# UC Merced

## Proceedings of the Annual Meeting of the Cognitive Science Society

### Title

Discovering Low-Dimensional Causal Pathways between Multiple Interacting Neuronal Populations

### Permalink

https://escholarship.org/uc/item/3r68f33q

### Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

### Authors

Sigalas, Evangelos
Vo, Thanh Vinh
Leong, Tze Yun
et al.

### Publication Date

2023

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

# Discovering Low-Dimensional Causal Pathways between Multiple Interacting Neuronal Populations

**Evangelos Sigalas**[1] **(esigalas@u.nus.edu), Thanh Vinh Vo**[1] **(votv@comp.nus.edu.sg),**
**Tze-Yun Leong**[1] **(leongty@nus.edu.sg), Camilo Libedinsky**[2] **(camilo@nus.edu.sg)**
[1]School of Computing, National University of Singapore
[2]Department of Psychology, National University of Singapore

## Abstract

Understanding the nature of neural activity and computations in the brain will help us build better decision-making models to facilitate human-AI collaboration. Recording the neural activity of multiple and large neural populations in the brain is becoming widely available with modern recording techniques. It still remains a challenge, however, to understand how distinct and anatomically different neural populations interact with each other to control behaviour. We propose a new method to discover causal interactions between neural populations based on recurrent switching dynamical systems. We introduce an extended dynamics model that incorporates the current time-step when calculating the latent state variables. We also introduce an acyclicity constraint in learning the parameters of the model. These mechanisms enable rich causal interactions between neural populations to be identified from the learned model. Our model outperforms previous work on discovering interactions between neural populations in simulated datasets, without sacrificing the prediction performance of firing rates. We also apply our method on real neural recordings from two Macaque monkey brains performing a behavioral task, and show that the proposed method is able to detect causal interactions between brain regions related to the different time windows of the task.

**Keywords:** causal discovery; neural population interactions; recurrent switching dynamical systems

## Introduction

Recent advances in neural recording techniques allow us to monitor thousands of neurons in multiple brain regions at the same time with single-cell resolution (Jun et al., 2017). The interactions and flow of information between these neural populations enable the sensory, motor, and cognitive functions of the brain (Kohn et al., 2020). A key challenge of utilizing multi-region recordings to understand how these populations communicate still remains.

The neural activity across a population of neurons can be depicted as a point in a coordinate system where each axis corresponds to the activity of one neuron. Since the number of neurons we are recording from is in the hundreds or thousands, this coordinate system defines a high-dimensional space which is known as the neural state (Remington, Egger, Narain, Wang, & Jazayeri, 2018). Analyzing the dynamics between or within brain regions can be challenging in such a space. Cunningham and Yu (2014) showed that neural activities within a population can be correlated and reflect their dynamics in a low-dimensional latent space, suggesting that dimensionality reduction methods can be used to study the local dynamics within a brain region. Gallego, Perich, Miller, and Solla (2017) utilized such a method to analyze how motor cortex controls movement in the brain of a Macaque monkey.

Dimensionality reduction techniques have also been used to study the interactions between brain regions. J. D. Semedo, Zandvakili, Machens, Yu, and Kohn (2019) applied the reduced-rank regression method to reveal that the communication between two visual cortex areas V1 and V2 occurs in a lower-dimension latent space. Rodu, Klein, Brincat, Miller, and Kass (2018) and J. Semedo, Zandvakili, Kohn, Machens, and Yu (2014) used a framework where latent variables summarize the population activity within each area or shared across areas. The interactions of these latent variables through time were used to describe the interactions between brain areas.

Glaser, Whiteway, Cunningham, Paninski, and Linderman (2020) proposed an extended version of recurrent switching linear dynamical system (rSLDS) where they enforced population-specific latent variables when analyzing multiple-population interactions. They also introduced a set of discrete states each with its own unique dynamics. Furthermore, the probabilities in the transition model between these discrete states were governed by the continuous latent states of the neural populations. This parametrization increases the interpretability of the final model.

