SPARSE COUPLED HIDDEN MARKOV MODELS SHED LIGHT ON RESTING-STATE FMRI CROSS-NETWORK INTERACTIONS

Thomas Bolton[1,2] *and Dimitri Van De Ville*[1,2]

[1] Institute of Bioengineering, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
[2] Department of Radiology and Medical Informatics, University of Geneva (UNIGE), Switzerland

ABSTRACT

Dynamic functional connectivity (dFC) analysis aims at understanding how interactions across the brain resting-state networks (RSNs) evolve over time. Here, we introduce a novel methodological framework operating at the level of RSN activity time courses. Through the use of coupled hidden Markov models (CHMMs), we model cross-network couplings, i.e. the ability of one RSN to influence state transitions of the others. Because such modulatory influences are not expected across all possible pairs of RSNs, we combine this modeling strategy with ℓ_1 regularisation to derive a sparse set of cross-network modulatory coefficients.
As a validation of this framework, we first demonstrate the ability of the sparse CHMM approach to disentangle intrinsic state transition probabilities from external modulatory influences on an artificially generated dataset. We then perform preliminary analyses on a real resting-state dataset, using RSN activity time courses derived from a state-of-the-art deconvolution technique as inputs to our framework, and shed light on several significant cross-network couplings across major RSNs.

***Index Terms*—** Functional Magnetic Resonance Imaging, Dynamic Functional Connectivity, Coupled Hidden Markov Models, ℓ_1 Regularisation, Cross-Network Coupling

1 Introduction

The brain is active even at rest, and studying the relationship between its core resting-state networks (RSNs) by functional magnetic resonance imaging (fMRI) is an active field of research. Recently, the non-stationary nature of those interactions has been put forward [1], giving rise to the growing field of dynamic functional connectivity (dFC) [2].
Resting-state activity is at least partly made of spatiotemporal profiles that repeatedly occur over time [3]; thus, the brain evolves towards precise state configurations that depend on the previous time points. To date, only few efforts have been undertaken to explicitly model this temporal evolution: one strategy is to impose smoothness on the temporal evolution of connectivity estimates computed between brain regions [4], while another is the assumption that RSNs evolve in activity over time independently from each other [5].
A third strategy lies in modeling both individual RSN temporal behaviours, and also causal relationships across networks, where the activity level of one RSN will influence the others. This has recently been tackled through the use of coupled hidden Markov models (CHMMs) [6], with three limitations: (1) modeling was performed at the level of connectivity time courses derived with a sliding window approach [7], diminishing temporal resolution compared to a frame-wise analysis of activity; (2) RSNs and associated activity profiles were generated using spatial Independent Component Analysis (SICA), a stationary tool for which obtained RSNs are not always optimal; (3) to alleviate computational burden, modeling was downscaled to a subset of selected RSNs, in a one versus the others scheme.
In order to overcome those limitations, inspired by previous work in the genomics field [8], we introduce a novel sparse CHMM framework applied at the level of RSN activity time courses, with embedded ℓ_1 regularisation [9]. This effectively enables larger-scale analyses of RSN-to-RSN relationships, while conserving frame-wise temporal resolution. We first validate the implementation of this framework on artificially generated data with added cross-network coupling. We then apply it to real spontaneous activity traces recovered by a state-of-the-art deconvolution approach combining total activation (TA) [10] and subsequent generation of innovation-driven co-activation patterns (iCAPs) [11], which stand as our RSNs here. In this framework, the iCAPs are extracted through the detection of transient activity, and the deconvolved time courses are undone from the hemodynamic effect. This allows for temporal overlap, opening the door for the investigation of RSN coupling.

2 Methods

2.1 Sparse CHMM Framework

Let $h_t^{(k)}$ be the hidden state of network k out of N at time point t; in the following, we make the simplifying assumption that each network can be in one of three states: *deactive* ($h_t^{(k)} = -1$), *baseline* ($h_t^{(k)} = 0$), or *active* ($h_t^{(k)} = +1$).

