Using Canonical Correlation Analysis (CCA) to analyse neural data

Canonical correlation analysis (CCA) and its more advanced machine learning variants (kernel CCA) are powerful methods to analyse data, in particular neural data. CCA finds the maximal correlation between two sets of data (input and output) by finding a linear combination for each such that the resulting two projections are maximally correlated. In other words, CCA can be seen as the problem of finding basis vectors for two sets of variables such that the correlations between the projections of the variables onto these basis vectors are mutually maximized.

This proposal suggests to use this method to analyse spike trains generated by a network of spiking neurons, with an architecture proposed by Deneve et al. called the predictive coding framework (http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003258). These neurons encode a stimulus collaboratively in such a way that the stimulus can be decoded from them linearly. The encoding, however, is non-linear.

The purpose of this project would be to determine how well CCA can be used to recover the encoding and decoding filters of such a network. If time permits, this analysis can be extended to real data from the retina, and/or simulated data from a popular model of neural responses called the generalized linear model (GLM). Furthermore, a more sophisticated extension called kernel CCA can be used to investigate whether this aids in recovering neural filters.

For more information about CCA, consult e.g. http://davidroihardoon.com/Professional/Publications_files/NC_Hardoon_2817_reg.pdf

This project would suit students who are mathematically inclined and like to build computational models in order to analyse and understand neural data.