None of the aforementioned works considers causality when they study neural population interactions. Here, causality refers to the direct influence that one neural population exerts over another neural population. This can be interpreted as the passing or exchanging of information between them over time. Discovering the causality between neural populations especially in low-dimensional spaces might be the key to understanding the link between neural activity and brain behavior (Jazayeri & Afraz, 2017), (Saxena & Cunningham, 2019). We propose an extension of the rSLDS similar to Glaser et al. (2020) to identify causal interactions between the population-specific latent variables. Driven by Cohen and Kohn (2011), who showed that timescale window of calculated spikes of the neural activity affects the measurements of interactions between neurons, we extend the model to include the current time step in the calculation of latent variables. This introduces cycles in the directed acyclic graph (DAG) of the model. Learning the DAG of a model is an NP-Hard problem (Chickering, 1996), mainly because it is very difficult to efficiently enforce the combinatorial acyclicity constraint.



In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.* ©2023 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

Zheng, Aragam, Ravikumar, and Xing (2018) proposed a new approach for score-based learning of DAGs by converting the traditional combinatorial optimization problem into a continuous program. We adapt their suggested smooth function in our problem to impose the acyclicity constraint in our model and learn the underlying DAG.

We fit our model with a variational expectation-maximization algorithm that leverages the standard smooth optimization scheme L-BFGS (Mikosch, Wright Stephen, & Jorge, 2006). We show that our model is able to detect causal interactions between multiple neural populations in simulated and real data recordings. Discovering such interactions can help us understand how brain regions communicate while the brain performs a behavioral task and what computations take place in the brain.

## Background

In an example of two interacting neural populations, we wish to study the interactions between the *source* population and the *target* population. In particular, we want to understand how the neural activity of the source population relates to the activity of the target population. One way to determine that relationship is to measure how well the neural activity of the source population predicts the activity of the target population. Given the high-dimensionality of the neural activity of both populations, however, it is a challenge to avoid overfitting. There are various types of models that use dimensionality reduction techniques with latent variables to overcome this overfitting problem for multiple neural populations.

The linear dynamical system model (LDS) uses linear dynamics to model a lower-dimensional latent space representation of the multi-dimensional time series. An extended version of LDS is the switching linear dynamical system (SLDS) which includes a set of discrete states, each with its own linear dynamics. This allows the model to switch between different linear dynamics over time, thus modeling nonlinear dynamics in a locally linear manner. A Markov transition matrix specifies the probabilities of switching between the discrete states, similar to a hidden Markov model. A further extension of SLDS is the rSLDS model, where the discrete state transitions are dependent on the the low-dimensional latent space that represents the time series (Linderman et al., 2017).

More specifically, in rSLDS, a vector $y_t$ represents the neural activity of $N$ neurons at time $t$. $x_t \in \mathbb{R}^D$ denotes the continuous latent state space with a dimension much lower compared to that of the observed neural activity ($D \ll N$). A Generalized Linear Model (GLM) models the observed neural activity as $E[y_t] = f(Cx_t + d)$, where $C \in \mathbb{R}^{N\times D}$ and $d \in \mathbb{R}^N$ are defining a linear mapping between the latent states and the neural activity. The function $f$ is applied to map the output to the desired space. Typically, for spike counts, the softplus function $f(x) = \log(1+e^x)$ is used along with a Poisson distribution. A Markov process with conditional dependencies on the continuous latent states models the discrete states $z_t \in \{1,\dots,K\}$, where $K$ is the total number of discrete states which is predetermined. Note that each discrete state $K$ has its own unique set of linear dynamics. Furthermore, the discrete state at time $t$ is sampled from a Categorical Distribution,

$$z_t \sim \mathsf{Cat}(\pi_t), \quad \pi_t = \mathsf{softmax}(R_{z_{t-1}} x_{t-1} + r_{z_{t-1}}), \qquad (1)$$

where $R_{z_{t-1}} \in \mathbb{R}^{K\times D}$ and $r_{z_{t-1}} \in \mathbb{R}^K$ define a GLM that determines the influence of the continuous latent state on the discrete state transitions.