This work was supported by the Swiss National Science Foundation (grant 205321-163376), the Bertarelli Foundation, the Center for Biomedical Imaging (CIBM) of Geneva and Lausanne Universities and the Leenaards and Louis-Jeantet Foundations. We thank F. I. Karahanoğlu for providing the resting-state data analysed here.

In a classical parallel hidden Markov model (HMM) framework, each network is parameterised by its probability to start in state i, $\Pi_i^{(k)} = \mathcal{P}(h_1^{(k)} = i)$, and its probability to transit from state i to state j (notwithstanding the activity of the other networks), $A_{i\to j}^{(k)} = \mathcal{P}(h_{t+1}^{(k)} = j | h_t^{(k)} = i)$.

In order to sparsely model possible couplings across networks, we deploy a regularised multinomial framework [12], where the effective transition probability of network k from state i to state j at time point t depends on the activity level of the other networks, and is given by:

$$B_{t,i\to j}^{(k)} = \mathcal{P}(h_{t+1}^{(k)} = j | h_t^{(k)} = i, \mathbf{h_t^{(-k)}}) = \frac{e^{\beta_{0,i\to j}^{(k)} + \sum_{l\neq k} \beta_{l,i\to j}^{(k)} h_t^{(l)}}}{\sum_{m\in\mathcal{S}} e^{\beta_{0,i\to m}^{(k)} + \sum_{l\neq k} \beta_{l,i\to m}^{(k)} h_t^{(l)}}}.$$

In this equation, $\beta_{0,i\to j}^{(k)}$ is the baseline coefficient for the i to j transition, and $\beta_{l,i\to j}^{(k)}$ is the coefficient describing the influence of network l on network k, which impacts on the transition probability if $h_t^{(l)} \neq 0$. $\mathcal{S} = \{-1, 0, +1\}$ is the set of activity states, and $\mathbf{h_t^{(-k)}}$ depicts the activity level of all networks else than network k at time point t. Figure 1A provides a schematic view on the modulatory influences of RSNs onto each other when they are active/deactive.

If all coefficients were to be inferred from the data, this would amount to $9N^2$ values. Aside from the related computational burden, which can become difficult to handle for large values of N, it is also physiologically implausible that all RSNs influence each other; rather, we are more likely to observe only particular subsets of interactions amongst functionally related systems. Thus, we impose the following ℓ_1 regularisation to the set of coefficients:

$$\sum_{l\neq k} |\beta_{l,i\to j}^{(k)}| \leq \rho_k \forall k \in \{1, ..., N\}, i, j \in \{-1, 0, +1\}.$$

The regularisation parameter ρ_k is allowed to differ across networks, because there may be more modulatory influences onto some RSNs than others. To solve for a given network k and starting state i, we can individually consider the set of coefficients related to each end state j by forming a partial quadratic approximation to the log-likelihood. We must then minimise:

$$\frac{1}{2D} \sum_{t\in\mathcal{C}} [\omega_{t,j} (z_{t,j} - \beta_{0,i\to j}^{(k)} - \sum_{l\neq k} \beta_{l,i\to j}^{(k)} h_t^{(l)})^2] + \lambda_k \sum_{l\neq k} |\beta_{l,i\to j}^{(k)}|.$$

In the above equation, $\mathcal{C} = \{t : h_t^{(k)} = i\}$ is the set of considered data points, and D the number of such data points. The coefficients of this regularised least square problem read:

$$\begin{cases} \omega_{t,j} = \tilde{B}_{t,i\to j}^{(k)} (1 - \tilde{B}_{t,i\to j}^{(k)}) \\ z_{t,j} = \tilde{\beta}_{0,i\to j}^{(k)} + \sum_{l\neq k} \tilde{\beta}_{l,i\to j}^{(k)} h_t^{(l)} + \frac{y_{t,j}^{(k)} - \tilde{B}_{t,i\to j}^{(k)}}{\omega_{t,j}}. \end{cases} \quad (1)$$

In this expression, $y_{t,j}^{(k)} = \delta_{h_{t+1}^{(k)},j}$ (where δ is the Kronecker delta), $\tilde{\beta}$ denotes the current estimate of a regression coefficient, and $\tilde{B}$ represents the current estimate of an effective transition probability.

