Linear Factor Model for Discovering Lead-Lag Relationship between Two Brain Areas

Heejong Bong [1] Valerie Ventura [1] Eric Yttri [2] Robert E. Kass [1]

[1]Department of Statistics and Data Science, Carnegie Mellon University
[2]Department of Biological Sciences, Carnegie Mellon University

Carnegie Mellon University

Discovering Dynamic Features of Functional Connection

Experiment: Mice performed an uncued, goal-oriented joystick reaching task.

Data: We collected **spike trains** of neurons in **motor cortex** and **striatum**.

Goal: We want to evaluate **functional connection between two brain** areas. We expect **lead-lag relationship**.

Methods: We summarize the numerous pairwise correlations between the two areas using a **multiset inverse canonical correlation sparsification (MICCS)** and develop an **inferential framework** to assess the significance of the discovered functional connectivity.

Multiset probabilistic CCA (pCCA)

$$\underline{X}_i^t | Z_i^t \sim MVN\left(Z_i^t \underline{\beta}_i^t + \underline{\mu}_i^t, \Phi_i^t\right), t = 1, \ldots, T, i = 1, 2 \quad (1)$$

$$\begin{pmatrix} \underline{Z_1} \\ \underline{Z_2} \end{pmatrix} \sim MVN\left(\begin{pmatrix} \underline{0} \\ \underline{0} \end{pmatrix}, \Sigma\right) \quad (2)$$

Each of two neuron populations $\underline{X}_i^t$ are summarized by a 1D latent variable Z_i^t, which we use to assess connectivity between two areas.

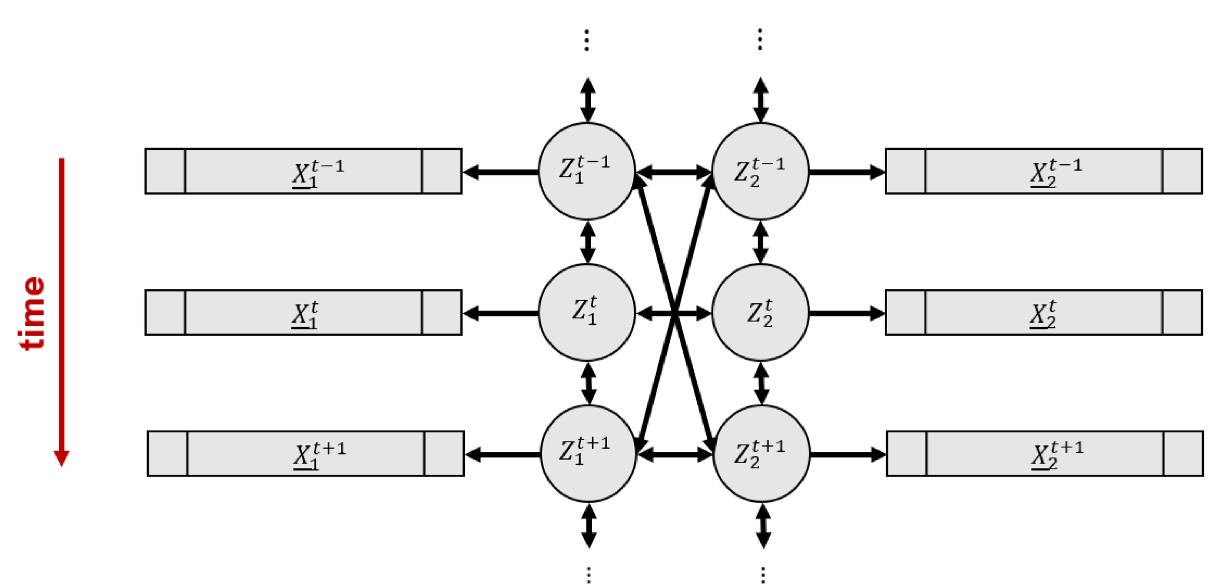


Figure 1: pCCA model

Our goal is to identify which connections between Z_1 and Z_2 are significant.

