

GRAPHICAL MODELS AND DYNAMIC LATENT FACTORS FOR MODELING FUNCTIONAL BRAIN CONNECTIVITY

Andersen Chang^a, Tianyi Yao^a, and Genevera I. Allen^{b,c}

^aDept. of Statistics, Rice University

^bDept. of Statistics, Computer Science, and Electrical and Computer Engineering, Rice University

^cNeurological Research Institute, Baylor College of Medicine
Houston, TX

ABSTRACT

With modern technology, the activity of thousands of neurons in the brain can be recorded simultaneously. Such data can potentially shed light on how neurons communicate with one another. These neuronal interactions are often viewed under the framework of functional connectivity, which is defined as the statistical dependence between recorded neuronal activity. Several have proposed to use graphical models to estimate functional connectivity between neurons directly from neuronal recording data. However, one challenge that can arise from this type of data that is not addressed by a traditional graphical model is the influence of dynamic latent brain states on recorded neuronal activity, as the neurons recorded in one experimental session constitute only a small subset of all the neurons in the brain. These latent states should be accounted for to get a more accurate estimate of functional connectivity. In this paper, we introduce two models, the dynamic mean operator (DYNAMO) and the dynamic covariance operator (DYNACO) conditional Gaussian graphical models, to infer functional connectivity from neuronal activity data after adjusting for dynamic latent brain states. We apply the DYNAMO and DYNACO models to a variety of simulation studies and demonstrate their superior performance over traditional, unconditional graphical models.

Index Terms— Latent variable conditional graphical model, Gaussian graphical model, functional PCA, functional connectivity

1. INTRODUCTION

A plethora of new neuroimaging technologies have allowed for measurements of large-scale neuronal activity in vivo [3]. Given such data, a fundamental question in modern neuroscience is how individual neurons in the brain interact with each other. Such neuronal interactions are often considered

under the framework of functional connectivity, which is defined as the statistical (as opposed to structural) relationships between different neurons [1]. Results from studies of functional connectivity can be used to help researchers better understand how brain processes information and how different parts of the brain communicate in ways that may not be captured solely by examining physical synaptic connections.

Modern technology has allowed for the simultaneous recording of the activity of thousands of neurons in localized chunks of the brain [3]. Functional connectivity is often estimated through calculating Pearson correlations between the activity of pairs of neurons [2]. The calculation of individual pairwise correlations is generally a poor choice for analyzing such high-dimensional neuron activity data, though, as it does not capture the full dependence structure amongst all of the observed neurons. One of the ways that multi-neuronal relationships can be more accurately modeled is through undirected graphical models. These models have been used in a wide variety of applications, including genetics, fMRI, and cell networks [5,6] to explore network structures and conditional dependency relationships in data in high-dimensional settings. In the neuroscience setting, graphical models can be used to examine conditional dependence structures between neurons by identifying pairs of neurons that are predictive of each other's behavior given the information from all of the other neurons observed in the data, helping shed light on functional connectivity structures.

Unconditional graphical models, though, may not be able account for all of the potential interactions between neurons in the brain. First, the number of neurons that can be observed simultaneously is still an extremely small proportion of the total number that are in the brain. These unobserved neurons could have an influence on the activities of the observed neurons in the data. Additionally, other dynamics, such as visual stimuli and physical movements, could affect how neurons interact with one another. Previous neuroscientific studies have tried to account for such latent effects using static conditional latent variable graphical models to estimate the dependency structure from neuronal activity data [3,4].

The authors acknowledge support from NSF DMS-1554821 and NSF NeuroNex-1707400.

However, other studies have shown that latent brain state effects can change over time. Thus, the effects of these latent states have on both the functional connectivity between observed neurons in the data as well as between unobserved and observed neurons can be dynamic [7], something that the static latent variable graphical model methods do not account for. It may be more pertinent to instead apply dynamic latent variable conditional graphical model methods, similar to those introduced in [8, 9], to estimate functional connectivity from high-dimensional neuronal activity data. For this specific application, we would like to estimate a static graph that serves as a summary of functional connectivity between observed neurons after accounting for dynamic latent states that represent the collective effects of the unobserved neurons in the brain on the conditional relationships between observed neurons.

Our goal for this paper is to investigate the potential use of dynamic latent variable conditional graphical models to estimate a static graph to represent functional connectivity between recorded neurons in the presence of dynamic latent brain states. Below, we introduce two possible representations of this problem, along with estimation methods for both. In particular, we modify existing dynamic latent variable conditional graphical model methods to estimate a static graph after adjusting for dynamic latent variables, rather than a dynamic graph as done in the original methods.