Glaser et al. (2020) proposed an extended version of rSLDS called multi-population srSLDS (mp-srSLDS) where they constrained the neural activity of each brain region to have its own continuous latent states. These are defined as:

$$x_t^{(j)} = A_{j\leftarrow j}^{(z_t)} x_{t-1}^{(j)} + \sum_{i\neq j} A_{j\leftarrow i}^{(z_t)} x_{t-1}^{(i)} + b_j^{(z_t)} + \varepsilon_t^{(j)}, \qquad (2)$$

where $\varepsilon_t = (\varepsilon_t^{(1)},\dots,\varepsilon_t^{(J)}) \sim \mathsf{N}(0, Q^{(z_t)})$ and $b_j^{(z_t)}$ is the bias. The matrices $A_{j\leftarrow j}^{(z_t)} x_{t-1}^{(j)}$ and $A_{j\leftarrow i}^{(z_t)} x_{t-1}^{(i)}$ capture the linear dynamics within and between brain regions, respectively. Note that each continuous latent state $x_t$ at time $t$ is associated with a different brain region $j$. It also depends on the previous time step continuous latent state $x_{t-1}$. Finally they extended the transition model to distinguish between self-transitions and transitions to other states. By dividing the neural population activity into groups based on brain region and constraining each region to have its own latent state, the interactions between those regions can be easily interpreted. Matrix $A^{(k)}$ captures these interactions for each discrete state. However, as mentioned in their work, these interactions should not be interpreted as inferred causal interactions within the brain.

## Method

In this work, we extend the mp-srSLDS model to identify causal relationships between neural populations based on their neural activity. We model the observed spiking activity of different neural populations in a continuous low-dimensional latent state space similar to the rSLDS we described. If the size of the time window of spiking activity is large enough, it can include causal interactions that are missed. To address this issue, we propose an extension of the dynamics in the continuous space that includes the current time step of the latent state. In that continuous latent space, we want to detect causal interactions between the neural populations and discover the underlying DAG that depicts the causal relations. The extended dynamics of our model takes into consideration the cycles between neural populations that occur in the brain. This however, hinders our ability to learn the underlying DAG. Hence, we introduce an acyclicity constraint that helps us solve this problem and discover causal interactions between neural populations.

**Extending the dynamics model** Typically, when calculating spike counts, we select a time bin size in the range of 20-100$ms$ based on various experimental conditions and the method we want to apply. This time bin size is an important factor that can affect the measurement of interactions between