The precision matrix $\Omega = \Sigma^{-1}$ provides a **graphical interpretation of functional connection**.

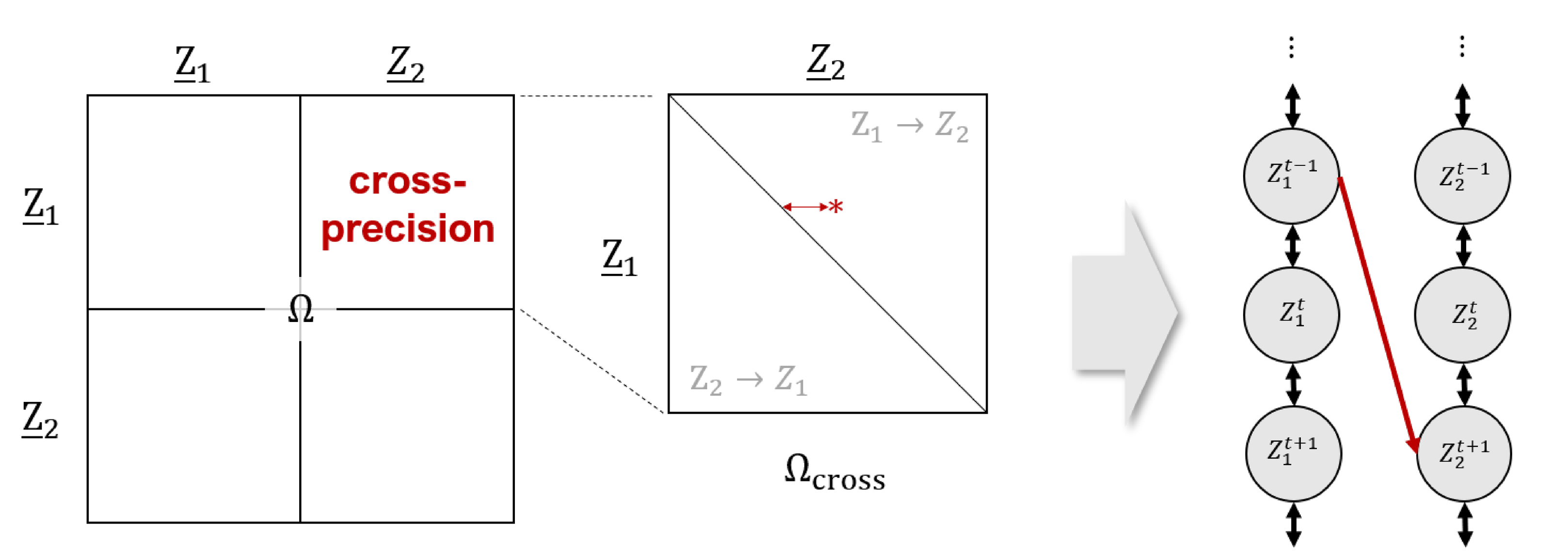


Figure 2: Graphical Interpretation from $\Omega = \Sigma^{-1}$: When a significant element of Omega is identified, e.g., the red star, its coordinates and distance from the diagonal indicate at what time in the experiment a connectivity happens between two brain areas at what lead or lag. $Z_i \to Z_j$ means that area i leads area j in that section of Ω.

MICCS: Multiset Inverse Canonical Correlation Sparsification

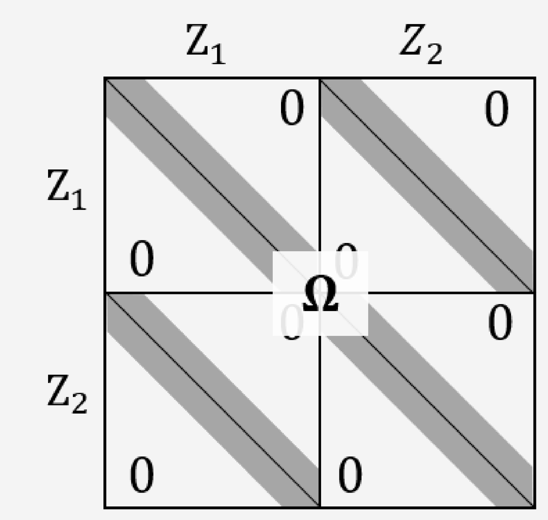
We impose **sparsity constraints** on Ω to stabilize its estimation:

- **forced sparsity:** interested in lead-lag relationship with delay less than 200ms
- **regularized sparsity:** assume dominant association is concentrated

Maximum penalized likelihood estimate:

$$\hat{\Omega} = argmin_{\Omega, \underline{\eta}_i^t} \underbrace{- logdet(\Omega) + tr(\Omega\hat{\Sigma})}_{\text{Negative log-likelihood}} + \underbrace{\lambda_1 \|\Omega\|_1}_{\text{GLASSO penalty}} \quad (3)$$

where Ω satisfies the following form:



(4)

We implemented **coordinate descent method** to find a local minimum.

Statistical inference:

- Entries of **de-sparsified precision** have Gaussian null limiting distributions[1] with a variance estimated by **permuting trials** → We calculate p-value for each entry
- To find areas of concentrated associations, we control the false discovery rate (FDR) by the multiple hypothesis testing method, **STAR**[2].

.

Simulation Analysis

Artificial data with **known underlying functional connections** are simulated from:

- our multiset pCCA model above, e.g. in fig. 3a
- modified experimental spike train data, e.g., in fig. 3b

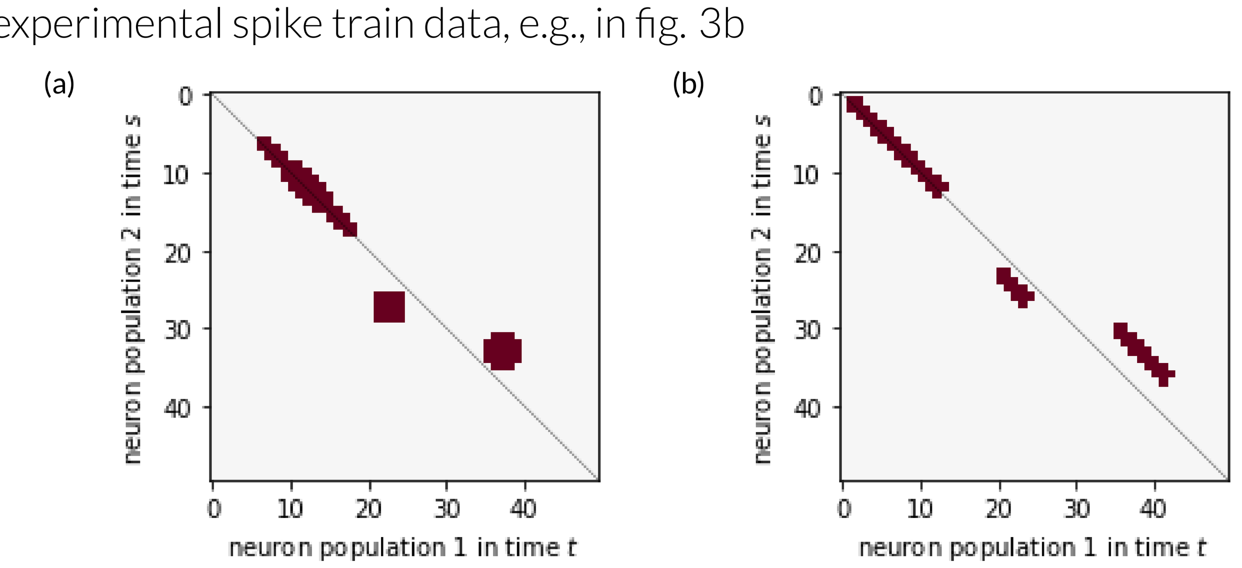


Figure 3: Artificial data are simulated from the pCCA model with true Ω in (a) and from experimental data with true Ω in (b)

STAR on de-sparsified precisions **controls FDR and discovers real signals in high probability**