The rest of the paper is organized as follows. In section 2, we describe in detail the models we use to represent functional connectivity adjusted for dynamic latent brain states and the associated estimation methods. We examine the performance of our proposed methods as well as traditional, unconditional graphical models in a variety of simulation studies in section 3. We conclude with a discussion of our results and potential future work in section 4.

2. DYNAMIC LATENT VARIABLE GRAPHICAL MODELS AND ESTIMATION

Below, we present two different approaches to estimate conditional independence relationships between observed neurons adjusted for dynamic latent brain states. We assume in both of these methods that the latent brain states vary smoothly with respect to time and are low rank, while the functional connectivity of the observed neurons conditioned on the latent brain states is static over the observed times in the data.

2.1. Dynamic Mean Operator (DYNAMO) Conditional Gaussian Graphical Model

Our DYNAMO model is closely related to the conditional Gaussian graphical model (cGGM) proposed in [19], which aims to estimate a sparse precision matrix representing conditional independence relationships among variables of interest after adjusting for the effects of covariates on the

mean structure. However, one key distinction between our DYNAMO model and the cGGM is that the effects of dynamic latent brain states in our case are not directly observable, whereas the covariates in the cGGM are usually observed quantities. The DYNAMO model assumes that the time-varying latent brain states have an effect on the mean of activity of the observed neurons and that these effects change smoothly as a function of time. Formally, let $\mathbf{X}_{p \times N} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_t, \dots, \mathbf{x}_N]$ be the observed activity for p neurons at N observed time points. Each column of $\mathbf{X}$ can then be modeled as:

$$\mathbf{x}_t = \mathbf{b}_t + \boldsymbol{\epsilon}_t, \\ \boldsymbol{\epsilon}_t \sim N(\mathbf{0}_p, \boldsymbol{\Theta}^{-1}).$$

where

$$\mathbf{B}_{p \times N} = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_t, \dots, \mathbf{b}_N],$$

with

$$\|\mathbf{B}\|_* \leq \tau \quad \text{and} \quad \text{trace}(\mathbf{B}\boldsymbol{\Omega}\mathbf{B}^T) \leq \alpha.$$

Here, the columns of $\mathbf{B}$ quantify the effects of the latent brain states on the observed neurons at time point $t = 1, \dots, N$. $\mathbf{B}$ is constrained to be low rank with $\|\mathbf{B}\|_* \leq \tau$ because the number of latent brain states is relatively small compared to the number of observed neurons. In addition, as the latent brain states are assumed to vary smoothly across time, we force the columns of $\mathbf{B}$ to change smoothly over time with $\text{trace}(\mathbf{B}\boldsymbol{\Omega}\mathbf{B}^T) \leq \alpha$, where $\boldsymbol{\Omega}_{N \times N} \succeq 0$ is a smoothing operator such as the second or fourth differences matrix [15]. Columns of $\mathbf{E}_{p \times N} = [\boldsymbol{\epsilon}_1, \dots, \boldsymbol{\epsilon}_N]$ are realizations of the conditional structures of the functional connectivity of observed neurons given the effects of the latent brain states $\mathbf{B}$. $\boldsymbol{\epsilon}_1, \dots, \boldsymbol{\epsilon}_N$ are assumed to be independently drawn from the same time-invariant p -variate Gaussian distribution $N(\mathbf{0}_p, \boldsymbol{\Theta}^{-1})$. The sparse, static precision matrix $\boldsymbol{\Theta}_{p \times p} \succ 0$ represents the true functional connectivity structures of observed neurons and $\boldsymbol{\Theta}$ is assumed to be independent of the latent brain dynamics $\mathbf{B}$.

Our ultimate goal in this case is to estimate the sparse, static precision matrix $\boldsymbol{\Theta}$, while also inferring and adjusting for the effects of latent brain states via the matrix $\mathbf{B}$. To this end, we could formulate our estimation problem in a penalized maximum likelihood framework to simultaneously estimate the underlying dynamic mean structure and the conditional graph. This would then become a biconvex problem in $\boldsymbol{\Theta}$ and $\mathbf{B}$, which could be solved via a flip-flop algorithm. However, such an iterative algorithm is known to be computationally intensive.