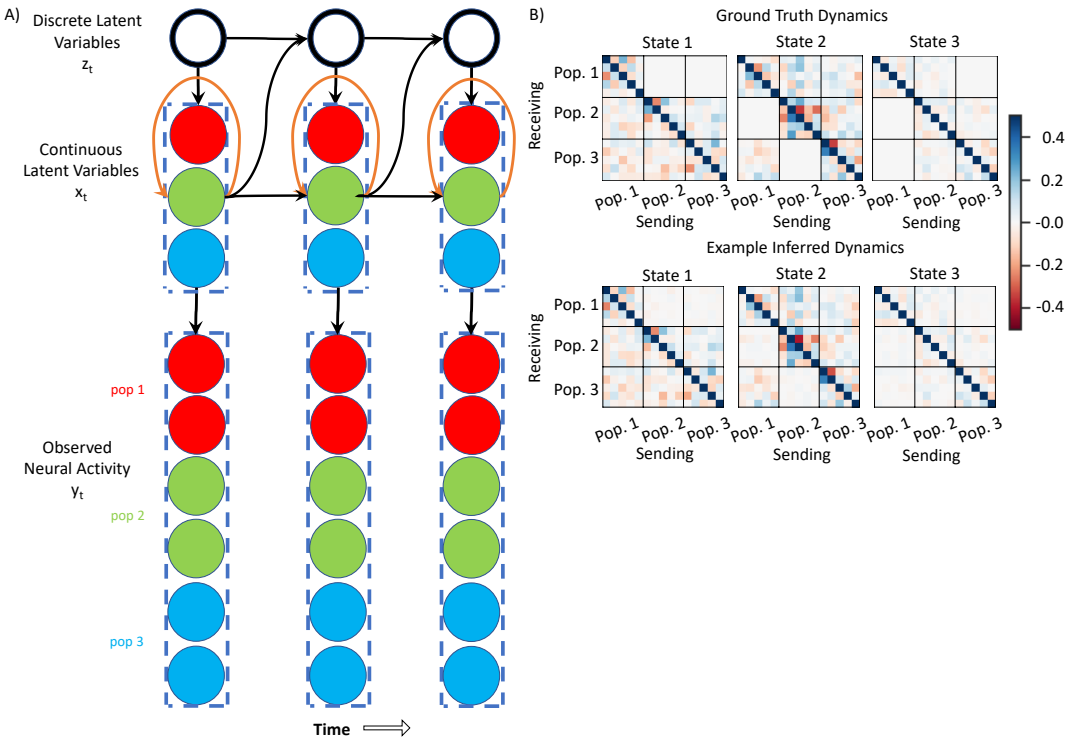


Figure 1: **A)** Graphical Model of our causal model, an extension of mp-srSLDS for multiple population interactions. Each color represents a different neural population. **B)** (top) Ground Truth Dynamics Matrices $A$ for each of the three states of simulated datasets. (bottom) Example of inferred Dynamics Matrices $A$.

populations of neurons. Small time bin sizes can reveal weak correlations between neural populations while big time bin sizes can reveal strong ones (Cohen & Kohn, 2011). This means that information in each current time bin of neural activities is important to accurately measure neural population interactions. As a result, we extend the multi-population srSLDS dynamics model to include the current time step of observed neural activity in the calculation of the continuous latent state variables (Fig1.A). Thus, the new temporal dynamics in our model are defined as:

$$x_t^{(j)} = A_{j\leftarrow j}^{(z_t)} x_{t-1}^{(j)} + A_{j\leftarrow j}^{(z_t)} x_t^{(j)} + \sum_{i\neq j} A_{j\leftarrow i}^{(z_t)} x_{t-1}^{(i)} + A_{j\leftarrow i}^{(z_t)} x_t^{(i)} + b_j^{(z_t)} + \varepsilon_t^{(j)}, \tag{3}$$

where $x_t \in \mathbb{R}^D$, $A \in \mathbb{R}^{K\times D\times D}$, $D$ is the size of the low-dimensional space and $K$ is the number of discrete states. Note that each discrete state $z_t$ has a different dynamics matrix $A^{(k)}$ with its own dynamics for the interactions between neural populations.

The two new introduced terms, $A_{j\leftarrow j}^{(z_t)} x_t^{(j)}$ and $A_{j\leftarrow i}^{(z_t)} x_t^{(i)}$, capture the linear dynamics of the current time bin within and between brain regions, respectively. Note that $A_{j\leftarrow j}^{(k)}$ and $A_{j\leftarrow i}^{(k)}$ form the blocks of the full dynamics matrix $A^{(k)}$. Now the dynamics matrix $A^{(k)}$ can capture richer interactions between neural populations, especially in cases where the time bin size is big ($\gg 25ms$).

**Acyclicity Constraint** Naturally this new definition of the dynamics model introduces cycles between the continuous latent states, thus creating a directed cyclic graph. This imposes difficulties in learning the underlying DAG. Our second extension of the model inspired by the work of Zheng et al. (2018) where they proposed a new definition of acyclicity which is expressed through a smooth function constraint $h(A) = 0$. This function ideally should satisfy the following criteria:

1. $h(A) = 0$ if and only if $A$ is acyclic;
2. The values of $h$ quantify the "DAG-ness" of the graph;
3. $h$ is smooth;
4. $h$ and its derivatives are easy to compute.

"DAG-ness" is a quantification of how much the matrix $A$ violates acyclicity. A function that satisfies all the above is:

$$h(A) = tr(e^{A\circ A}) - d, \tag{4}$$

where $\circ$ is the element-wise product. The parameter $d$ is a constant that is equal to the dimension of matrix $A$. When there is no cycle in the underlying graph then $h(A) = 0$. By incorporating this function in our fitting method while learning the parameters of our model, we can learn a matrix $A$ that

depicts the causal interactions between the different neural populations in the lower latent space $x$. We also introduce a regularizer parameter $\lambda h(A)$ that will control the level of acyclicity in our model. In our experiments we try different values of $\lambda$ and evaluate their performance, but in general the choice of $\lambda$ should depend on the type of observed data we train the model with.