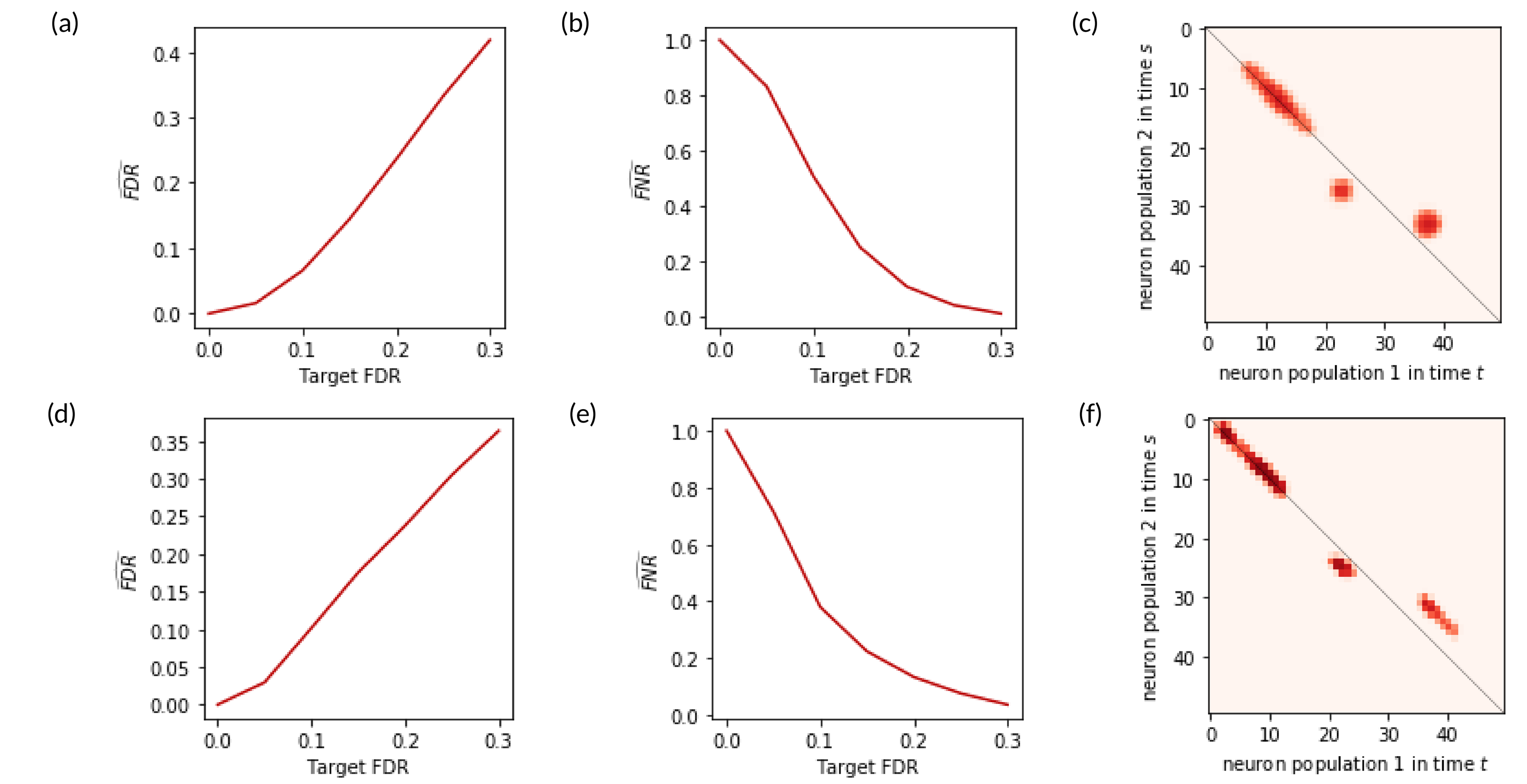


Figure 4: (a) Estimated false discovery rate ($\widehat{FDR}$), (b) false non-discovery rate ($\widehat{FNR}$) in different target FDRs, and (c) probabilities of each precision entry to be discovered at the target $FDR = 0.1$ in simulation from our pCCA model with true Ω in fig. 3a. (d, e, f) The same for simulation from modified experimental spike train data with true Ω in fig. 3b.

