] U. Von Luxburg, A tutorial on spectral clustering, *Statistics and computing* 17 (4) (2007) 395–416.
- [35] A. Y. Ng, M. I. Jordan, Y. Weiss, On spectral clustering: Analysis and an algorithm, in: *Advances in neural information processing systems*, 2002, pp. 849–856.
- [36] J. Shi, J. Malik, Normalized cuts and image segmentation, *IEEE Transactions on pattern analysis and machine intelligence* 22 (8) (2000) 888–905.
- [37] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer-Verlag New York, 2016.
- [38] F. R. Chung, F. C. Graham, *Spectral graph theory*, no. 92, American Mathematical Soc., 1997.
- [39] S. Moeller, E. Yacoub, C. A. Olman, E. Auerbach, J. Strupp, N. Harel, K. Ugurbil, Multiband multislice ge-epi at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fmri, *Magnetic resonance in medicine* 63 (5) (2010) 1144–1153.
- [40] D. A. Feinberg, S. Moeller, S. M. Smith, E. Auerbach, S. Ramanna, M. F. Glasser, K. L. Miller, K. Ugurbil, E. Yacoub, Multiplexed echo planar imaging for sub-second whole brain fmri and fast diffusion imaging, *PloS one* 5 (12) (2010) e15710.
- [41] K. Setsompop, B. A. Gagoski, J. R. Polimeni, T. Witzel, V. J. Wedeen, L. L. Wald, Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty, *Magnetic resonance in medicine* 67 (5) (2012) 1210–1224.

- [42] J. Xu, S. Moeller, J. Strupp, E. Auerbach, L. Chen, D. Feinberg, K. Ugurbil, E. Yacoub, Highly accelerated whole brain imaging using aligned-blipped-controlled-aliasing multiband epi, in: Proceedings of the 20th Annual Meeting of ISMRM, Vol. 2306, 2012.
- [43] M. Rubinov, O. Sporns, Complex network measures of brain connectivity: uses and interpretations, *Neuroimage* 52 (3) (2010) 1059–1069.
- [44] M. Xia, J. Wang, Y. He, Brainnet viewer: a network visualization tool for human brain connectomics, *PloS one* 8 (7) (2013) e68910.
- [45] E. Damaraju, E. A. Allen, A. Belger, J. Ford, S. McEwen, D. Mathalon, B. Mueller, G. Pearlson, S. Potkin, A. Preda, et al., Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia, *NeuroImage: Clinical* 5 (2014) 298–308.
- [46] B. Rashid, E. Damaraju, G. D. Pearlson, V. D. Calhoun, Dynamic connectivity states estimated from resting fmri identify differences among schizophrenia, bipolar disorder, and healthy control subjects, *Frontiers in human neuroscience* 8 (2014) 897.
- [47] R. Hindriks, M. H. Adhikari, Y. Murayama, M. Ganzetti, D. Mantini, N. K. Logothetis, G. Deco, Can sliding-window correlations reveal dynamic functional connectivity in resting-state fmri?, *Neuroimage* 127 (2016) 242–256.