In practice, we first run parallel HMMs on all networks to derive hidden state estimates and transition probabilities. Then, we set $\tilde{\beta}_{l,i\to j}^{(k)} = 0$ and $\tilde{\beta}_{0,i\to j}^{(k)} = A_{i\to j}^{(k)}$, and perform the following sequentially for all networks: select the data points $y_{t,j}^{(k)}$ and $\{h_t^{(l)}\}_{l\neq k}$ within the working set $\mathcal{C}$; loop through the end state j, which involves the update of $\omega_{t,j}$ and $z_{t,j}$ according to (1) and the extraction of new coefficient values by coordinate descent [13]; exit the loop if the log-likelihood change is below $\epsilon = 10^{-6}$.

In order to select the regularisation parameters to use, the above process is iterated for gradually decreasing values of λ_k, using warm restarts for coefficient estimates. For all the presented results, we selected λ_k giving the lowest Bayesian Information Criterion (BIC) [14] for each network, probing the interval of values $[1000, 0.5]$.

For simplicity, in all the presented analyses, we focused on transitions from the baseline activity state, that is, $\mathcal{C} = \{t : h_t^{(k)} = 0\}$. For real data analyses, we focused on modulatory coefficients from baseline to active state ($\beta_{l,0\to+1}^{(k)}$).

3 Experimental Results

3.1 Artificial Data

To validate our framework, we first generated $n_s = 20$ sets of $N = 3$ artificial network time courses long of $T = 300$ time points, characterised by the following transition matrices:

$$A^{(1)} = A^{(3)} = \begin{bmatrix} 0.1 & 0.9 & 0 \\ 0.05 & 0.9 & 0.05 \\ 0 & 0.9 & 0.1 \end{bmatrix}, A^{(2)} = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.3 & 0.4 & 0.3 \\ 0 & 0.5 & 0.5 \end{bmatrix}.$$

Figure 1B (top left panel) illustrates a portion of one of the generated time courses. With those parameters, network 2 often visits the active and deactive states, whereas networks 1 and 3 remain at baseline activity most of the time and exhibit a similar behaviour.

To study the ability of our sparse CHMM framework to disentangle intrinsic transition propensity from external coupling effects, we also included an influence of network 2 onto network 1 modulating transition probabilities as follows:

$$\hat{A}^{(1)} = \begin{bmatrix} 0.1 - 0.1h_t^{(2)} & 0.9 + 0.1h_t^{(2)} & 0 \\ 0.05 & 0.9 - 0.8h_t^{(2)} & 0.05 + 0.8h_t^{(2)} \\ 0 & 0.9 - 0.1h_t^{(2)} & 0.1 + 0.1h_t^{(2)} \end{bmatrix} \text{ if } h_t^{(2)} = +1.$$

As can be seen from Figure 1B (bottom left panel), network 1 is then attracted into the active state when network 2 is itself active, and the time course of network 1 now differs from the one of network 3.

We ran our sparse CHMM framework on this artificial case, and to establish significance of the modulatory coefficients, compared our results against the ones generated from null data in which each network time course from each set was independently circularly shifted along time by

$\Delta S \in \{0, ..., T\}$. We generated 40 null datasets, and for each modulatory coefficient, used the minimal and maximal obtained values as the 2.5$^{\text{th}}$ and 97.5$^{\text{th}}$ percentiles for significance assessment (α-level 0.05).
The regularisation parameters minimising the BIC were $\lambda_1 = \lambda_2 = \lambda_3 = 105$. The only significant modulatory coefficient was $\beta^{(1)}_{2,0\to+1} = 4.2933$, which makes sense as we simulated an attraction of network 1 into the active state by network 2.