**Model Fitting** We fit our model using a variational expectation-maximization algorithm (Zoltowski, Pillow, & Linderman, 2020). We use a structured mean-field approximation (Jordan, Ghahramani, Jaakkola, & Saul, 1999) to approximate the posterior distribution on latent variables given the model parameters and observed data similar to Glaser et al. (2020). Finally, we introduce our acyclicity constraint $\lambda h(A)$ when maximizing the expected log probability.

$$L(\Theta) = \mathbb{E}\Big[\log p(z_{1:T}, \{x^{(j)}_{1:T}, y^{(j)}_{1:T}\}_{j=1}^{J})\Big] - \sum_{k=1}^{K} \lambda h(A^{(k)}), \quad (5)$$

where $\Theta$ is the set of all model parameters. Since in our case each discrete state has its own dynamics matrix $A$ we apply the constraint separately in each one of the $A^{(k)}$ matrices and calculate their summation in the final term of the equation (5).

## Results

In our experiments we test if our model is able to discover causal interactions between neural populations. We quantify the model performance in two ways: First, in simulated datasets of neural spike activity, where ground truth is known, we evaluate the accuracy of the inferred dynamics matrix A and the inferred firing rates for different model versions. Second, in real data recordings of neural spike activity, where ground truth is not available, we do a qualitative interpretation of the results.

### Simulated Data

We performed experiments on two types of simulated datasets generated by two different generative models. The first generative model did not contain cycles like the original mp-srSLDS model while the second one contained cycles as we suggested in our extension of the model. We used $J = 3$ neural populations with $D = 5$ size for the latent space, $N = 75$ neurons per population and $K = 3$ discrete states of dynamics each with its own dynamics matrix $A$ as shown in Fig1.B. We experimented with different number of time bins for the time series of neural activity which we set to $T \in \{1000, 2000, 3000, 4000, 5000\}$. Finally, we also tried various bin sizes $\{25, 50, 100\}$ each with an average of 20 spikes per bin. The aim of the experiments is to show the consistency of the model as we increase the size of available data.

We trained three versions of our extended model with different values of our proposed regularizer $\lambda \in \{0.5, 2, 20\}$ to explore different levels of acyclicity. We also trained a version of mp-srSLDS on the same datasets for comparison. Additionally, we trained a version of srSLDS model which does not constrain the neural activity of each brain region to have its own continuous latent states.

To evaluate the performance of our models, first, we measured the Mean Square Error (MSE) between the ground truth and the inferred dynamics matrix $A$ (Fig2.A). Secondly, we measured the predictive accuracy of each model by calculating the MSE between the firing rate of the simulated neural activities and the predicted firing rate of the model (Fig2.B). More specifically, we created new time series of neural activities sampled from the same Poisson distribution as the neural activities used for training. Then, we tested the accuracy of the model on the prediction of the last 1000 time bins of each dataset.

We show that all three versions of our extended models that include the acyclicity constraint outperform both the mp-srSLDS and srSLDS models when inferring the dynamics matrix $A$ (Fig2.A). This is even more clear on the simulated datasets with cycles. Additionally, when we compare the predictive accuracy on firing rates of simulated neural activity we see that the proposed models do not sacrifice any prediction performance and are marginally better for datasets from both generative models, escpecially in cases with high number of time bins. These results indicate that the inclusion of the acyclicity constraint and the extended dynamics are improving the overall performance of our model, allowing it to better detect causal interactions between neural populations. Unfortunately, our model is not performing as well as mp-srSLDS when predicting the discrete state $z_t$ of the dataset. That means that the model cannot detect more than one discrete state in the data with high accuracy. This is mostly because the transition model does not include any similar extension as our proposed dynamics that include the current time step in the calculation of latent variables $x_t$. Future work on extending the transition model can help improve the performance of our model when predicting the discrete states $z_t$.