For simplicity, we propose a one-flop implementation as an approximation to the flip-flop algorithm. This one-flop estimation procedure is much less computationally intensive compared to flip-flop procedures. Specifically, with $\boldsymbol{\Theta}$ fixed

as the identity matrix, our problem essentially becomes the functional principal component analysis (PCA) problem [15], which produces estimates of functional principal loadings $\hat{\mathbf{V}}$ by solving the generalized eigenvalue problem $(\mathbf{X}^T \mathbf{X})\mathbf{V} = \alpha \Omega \mathbf{V}$. As a result, we can obtain estimates of the dynamic latent effects $\hat{\mathbf{B}} = \mathbf{X} \hat{\mathbf{V}}_K \hat{\mathbf{V}}_K^T$ with columns of $\hat{\mathbf{V}}_K$ consisting of the first K functional principal loadings. Subsequently, we can obtain an estimate of the sparse, static functional connectivity graph $\hat{\Theta}$ by fitting a graphical model, such as the graphical Lasso [14] or neighborhood selection approach [18], to the residuals $\hat{\mathbf{E}} = \mathbf{X} - \hat{\mathbf{B}}$ after adjusting for estimated dynamic latent effects. Thus, the model gives us estimates of both effects of dynamic latent brain states $\hat{\mathbf{B}}$ and a static conditional graph $\hat{\Theta}$ representing the functional connectivity of observed neurons. The method is analogous to centering the data before fitting a graphical model, except that functional PCA is used to remove a smooth, time-varying mean structure. Our one-flop estimation procedure is summarized in Algorithm 1.

Algorithm 1: One-Flop DYNAMO Algorithm

Input : $\mathbf{X} \in \mathbb{R}^{p \times N}$

(i) Estimate dynamic latent effects via functional PCA:

$\hat{\mathbf{B}} = \mathbf{X} \hat{\mathbf{V}}_K \hat{\mathbf{V}}_K^T$, where columns of $\hat{\mathbf{V}}_K$ consist of the first K functional principal loadings (the optimal K can be determined by a scree plot);

(ii) Estimate static functional connectivity graph:

Obtain estimated precision matrix $\hat{\Theta}$ by fitting a graphical model to the residual $\hat{\mathbf{E}} = \mathbf{X} - \hat{\mathbf{B}}$;

return $\hat{\mathbf{B}}, \hat{\Theta}$.

2.2. Dynamic Covariance Operator (DYNACO) Conditional Gaussian Graphical Model

As opposed to the DYNAMO model, in which we assume that the dynamic latent brain states affect the mean observed neural activity, the DYNACO model we introduce in this section posits that the dynamic latent brain states directly affect the functional connectivity of the observed neurons by changing the conditional relationships between latent brain states and observed neurons. The model follows from the Schur complement approach described in [8, 10]. As above, let $\mathbf{X}_{p \times N} = [\mathbf{x}_1, \dots, \mathbf{x}_N]$ be the observed activity for $i = 1, \dots, p$ neurons at $t = 1, \dots, N$ observed time points. In this situation, we model columns of $\mathbf{X}$ as:

$$\mathbf{x}_t \sim N(\mathbf{0}_p, \Theta_t^{-1}),$$

$$\Theta_t = \mathbf{S} - \mathbf{L}_t.$$

The distribution of $\mathbf{x}_t$ in the model has a dynamic marginal precision matrix Θ_t which changes at every time point. The Θ_t can be decomposed into the sum of a dynamic, low rank, smoothly time varying matrix $\mathbf{L}_t$ and a static, sparse conditional precision matrix $\mathbf{S}$. In this case, $\mathbf{L}_t$ serves as a summary of collective effects of the unobserved neurons in the brain on the conditional relationships among observed neurons and is assumed to vary smoothly over time. In addition, $\mathbf{L}_t$ is assumed to be low rank because we hypothesize that the number of collective latent brain states is relatively small compared to number of observed neurons. For the DYNACO model, we are mostly interested in estimating the static, sparse conditional precision matrix $\mathbf{S}$, which serves as the estimate of the functional connectivity among observed neurons conditional on effects of dynamic latent brain states.

Our approach to estimation for the above model is an extension of the latent variable graphical model proposed in [8] as well as its time-varying counterparts proposed in [10, 11]. In order to restrict the sparse component $\mathbf{S}$ to be static while allowing the low rank latent components $\mathbf{L}_t$ to change smoothly over time, we adopt the approach in [16] and use a kernel-smoothed estimate of the covariance matrix $\hat{\Sigma}_t = \frac{\sum_s \mathcal{K}(\frac{|s-t|}{h}) \mathbf{x}_s \mathbf{x}_s^T}{\sum_s \mathcal{K}(\frac{|s-t|}{h})}$ at each time point t with $\mathcal{K}(\frac{|s-t|}{h})$ given by a symmetric nonnegative kernel function of bandwidth h . Following from [8, 10], we propose the following penalized maximum likelihood approach to estimate both $\mathbf{S}$ and $\mathbf{L}_1, \dots, \mathbf{L}_N$ given a sequence of kernel-smoothed empirical covariance matrices $\hat{\Sigma}_1, \dots, \hat{\Sigma}_N$:

$$\begin{aligned} \underset{\mathbf{S}, \{\mathbf{L}_t\} \in \mathbb{S}_{++}^p}{\text{minimize}} \quad & \sum_{t=1}^N \left[-\log(\det(\mathbf{S} - \mathbf{L}_t)) + \text{trace}(\hat{\Sigma}_t(\mathbf{S} - \mathbf{L}_t)) \right] \\ & + \tau \sum_{t=1}^N \|\mathbf{L}_t\|_* + \alpha \|\mathbf{S}\|_{1,\text{off}} \end{aligned} \quad (1)$$