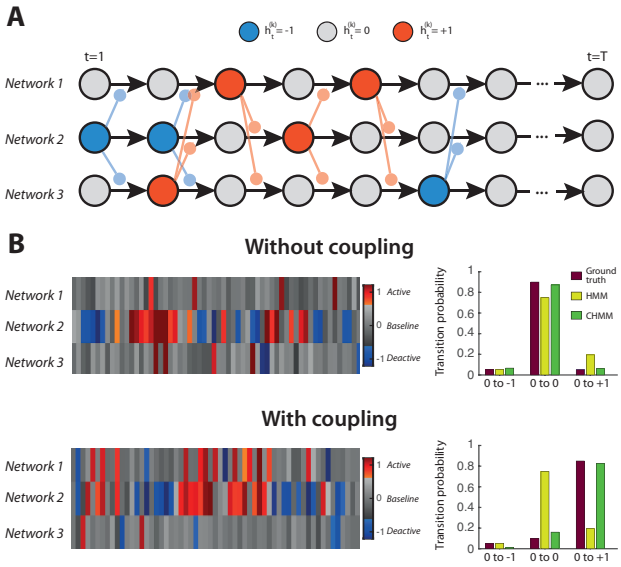

Fig. 1: **(A)** Schematic depiction of transitions (as denoted by black arrows) across three states of activity (blue: deactive, gray: baseline, red: active) for three networks. When a given network is (de)active, it can influence the transition probability of the others for their subsequent transition (red and blue modulatory arrows). **(B)** Comparison between parallel HMM and sparse CHMM approaches on an artificially generated dataset. Networks 1 and 3 have the same transition probabilities without external modulatory influences (top left panel), but we also introduce an influence of network 2 onto network 1 that attracts it into the active state (bottom left panel). Whereas the sparse CHMM framework can disentangle intrinsic transition probabilities from modulatory influences, parallel HMMs cannot as they only provide one transition probability estimate (right bar graphs).

To determine the ability of our sparse CHMM framework to disentangle intrinsic transition propensity from external influences, we compared real transition probabilities from the baseline state to the ones extracted by parallel HMMs, and by the sparse CHMM approach (Figure 1B, right panel). Parallel HMMs cannot disentangle both effects, and so in a setting where no modulation is present, the $0 \to 0$ probability is underestimated, whereas the $0 \to +1$ probability is overestimated (top bar graph). In the case where coupling is introduced, the converse is observed (bottom bar graph). With a sparse CHMM approach, however, excellent probability estimates are retrieved for both settings.

3.2 Real Data

For real data analyses, we applied our sparse CHMM framework to the resting-state data from $n_s = 20$ healthy volunteers (38.4±6 years old) acquired with a Siemens 3T Trio TIM scanner, using a 32-channel head coil and gradient-echo echo-planar imaging (TR/TE/FA=1.1s/27ms/90°, matrix=64x64, voxel size=3.75 x 3.75 x 5.63 mm^3, 21 slices). We analysed $T = 264$ volumes from the iCAP time courses previously extracted from this dataset in [11].
In order to maximise the number of observed configurations of the system, we selected the $N = 9$ networks displaying most apparent state transitions for our analyses (see Figure 2 for illustrations). We derived a set of significant modulatory coefficients using the same null data strategy as outlined in the artificial data section. For clarity, we then focused our interpretation on only the largest significant coefficients ($|\beta^{(k)}_{l,i\to j}| > 0.5$).
Optimal regularisation parameters were: $\lambda_1 = 9$, $\lambda_2 = 16$, $\lambda_3 = 7$, $\lambda_4 = 7$, $\lambda_5 = 5$, $\lambda_6 = 7$, $\lambda_7 = 16$, $\lambda_8 = 5$ and $\lambda_9 = 9$. Those values are different across networks and lower than in the artificial case, hinting at the facts that (1) different networks may be modulated to different extents, and (2) sparsity of modulatory influences may be less strong in real data compared to our simulated example.

