**Efficient Spike Sorting for Long Electrophysiology Recordings Using Batch Selection for Template Generation**

Abstract

Spike sorting is a critical process in neural data analysis that aims to identify and classify the actional potential activity (“spikes”) of individual neurons from extracellular electrophysiological recordings. Traditional recordings might last 8 hours, but to capture long-term brain changes critical to mental health, multi-day or multi-week recordings will be necessary. Current spike sorting methods such as KiloSort rely on processing whole recordings, making the processing of longer recordings very computationally expensive - thereby preventing sorting of the longest recordings. We present a novel method for spike sorting that utilizes KiloSort's template-matching algorithm while introducing subsampling to allow the handling of larger datasets. In particular, rather than spike-sorting multiple individual segments of 12-24 hours, followed by subsequent spike-matching, which can introduce error, we sought a means to sort across the full span of a recording to avoid this post-hoc processing. To do this, we spike sort on only sections of the data that are spaced evenly across the full recording - for example, one minute out of every ten. Initial findings indicate that this approach yields high-quality replications of full dataset spike sorting. Detailed quantifications of optimal parameters are ongoing. In a second approach, we seek to select data segments that are optimal for spike template development due to having the highest information content. We then plan to focus computational resources on these selected data segments Overall these approaches open the door to reliably and efficiently spike sorting datasets spanning days or weeks.

Introduction

* Talk about importance of long-term spikesorting

Methods

Results

Discussion