Here, $\mathbb{S}_{++}^p$ denotes the set of positive definite matrices of size p and $\{\mathbf{L}_t\}$ represents the sequence of dynamic, low rank matrices $\mathbf{L}_1, \dots, \mathbf{L}_N$. $\|\cdot\|_{1,\text{off}}$ is defined as an ℓ_1 -penalty on the off-diagonal terms of the matrix, which encourages sparsity in estimate of the sparse functional connectivity graph $\mathbf{S}$ of observed neurons conditional on effects of dynamic latent brain states. In addition, $\|\cdot\|_*$ denotes the nuclear norm, which encourages low-rankness in estimates of dynamic latent effects $\{\mathbf{L}_t\}$. Both α and τ are nonnegative regularization parameters, which determine the sparsity levels of $\hat{\mathbf{S}}$ and ranks of $\{\hat{\mathbf{L}}_t\}$, respectively.

In order to decouple some of the terms in problem (1) that are difficult to jointly optimize and develop a tractable algorithm to solve problem (1), we adopt the Alternating Direction Method of Multipliers (ADMM) framework described in [12] as well as the variable splitting scheme used in [10]. Specifically, we introduce a set of auxiliary variables $\{\Psi_t\}$ and recast problem (1) as follows:

$$\begin{aligned}
& \underset{\mathbf{S}, \{\mathbf{L}_t\}, \{\Psi_t\} \in \mathbb{S}_{++}^p}{\text{minimize}} \quad \sum_{t=1}^N \left[-\log(\det(\Psi_t)) + \text{trace}(\hat{\Sigma}_t \Psi_t) \right] \\
& \quad + \tau \sum_{t=1}^N \|\mathbf{L}_t\|_* + \alpha \|\mathbf{S}\|_{1,\text{off}} \quad (2) \\
& \text{subject to} \quad \Psi_t - (\mathbf{S} - \mathbf{L}_t) = \mathbf{0}, \forall t = 1, \dots, N
\end{aligned}$$

The corresponding augmented Lagrangian in scaled form can be written as:

$$\begin{aligned}
& \mathcal{L}_\rho(\mathbf{S}, \{\mathbf{L}_t\}, \{\Psi_t\}, \{\mathbf{U}_t\}) \\
& = \sum_{t=1}^N \left[-\log(\det(\Psi_t)) + \text{trace}(\hat{\Sigma}_t \Psi_t) \right] \\
& \quad + \tau \sum_{t=1}^N \|\mathbf{L}_t\|_* + \alpha \|\mathbf{S}\|_{1,\text{off}} \\
& \quad + \frac{\rho}{2} \sum_{t=1}^N [\|\Psi_t - \mathbf{S} + \mathbf{L}_t + \mathbf{U}_t\|_F^2 - \|\mathbf{U}_t\|_F^2]
\end{aligned}$$

where $\{\mathbf{U}_t\}$ are the set of scaled dual variables. Following from [10, 12], we give Algorithm 2 to solve problem (1): $\text{prox}_{\frac{\alpha}{N\rho}\|\cdot\|_{1,\text{off}}}(\mathbf{A})$ is the proximal operator for the off-diagonal ℓ_1 -norm, which performs soft-thresholding [17] of the off-diagonal elements of matrix $\mathbf{A}$ to encourage sparsity in the solutions. In addition, $\text{prox}_{\frac{\tau}{\rho}\|\cdot\|_*}(\mathbf{A})$ denotes the proximal operator for the nuclear norm, which performs soft-thresholding of the singular values of matrix $\mathbf{A}$ to encourage the solutions to be low rank. Empirical convergence of Algorithm 2 is measured by the norm of the primal and dual residuals [12]. Because problem (1) is convex, Algorithm 2 is guaranteed to converge to a global solution [12].

3. NUMERICAL STUDIES

We now compare the DYNAMO and DYNACO models to unconditional graphical model estimates from the Glasso or neighborhood selection methods in a variety of simulation studies. In the DYNAMO case, we evaluate performance of the one-flop algorithm and present results from using both neighborhood selection and the Glasso to estimate Θ after adjusting for the mean effects of dynamic latent brain states. The Glasso and neighborhood selection graphs are estimated using the *huge* package in R.