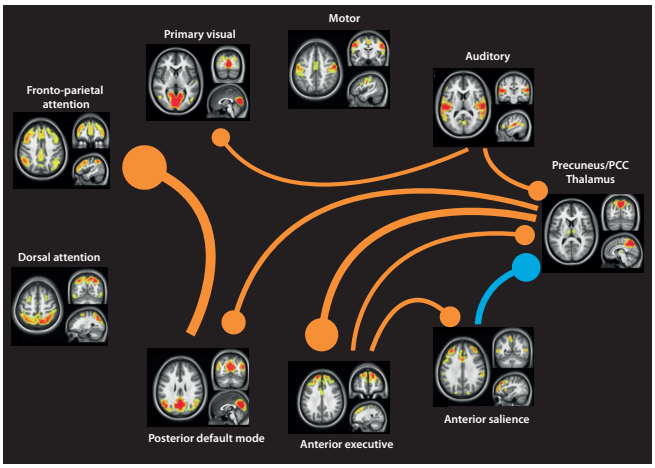

Fig. 2: The 9 iCAPs selected for sparse CHMM analysis, entitled as suggested in [11], and modulatory influences across them (orange/blue: positive/negative modulatory coefficient, meaning that if the modulatory network is active, the modulated network has a larger/lower probability to become active from baseline state. Thickness of the modulatory arrows is proportional to the strength of the modulatory coefficients.

19 significant modulatory coefficients were retrieved (out of the 72 possible ones; recall that we focus on transitions from the baseline state to the active state only). In Figure 2, for clarity, we only display the top 8 modulatory influences across iCAPs: 7 of those 8 coefficients were positive, with the modulated network attracted into the active state when the modulating network is active. Only the anterior salience iCAP had a negative modulatory influence on the precuneus/PCC/thalamus iCAP; i.e., decreasing its propensity to become active when being itself active.
Positive couplings were observed both between sensory networks (auditory to visual), as well as between several iCAPs reflective of subparts from the salience, executive and default mode networks. The largest such modulation was positive

coupling from the posterior default mode iCAP to the fronto-parietal attention one.

4 Discussion and Extensions

Through explicit modeling of cross-network couplings by a sparse CHMM framework, we could reveal a set of yet uncharacterised modulatory influences across some of the key brain RSNs. Many of those couplings unsurprisingly involved players from the salience/executive/default mode triad, which are known to dynamically interact across time upon resting-state to give rise to the content of mind wandering [15,16].

Particularly intriguing from those preliminary results is the positive coupling between the posterior default mode and fronto-parietal attention networks, which are classically believed to be anti-correlated at rest [17]. This may highlight the fact that particular subsystems of those networks may exhibit more complex relationships than the ones known to date. Alternatively, it may be that, while the posterior default mode network drives the fronto-parietal attention network towards an active state, that fronto-parietal attention network also promotes the shift of the posterior default mode network back to a baseline state of activity. Future analysis of modulatory influences from the active state ($\beta^{(k)}_{l,+1\rightarrow j}$) could help in answering this question.

We also note that the present framework holds some limitations: first, we assume that the transition probabilities and modulatory influences are the same across subjects, whereas there is in fact enough inter-individual heterogeneity in resting-state configurations to permit the accurate fingerprinting of single subjects [18]. Second, we assume that the modulatory strength of a deactive network is the same as when active with opposite sign, which may be an oversimplification. Finally, there are also alternatives to CHMM modeling, such as through fully-linked HMMs or dynamically multi-linked HMMs [19]; the comparison of those different approaches may consist in an interesting direction for future work.

5 Conclusion

Here, we introduced a novel framework combining coupled hidden Markov models and ℓ_1 regularisation, to describe modulatory influences across RSNs. We showed the presence of such interactions in resting-state, and presented preliminary results demonstrating the value of this added characterisation. We believe that more complete analyses with this framework, and possible further methodological extensions, shall pave the way towards more accurate dFC analyses.

6 References

[1] C. Chang, G. Glover, "Time-frequency dynamics of resting-state brain connectivity measured with fMRI", *NeuroImage*, vol. 50, pp. 81-98, 2010.

[2] R. M. Hutchison, T. Womelsdorf, E. A. Allen, P. A. Bandettini, V. D. Calhoun et al., "Dynamic functional connectivity: promise, issues, and interpretations", *Neuroimage*, vol. 80, pp. 360-378, 2013.