### Real Data

Next we applied our method on a dataset that consists of spiking activity recorded from the Dorsolateral Prefrontal Cortex (DLPFC) and Front Eye Fields (FEF) areas of a Macaque monkey brain while the monkey performed a delayed saccade task (Parthasarathy et al., 2017). In this behavioral task the monkey was placed in front of a monitor displaying a grid. The trial began when the monkey's gaze was fixated on the fixation point at the center of the grid. The goal was to remember the location of a target stimulus and make an eye movement at the end of the trial to the correct location to receive juice as reward. In Fig3.A, we present the various time windows of the task and their corresponding duration. Particularly, a target (purple square) was presented for 300ms in a random place inside the grid in the monitor, called the Target Stimulus Display (TSD). TSD was followed by a 1s delay (D1). A distractor (orange square) was then presented for 300ms (termed the Distractor Stimulus Display (DSD)), followed by a another 1s delay (D2). After D2, the fixation point disappeared from the monitor which was the go cue for

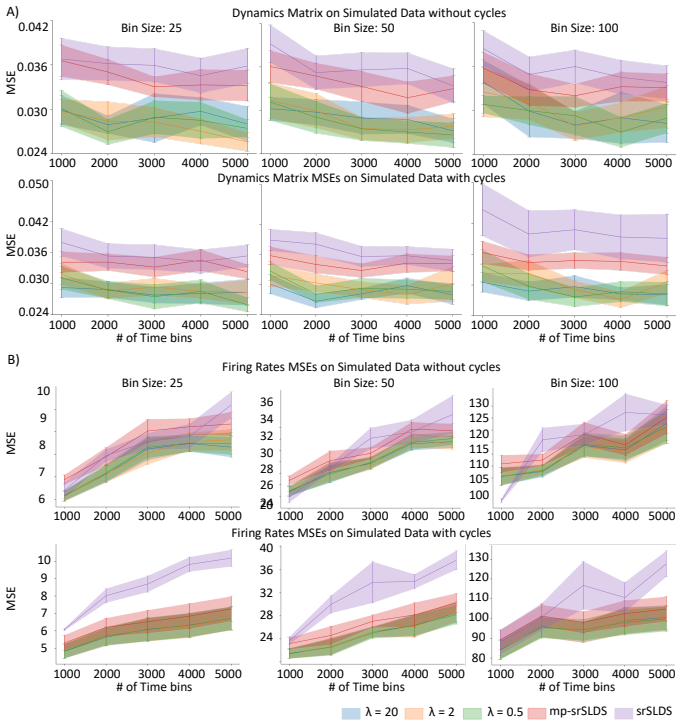


Figure 2: Performance of Models. Each color represents a different version of a trained model. **A)** The MSEs and variances of inferring Dynamics Matrix $A$ for the models trained on datasets with and without cycles in the $y$–axis. In the $x$–axis are the different number of time bins. **B)** Same as A) for the MSEs and variances of the Firing Rates.

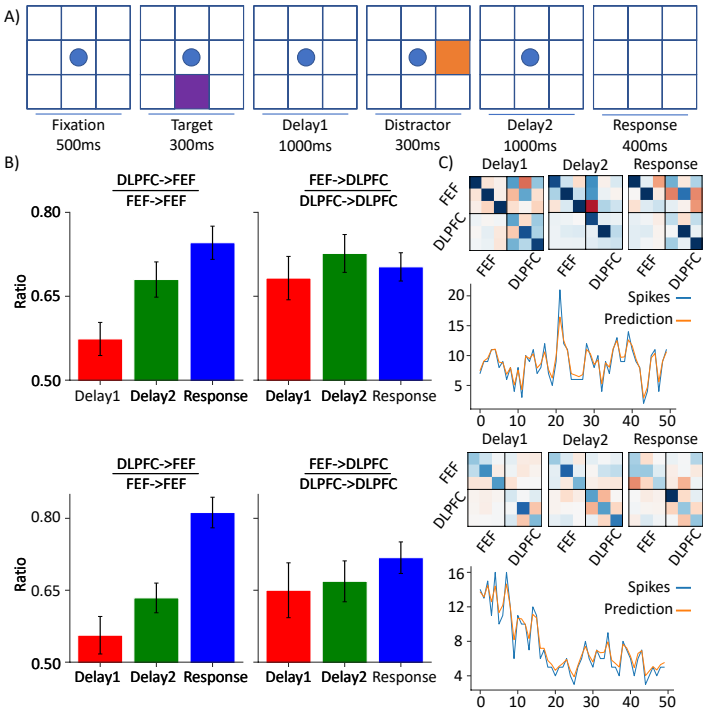


Figure 3: **A)** The time windows of the behavioral task and their duration in ms. **B)** Results of the learned model parameters for each monkey. (Top) The ratio of magnitude of external to internal interactions between the two brain regions for each time window of the trial from first monkey. (Bottom) Same for the second monkey **C)** (Top) Learned Dynamics Matrices for each time window and an example plot of the predicted neural activity. (Bottom) Same for the second monkey

the monkey to report the location of the target using an eye movement (Response). If the movement was in the correct location, it was rewarded with a drop of juice.