We generate data from 4 different simulation scenarios, with 5 replicate data sets for each. Scenarios 1 and 2 are generated from the DYNAMO model with $N = 500$ and $p = 50$ and $N = 500$ and $p = 200$, respectively. Similarly, scenarios 3 and 4 are generated from the DYNACO model with $N = 500$ and $p = 50$ and $N = 500$ and $p = 200$, respectively. In each simulation scenario, true sparse precision matrices are simulated to have a connected small-

Algorithm 2: ADMM algorithm for DYNACO

Input : $\mathbf{X} \in \mathbb{R}^{p \times N}$, $\alpha, \tau \geq 0$, $\rho > 0$;

Initialize: Primal variables to identity matrices and dual variables to zero matrices;

Precompute: $\hat{\Sigma}_t, \forall t = 1, \dots, N$;

while not converged do

(i) Update Ψ_t ($\forall t$ in parallel):

(a) Compute the eigenvalue decomposition:

$$\Gamma_{\Psi_t}^{(k)} \Lambda_{\Psi_t}^{(k)} (\Gamma_{\Psi_t}^{(k)})^T = \hat{\Sigma}_t - \rho(\mathbf{S}^{(k)} - \mathbf{L}_t^{(k)} - \mathbf{U}_t^{(k)})$$

(b) Update Ψ_t :

$$\Psi_t^{(k+1)} = \frac{1}{2\rho} \Gamma_{\Psi_t}^{(k)} (-\Lambda_{\Psi_t}^{(k)} + \sqrt{(\Lambda_{\Psi_t}^{(k)})^2 + 4\rho\mathbf{I}}) (\Gamma_{\Psi_t}^{(k)})^T$$

(ii) Update sparse $\mathbf{S}$:

$$\mathbf{S}^{(k+1)} = \text{prox}_{\frac{\alpha}{N\rho}\|\cdot\|_{1,\text{off}}} \left(\frac{1}{N} \sum_{t=1}^N \Psi_t^{(k+1)} + \mathbf{L}_t^{(k)} + \mathbf{U}_t^{(k)} \right)$$

(iii) Update low rank $\mathbf{L}_t$ ($\forall t$ in parallel):

$$\mathbf{L}_t^{(k+1)} = \text{prox}_{\frac{\tau}{\rho}\|\cdot\|_*} (\mathbf{S}^{(k+1)} - \Psi_t^{(k+1)} - \mathbf{U}_t^{(k)})$$

(iv) Update dual variables $\mathbf{U}_t$ ($\forall t$ in parallel):

$$\mathbf{U}_t^{(k+1)} = \mathbf{U}_t^{(k)} + (\Psi_t^{(k+1)} - \mathbf{S}^{(k+1)} + \mathbf{L}_t^{(k+1)})$$

end

return $\hat{\mathbf{S}} = \mathbf{S}^{(k+1)}$, $\hat{\mathbf{L}}_t = \mathbf{L}_t^{(k+1)}, \forall t = 1, \dots, N$

world structure, which has been widely postulated as a possible structure for connections in the brain [13]. The latent dynamics are simulated to vary over time according to some smoothly oscillating functions with various frequencies and amplitudes. Specifically, for the DYNAMO model, we create an $p \times N$ latent variable matrix $\mathbf{B}$ of rank 2 whose columns vary smoothly over time according to 2 smooth sinusoids of different frequencies along with a $p \times p$ sparse precision matrix Θ with a connected small-world topology. We then generate $\epsilon_1, \dots, \epsilon_N \stackrel{i.i.d.}{\sim} N(\mathbf{0}_p, \Theta^{-1})$ and create the simulated data matrix $\mathbf{X}_{p \times N} = \mathbf{B} + \mathbf{E}$ with $\mathbf{E}_{p \times N} = [\epsilon_1, \dots, \epsilon_N]$. For the DYNACO model, we generate $N \times p \times 2$ matrices $\mathbf{H}_t$ that represent the dynamic conditional relationships between the observed and unobserved variables, as well as a $p \times p$ static, sparse conditional precision matrix $\mathbf{S}$. Following the simulation studies in [10], which assumes the individual latent dynamics are independent from one another, we simulate the true dynamic marginal precision matrices at each time point $\Theta_t = \mathbf{S} - \mathbf{L}_t$ where $\mathbf{L}_t = \mathbf{H}_t \mathbf{H}_t^T$. The simulated data are then sampled from multivariate Gaussian distribu-

Table 1. Results from simulation scenarios 1 and 2 averaged over 5 replicates. Best performing methods are boldfaced.