Results: Application to Neuropixel data

MICCS on spike counts of 22 M1 and 20 striatum neurons from -1 sec. to 1 sec. after each of 147 movement onsets.

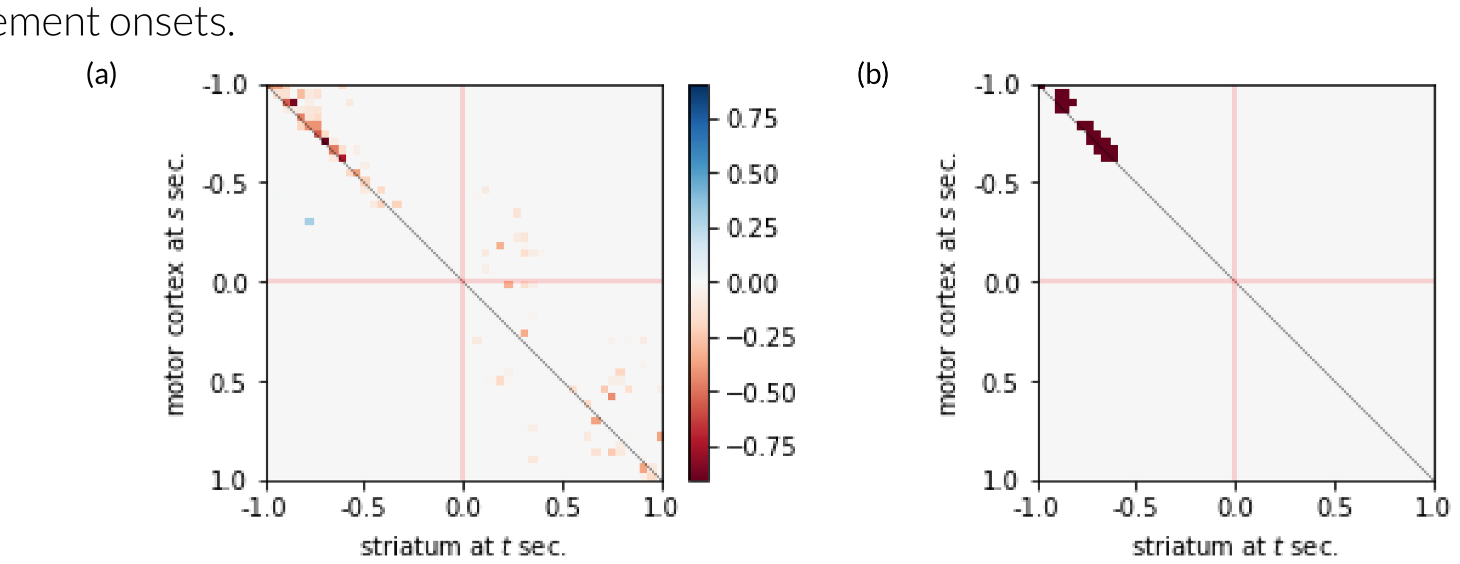


Figure 5: (a) Estimated Ω using MICCS and (b) discovered functional connections using STAR with target $FDR = 0.2$

Discussion:

- MICCS discovered simultaneous (zero lag) connectivity between brain areas before movement onsets.
- MICCS did not discover a hypothesized lead-lag relationship around movement onset: we may lack power due to correlation attenuation from spike train Poisson noise.

Future goal: increase statistical power by **accounting for Poisson noise**.

References

[1] Jana Jankova, Sara Van De Geer, et al. Confidence intervals for high-dimensional inverse covariance estimation. *Electronic Journal of Statistics*, 9(1):1205--1229, 2015.

[2] Lihua Lei, Aaditya Ramdas, and William Fithian. Star: A general interactive framework for fdr control under structural constraints. *arXiv preprint arXiv:1710.02776*, 2017.