Hierarchical Latent Variable Models for Neural Data Analysis Quarter 2 Proposal

Charlie Gillet cgillet@ucsd.edu
Mohamed Elshazly melshazly@ucsd.edu
Nathen Lee nal008@ucsd.edu
Mikio Aoi maoi@ucsd.edu

All our decisions, from simple reflexes to complex reasoning, are driven by patterns of neuronal firing, where networks of neurons communicate through precisely timed spikes to process information and generate actions. Neurons fire by transmitting electrical signals through an action potential, where a rapid influx and outflow of ions from the dendrites across the cell membrane generates an electrical spike that travels down the axon. We measure this activity using techniques like electrophysiology (e.g., probes) or imaging methods such as calcium or voltage-sensitive dyes. The goal of neural encoding and decoding is to find appropriate representations to describe the behavior or stimulus we record and observe through neural electrodes and their subsequent spike trains. So, what is the best way of representing these results?

With the maturation of data collecting technology, researchers can now observe and record hundreds of neurons at once simultaneously, creating bottlenecks in the analytical methods we currently possess. How do we find appropriate representations that can be easily interpreted from multiple large sets of neuron recordings? What methods could be used to create a model that effectively performs analysis on the vast amounts of neural data being collected?

We aim to compound upon established techniques to develop new ways to infer and analyze datasets of different observations in order to find relationships between each modality and interpretable representations of their shared and exclusive patterns. These techniques include factor analysis models (FA), variational inference (VI), probabilistic canonical correlation analysis (PCCA), and gaussian processes (GPs). These 4 methods will all be used as building blocks in constructing a latent variable model that effeectively processes neural data. As for our data, we intend on working with data collected by the International Brain Laboratory (IBL) on experiments performed on mice. These experiments places mice in a cylinder facing a screen with two images displayed and tasks them with turning a wheel towards the image with higher contrast. As these mice are starved of quality water, upon successfully completing each trial, the mice are rewarded with better quality water. The dataset generated from this experiment contains data on various different mice over numerous trials. We intend to train the model on this data and eventually have it perform to a level to which it can be applicable to any neural data analysis.