	Scenario 1 (DYNAMO $p = 50, N = 500$)			Scenario 2 (DYNAMO $p = 200, N = 500$)		
Method	Recall	Precision	F-1	Recall	Precision	F-1
Glasso Oracle	0.15 (0.01)	0.16 (0.01)	0.15 (0.01)	0.15 (0.01)	0.16 (0.01)	0.15 (0.01)
NS Oracle	0.30 (0.03)	0.30 (0.03)	0.30 (0.03)	0.28 (0.01)	0.27 (0.01)	0.27 (0.01)
Glasso	0.55 (0.01)	0.21 (0.01)	0.30 (0.01)	0.68 (0.01)	0.13 (0.01)	0.21 (0.01)
NS	0.04 (0.04)	0.19 (0.17)	0.07 (0.07)	0.05 (0.03)	0.47 (0.12)	0.08 (0.05)
Dynamo Glasso Oracle	0.71 (0.02)	0.71 (0.02)	0.71 (0.02)	0.50 (0.03)	0.50 (0.03)	0.50 (0.03)
Dynamo NS Oracle	0.84 (0.02)	0.84 (0.02)	0.84 (0.02)	0.81 (0.03)	0.82 (0.01)	0.82 (0.02)
Dynamo Glasso	0.85 (0.02)	0.62 (0.02)	0.72 (0.01)	0.87 (0.01)	0.45 (0.01)	0.60 (0.01)
Dynamo NS	0.88 (0.02)	0.80 (0.03)	0.84 (0.02)	0.16 (0.02)	0.88 (0.02)	0.27 (0.02)
Dynaco Oracle	0.28 (0.18)	0.28 (0.18)	0.28 (0.18)	0.05 (0.01)	0.055 (0.01)	0.0525 (0.01)
Dynaco Stability	0.29 (0.01)	0.34 (0.01)	0.32 (0.01)	0.18 (0.01)	0.22 (0.02)	0.19 (0.01)

Table 2. Results from simulation scenarios 3 and 4 averaged over 5 replicates. Best performing methods are boldfaced.

	Scenario 3 (DYNACO $p = 50, N = 500$)			Scenario 4 (DYNACO $p = 200, N = 500$)		
Method	Recall	Precision	F-1	Recall	Precision	F-1
Glasso Oracle	0.46 (0.04)	0.48 (0.03)	0.47 (0.04)	0.46 (0.03)	0.43(0.08)	0.44 (0.03)
NS Oracle	0.52 (0.04)	0.53 (0.04)	0.53 (0.03)	0.46 (0.03)	0.44 (0.01)	0.45 (0.01)
Glasso	0.76 (0.06)	0.34 (0.02)	0.47 (0.02)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
NS	0.67 (0.04)	0.45 (0.01)	0.54 (0.02)	0.66 (0.01)	0.23 (0.02)	0.34 (0.02)
Dynamo Glasso Oracle	0.47 (0.03)	0.47 (0.03)	0.47 (0.03)	0.44 (0.04)	0.44 (0.05)	0.44 (0.04)
Dynamo NS Oracle	0.52 (0.02)	0.52 (0.02)	0.52 (0.02)	0.45 (0.03)	0.47 (0.05)	0.46 (0.04)
Dynamo Glasso	0.77 (0.06)	0.34 (0.01)	0.47 (0.01)	0.67 (0.01)	0.17 (0.02)	0.27 (0.03)
Dynamo NS	0.73 (0.03)	0.42 (0.01)	0.54 (0.01)	0.68 (0.01)	0.22 (0.01)	0.33 (0.02)
Dynaco Oracle	0.58 (0.03)	0.59 (0.03)	0.58 (0.03)	0.33 (0.04)	0.39 (0.05)	0.36 (0.04)
Dynaco Stability	0.53 (0.02)	0.69 (0.02)	0.60 (0.02)	0.30 (0.02)	0.35 (0.03)	0.32 (0.02)

tions $\mathbf{x}_t \sim N(\mathbf{0}_p, \Theta_t^{-1}), \forall t = 1, \dots, N$ and arranged into a data matrix $\mathbf{X}_{p \times N} = [\mathbf{x}_1, \dots, \mathbf{x}_N]$.

In each of the simulation scenarios, we present results from unconditional graphical models (i.e. Glasso, neighborhood selection) and the DYNAMO and DYNACO models, using both oracle model selection based on true graph sparsity and latent variable rank (where applicable) as well as data-driven model selection methods. The extended Bayes information criteria (EBIC) and rotational information criteria (RIC) are used for data-driven model selection for the Glasso, neighborhood selection, and DYNAMO models, while stability selection with block bootstrapping is used for the DYNACO model. The average recall, precision, and F-1 scores from the model estimates for each of the simulation scenarios, as well as the standard error for the aforementioned metrics across 5 replicate data sets for each scenario, are shown in Table 1 and 2. In scenarios 1 and 2, the estimates from the DYNAMO model achieve much higher F-1 scores than the estimates from both the unconditional graphical models and the DYNACO model. The DYNACO model performs much better than the unconditional graphs in scenario 3, while DY-

NAMO still attains decent F-1 scores in both scenario 3 and 4. Such simulation results show that the DYNAMO model can perform relatively well in the presence of either dynamic latent mean structure or dynamic latent brain states that directly influence the covariance structures of observed neuronal activity and that the DYNAMO model is more robust to different situations and to perturbations in data.