The task was performed by two monkeys. Each dataset consisted of 525 and 365 trials for each monkey. Spikes from 127 FEF and 173 DLPFC neurons were recorded in total from both monkeys. The spikes were counted in timebins of size 100ms. In our experiments we only modeled the neural activity of D1, D2, and Response time windows.

First we cross-validated our models with different $D$ latent space sizes to decide which one accurately predicts the firing rate of real recordings. We ended up picking $D = 3$. Then, we used that model to fit the fully observed data recordings from each of the three time windows. We assumed that each time window of the task contained only one discrete state related to it, thus, we chose $K = 1$ as the latent discrete state size. We picked a low value for $\lambda = 0.5$ to allow cycles between the neural populations, since we know that the brain has a high number of recurrent connections. Our goal was to learn the dynamics matrix $A$ that contains the causal interactions between the two brain regions for the different time windows of the behavioral task. In Fig3.C we show an example of inferred dynamics matrices $A$ and the predictive firing rates of the model for each monkey. We can see that our model is able to fit the firing rates of the data with high accuracy.

To quantify the causal interactions represented in matrix $A$ we measured the ratio of external to internal interactions for each brain area across all the three time windows. Our results in Fig3.B show an increasing trend of causal interactions from DLPFC to FEF as we move from D1 to D2 and Response, compared to the interactions from FEF to DLPFC where such a trend is not present. This can be explained by the transferring of information related to the location of the target from DLPFC to the FEF, since DLPFC contains the memory information of the correct target location and FEF is responsible for controlling the movement of the eyes of the monkey. The causal interactions peak during the response time window where the monkey has to make an eye movement to the correct target location. These results demonstrate the ability of our model to detect the causality between the two neural populations.

## Subjects and surgical procedures

We used two adult male macaques (Macaca fascicularis) in this experiment; Animal J (age 4) and Animal W (age 12). All animal procedures were approved by, and conducted in compliance with, the standard of the Agri-Food and Veterinary Authority of Singapore and Singapore Health Services Institutional Animal Care and Use Committee (SingHealth IACUC #2012/SHS/757), and the National University of Singapore Institutional Animal Care and Use Committee (NUS IACUC #R18-0295). Procedures also conformed to the recommendations described in Guidelines for the Care and Use of Mammals in Neuroscience and Behavioral Research (National Academies Press, 2003). Each animal was first implanted with a titanium head-post (Crist Instruments, MD, USA) before arrays of intracortical microelectrodes (Micro-Probes, MD, USA) were implanted in multiple regions of the left frontal cortex. In Animal W, one array of 32 electrodes was placed over the FEF, while in Animal J, 2 arrays of 32 electrodes were placed over the dorsal and ventral aspect of the FEF, respectively. The arrays consisted of platinum-iridium wires with either 200- or 400$\mu$m separation, 1-5.5 mm long and with 0.5M$\Omega$ of impedance, arranged in $4 \times 4$ or $8 \times 4$ grids.

## Discussion

Understanding how computations and causal interactions occur in different brain regions regulating different behavioral functions can help us build more effective decision-making models that support human-AI collaboration. In this work, we proposed an extended version of recurrent switching dynamical systems model to infer causal interactions between neural populations. We demonstrated increased efficacy of our model compared to similar previous work on simulated datasets. We achieved this by introducing an extended dynamics model of the latent state variables along with an acyclicity constraint to learn the parameters of the model. Furthermore, we showed that our model is able to detect causal interactions of real data recordings from the Macaque monkey brains in an interpretable way. A limitation of our model is that it cannot match the performance of previous works when detecting the discrete state of the observed data. However, we believe that future work to extend the current discrete state model can improve its performance.