[3] W. Majeed, M. Magnuson, W. Hasenkamp, H. Schwarb, E. H. Schumacher et al., "Spatiotemporal dynamics of low frequency BOLD fluctuations in rats and humans", *Neuroimage*, vol. 54(2), pp. 1140-1150, 2011.

[4] R. P. Monti, P. Hellyer, D. Sharp, R. Leech, C. Anagnostopoulos et al., "Estimating time-varying brain connectivity networks from functional MRI time series", *Neuroimage*, vol. 103, pp. 427-443, 2014.

[5] S. M. Smith, K. L. Miller, S. Moeller, J. Xu, E. J. Auerbach et al., "Temporally-independent functional modes of spontaneous brain activity", *Proc Natl Acad Sci U S A*, vol. 109(8), pp. 3131-3136, 2012.

[6] M. Sourty, L. Thoraval, D. Roquet, J. P. Armspach, J. Foucher et al., "Identifying dynamic functional connectivity changes in dementia with lewy bodies based on product hidden markov models", *Front Comput Neurosci*, vol. 10(60), pp. 1-11, 2016.

[7] E. A. Allen, E. Damaraju, S. M. Plis, E. B. Erhardt, T. Eichele et al., "Tracking whole-brain connectivity dynamics in the resting state", *Cereb Cortex*, vol. 24(3), pp. 663-676, 2014.

[8] H. Choi, D. Fermin, A. I. Nesvizhskii, D. Ghosh, Z. S. Qin, "Sparsely correlated hidden Markov models with application to genome-wide location studies", *Bioinformatics*, vol. 29(5), pp. 533-541, 2013.

[9] R. Tibshirani, "Regression shrinkage and selection via the lasso", *J R Stat Soc Series B Stat Methodol*, vol. 58(1), pp. 267-288, 1996.

[10] F. I. Karahanoğlu, C. Caballero-Gaudes, F. Lazeyras, D. Van De Ville, "Total activation: fMRI deconvolution through spatio-temporal regularization", *Neuroimage*, vol. 73, pp. 121-134, 2013.

[11] F. I. Karahanoğlu, D. Van De Ville, "Transient brain activity disentangles fMRI resting-state dynamics in terms of spatially and temporally overlapping networks", *Nat Commun*, vol. 6, pp. 1-10, 2015.

[12] J. Friedman, T. Hastie, R. Tibshirani, "Regularization paths for generalized linear models via coordinate descent", *J Stat Softw*, vol. 33(1), pp. 1-22, 2010.

[13] J. Friedman, T. Hastie, H. Höfling, R. Tibshirani, "Pathwise coordinate optimization", *Ann Appl Stat*, vol. 1(2), pp. 302-332, 2007.

[14] G. Schwarz, "Estimating the dimension of a model", *Ann Appl Stat*, vol. 6(2), pp. 461-464, 1978.

[15] V. Menon, "Large-scale brain networks and psychopathology: a unifying triple network model", *Trends Cogn Sci*, vol. 15(10), pp. 483-506, 2011.

[16] K. Christoff, Z. C. Irving, K. C. Fox, R. N. Spreng, J. R. Andrews-Hanna, "Mind-wandering as spontaneous thought: a dynamic framework", *Nat Rev Neurosci*, vol. 17, pp. 718-731, 2016.

[17] M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen et al., "The human brain is intrinsically organized into dynamic, anticorrelated functional networks", *Proc Natl Acad Sci U S A*, vol. 102(27), pp. 9673-9678, 2005.

[18] E. S. Finn, X. Shen, D. Scheinost, M. D. Rosenberg, J. Huang et al., "Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity", *Nat Neurosci*, vol. 18, pp. 1664-1671, 2015.

[19] L. Zhang, D. Samaras, N. Alia-Klein, N. Volkow, R. Goldstein, "Modeling neuronal interactivity using dynamic bayesian networks", *Adv Neural Inf Process Syst*, vol. 18, pp. 1593-1600, 2006.