The true adjacency matrices and the corresponding estimated adjacency matrices from the best performing models in scenario 1 are shown in Fig. 1. It is clear that both DYNAMO and DYNACO outperform the unconditional graphical models in terms of recovery of the true graph structure. In particular, the graph estimated by the DYNAMO model most closely captures the true graph structure while the DYNACO model still gives a relatively more accurate estimate of the true graph structure compared to the unconditional graphical models. On the other hand, the unconditional neighborhood selection graph estimate fails to detect the small world structure of the simulated functional connectivity graph. In the neuroscience context, such unconditional graphical models could fail to detect highly interconnected groups of neurons,

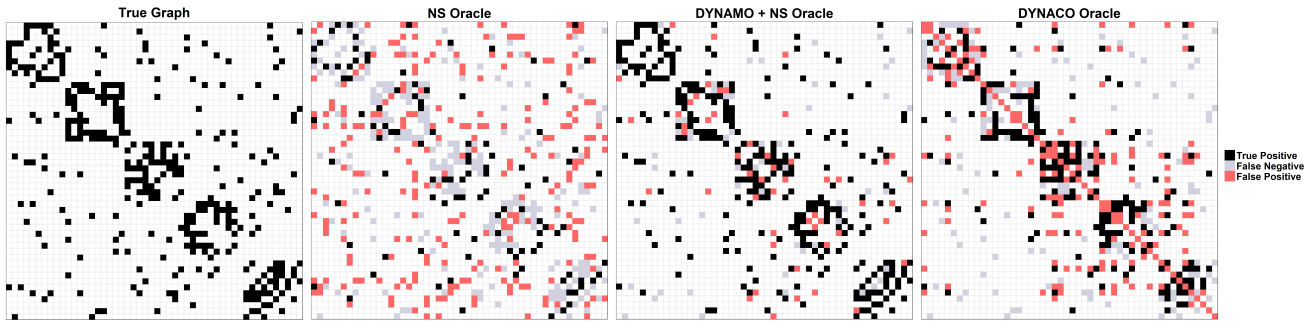


Fig. 1. Simulation results from simulation scenario 1. From the leftmost panel to the rightmost panel, we have the true adjacency matrix according to Θ , the best estimated unconditional graphs (from neighborhood selection with oracle model selection), the best graph estimated by DYNAMO (with neighborhood selection and oracle model selection), and the best graph estimated by DYNACO (with oracle model selection). We can see that both DYNAMO and DYNACO outperform the unconditional graphical models as DYNAMO and DYNACO achieve higher true positive rates with relatively low false positive and false negative rates.

which may otherwise shed light on organizations of neural circuits.

4. DISCUSSION

In this work, we have presented two potential frameworks and estimation procedures for dynamic latent variable conditional graphical models, the DYNAMO and DYNACO models, for estimating a static graph for functional connectivity from neuronal activity data while accounting for the potential effects of unobserved neurons, modeled as dynamic latent brain states. The DYNAMO model adjusts for dynamic latent brain states in terms of smoothly changing mean structures while the DYNACO model accounts for dynamic latent states which have a time-varying influence on the covariance structure of observed neurons. We demonstrate the effectiveness of our proposed models through a variety of simulation studies. In particular, our studies reveal that the DYNAMO model could potentially provide more accurate estimates of functional connectivity graph of observed neurons after adjusting for dynamic latent brain states compared to traditional, unconditional graphical models. Moreover, the DYNAMO model is easy to implement and produce estimates that appear to be more robust to different generative models, especially in higher dimensional cases.

There are many potential methodological extensions to our work that can be explored further in future work. For example, neuronal activity data are often collected along with auxiliary information, such as visual and auditory stimuli or body movements. Incorporating such auxiliary information in corresponding models could potentially provide more accurate estimation of the underlying dynamic latent brain states. Moreover, there are also further studies that can be built on the work done in this paper from a neuroscientific

perspective. Investigation can be done to see if the estimated functional connectivity graphs from dynamic latent variable conditional graphical models on real-world neuronal activity recording data provide neuroscientifically meaningful findings compared to unconditional graphical models. Further work can also be done to see if different variants of graphical model methods could be used in the neuroscience field, and to possibly develop new classes of graphical models that will be more useful specifically to neuroscientific problems. In conclusion, the application of graphical models in neuroscience has the potential to help us gain insight into how neurons in the brain communicate with one another, and our work is a step towards better understanding the types of graphical models that should be used in the field.