## Acknowledgments

This research was partially supported by an Academic Research Grant No. MOE-T2EP30121-0010 and a Research Scholarship from the Ministry of Education in Singapore. This research was also partially supported by the National Research Foundation Singapore and DSO National Laboratories under the AI Singapore Programme (AISG Award No: AISG2-RP-2020-016).

## References

Chickering, D. M. (1996). Learning bayesian networks is np-complete. In *Learning from data* (pp. 121–130). Springer.

Cohen, M. R., & Kohn, A. (2011). Measuring and interpreting neuronal correlations. *Nature Neuroscience*, *14*(7), 811–819.

Cunningham, J. P., & Yu, B. M. (2014). *Dimensionality reduction for large-scale neural recordings* (Vol. 17) (No. 11). doi: 10.1038/nn.3776

Gallego, J. A., Perich, M. G., Miller, L. E., & Solla, S. A. (2017). Neural manifolds for the control of movement. *Neuron*, *94*(5), 978–984.

Glaser, J., Whiteway, M., Cunningham, J. P., Paninski, L., & Linderman, S. (2020). Recurrent switching dynamical systems models for multiple interacting neural populations.

*Advances in Neural Information Processing Systems*, *33*, 14867–14878.

Jazayeri, M., & Afraz, A. (2017). Navigating the neural space in search of the neural code. *Neuron*, *93*(5), 1003–1014.

Jordan, M. I., Ghahramani, Z., Jaakkola, T. S., & Saul, L. K. (1999). An introduction to variational methods for graphical models. *Machine Learning*, *37*(2), 183–233.

Jun, J. J., Steinmetz, N. A., Siegle, J. H., Denman, D. J., Bauza, M., Barbarits, B., ... others (2017). Fully integrated silicon probes for high-density recording of neural activity. *Nature*, *551*(7679), 232–236.

Kohn, A., Jasper, A. I., Semedo, J. D., Gokcen, E., Machens, C. K., & Yu, B. M. (2020). Principles of Corticocortical Communication: Proposed Schemes and Design Considerations. *Trends in Neurosciences*, *43*(9), 725–737.

Linderman, S., Johnson, M., Miller, A., Adams, R., Blei, D., & Paninski, L. (2017). Bayesian learning and inference in recurrent switching linear dynamical systems. In *Proceedings of the 20th international conference on artificial intelligence and statistics* (pp. 914–922).

Mikosch, T. V., Wright Stephen, J., & Jorge, N. (2006). *Numerical optimization*. Springer New York.

Parthasarathy, A., Herikstad, R., Bong, J. H., Medina, F. S., Libedinsky, C., & Yen, S.-C. (2017). Mixed selectivity morphs population codes in prefrontal cortex. *Nature Neuroscience*, *20*(12), 1770–1779.

Remington, E. D., Egger, S. W., Narain, D., Wang, J., & Jazayeri, M. (2018). A dynamical systems perspective on flexible motor timing. *Trends in Cognitive Sciences*, *22*(10), 938–952.

Rodu, J., Klein, N., Brincat, S. L., Miller, E. K., & Kass, R. E. (2018). Detecting multivariate cross-correlation between brain regions. *Journal of Neurophysiology*, *120*(4), 1962–1972.

Saxena, S., & Cunningham, J. P. (2019). Towards the neural population doctrine. *Current Opinion in Neurobiology*, *55*, 103–111.

Semedo, J., Zandvakili, A., Kohn, A., Machens, C. K., & Yu, B. M. (2014). Extracting latent structure from multiple interacting neural populations. *Advances in Neural Information Processing Systems*, *27*.

Semedo, J. D., Zandvakili, A., Machens, C. K., Yu, B. M., & Kohn, A. (2019). Cortical areas interact through a communication subspace. *Neuron*, *102*(1), 249–259.

Zheng, X., Aragam, B., Ravikumar, P. K., & Xing, E. P. (2018). DAGs with NO TEARS: Continuous optimization for structure learning. *Advances in Neural Information Processing Systems*, *31*.

Zoltowski, D., Pillow, J., & Linderman, S. (2020). A general recurrent state space framework for modeling neural dynamics during decision-making. In *Proceedings of the 37th international conference on machine learning* (pp. 11680–11691).