5. REFERENCES

- [1] B. Biswal, F. Zerrin Yetkin, V. M. Haughton, and J. S. Hyde. "Functional connectivity in the motor cortex of resting human brain using echoplanar MRI." *Magnetic Resonance in Medicine*, vol. 34, no. 4, pp. 537-541, 1996.
- [2] B.B. Averback, P.E. Latham, and A. Pouget. "Neural correlations, population coding and computation." *Nature reviews neuroscience*, vol. 7, no.5, pp.358-366, 2006.
- [3] D. Yatsenko, K. Josi, A.S. Ecker, E. Froudarakis, R.J. Cotton, and A.S. Tolias. "Improved Estimation and Interpretation of Correlations in Neural Circuits." *PLOS Computational Biology*, vol. 11, no. 3, p. e1004083, 2015.
- [4] G. Vinci, V. Ventura, M. A. Smith, and R. E. Kass. "Adjusted regularization in latent graphical models: Application to multiple-neuron spike count data." *The Annals of Applied Statistics*, vol. 12, no. 2, pp. 1068-1095, 2018.
- [5] G.I. Allen and Z. Liu. "A Local Poisson Graphical Model for Inferring Networks From Sequencing Data." *IEEE Transactions on Nanobioscience*, vol. 12, no. 3, 2013
- [6] S. Ryali, T. Chen, K. Supekar, and V. Menon. "Estimation of functional connectivity in fMRI data using stability selection-based sparse partial correlation with elastic net penalty." *NeuroImage*, vol. 59, no. 4, pp.3852-3861, 2012.
- [7] M.J. McGinley, M. Vinck, J. Reimer, R. Batista-Brito, E. Zagha, C.R. Cadwell, A.S. Tolias, J.A. Cardin, and D.A. McCormick. "Waking state: rapid variations modulate neural and behavioral responses." *Neuron*, vol. 87, no. 6, pp.1143-1161, 2015.
- [8] V. Chandrasekaran, P. A. Parrilo, and A. S. Willsky. "Latent Variable Graphical Model Selection Via Convex Optimization." *The Annals of Statistics*, vol. 40, no. 4, pp. 1935-1967, 2012.
- [9] C. Sutton, A. McCallum, and K. Rohanimanesh. "Dynamic conditional random fields: Factorized probabilistic models for labeling and segmenting sequence data." *Journal of Machine Learning Research*, vol. 8, pp. 693-723, March 2007.
- [10] F. Tomasi, V. Tozzo, S. Salzo, and A. Verri. "Latent variable time-varying network inference." *arXiv preprint arXiv:1802.03987*, 2018.
- [11] A. Jalali and S. Sanghavi. "Learning the dependence graph of time series with latent factors." *arXiv preprint arXiv:1106.1887*, 2011.
- [12] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. "Distributed optimization and statistical learning via the alternating direction method of multipliers." *Foundations and Trends in Machine learning*, vol. 3, no. 1, pp. 1-122, 2011.
- [13] D.S. Bassett and E.T. Bullmore. "Small-world brain networks revisited." *The Neuroscientist*, vol. 23, no. 5 pp. 499-516, 2017.
- [14] M. Yuan and Y. Lin. "Model selection and estimation in the Gaussian graphical model." *Biometrika*, vol. 94, no. 1, pp. 1935, March 2007. <https://doi.org/10.1093/biomet/asm018>
- [15] J. A. Rice and B. W. Silverman. "Estimating the mean and covariance structure nonparametrically when the data are curves." *Journal of the Royal Statistical Society, Series B*, pp 233243, 1991.
- [16] S. Zhou, J. Lafferty, and L. Wasserman. "Time varying undirected graphs." *Machine Learning*, vol. 80, no. 2, pp. 295-319, 2010.
- [17] N Parikh and S. Boyd. "Proximal Algorithms." *Foundations and Trends in Optimization*, vol. 1, no. 3, pp. 127-239, 2014.
- [18] N. Meinshausen and P. Bhlmann. "High-dimensional graphs and variable selection with the Lasso." *Annals of Statistics*, vol. 34, no. 3, pp. 1436-1462, 2006.
- [19] Yin, Jianxin and Li, Hongzhe. "A sparse conditional Gaussian graphical model for analysis of genetical genomics data." *Ann. Appl. Stat.*, vol. 5, no. 4, pp. 2630-2650, 2011.