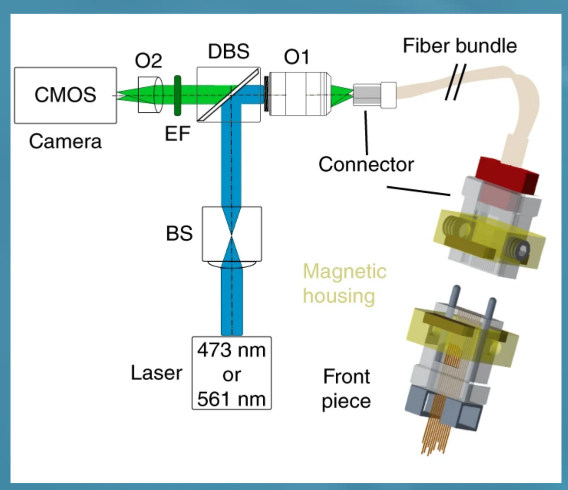
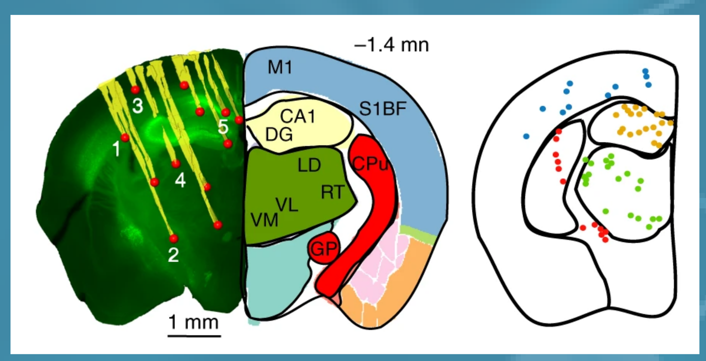
Comparison of functional connectivity estimator performance for whole-brain mesoscopic calcium indicator data in behaving mice

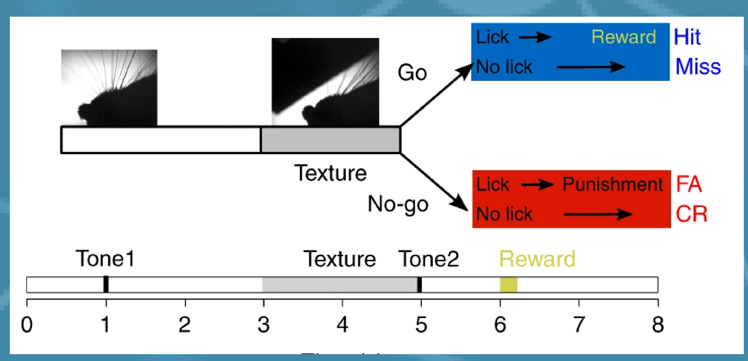
Aleksejs Fomins, Yaroslav Sych, Fritjof Helmchen



GCaMP6f mice are chronically implanted with 12 to 48 channel optical fibers, imaging a selection of task-related cortical and sub-cortical areas. Given fiber diameter of ~0.1mm, each channel records the average DF/F signal of ~10^3 neurons.

The head-fixed water-restricted mice perform 8 second trials of a texture discrimination task. They are presented an auditory cue followed by a texture which they discriminate by whisking. Following another auditory cue at the end of the texture presentation period, mice decide whether to lick at the water port. In case of a correct decision to lick (Hit) a water reward is applied at the port. In case of a wrong decision to lick (FA), a white noise punishment is applied. The other two outcomes (Miss and CR) are neither rewarded nor punished.

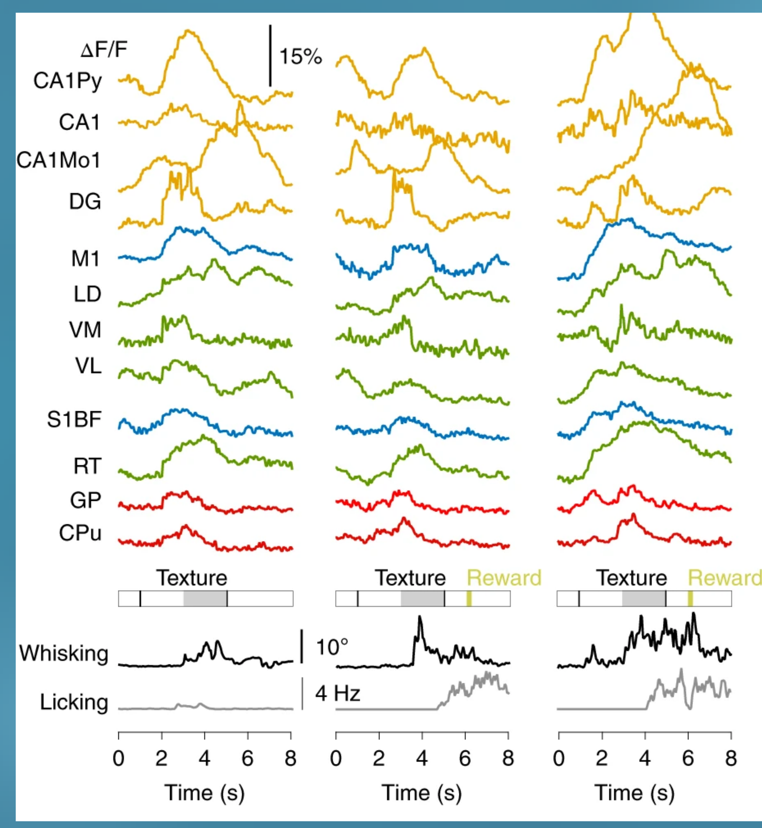




On each training day, several hundred trials are performed consecutively, with a few seconds break inbetween. A mouse is considered expert if it surpasses 70% accuracy on a given day.

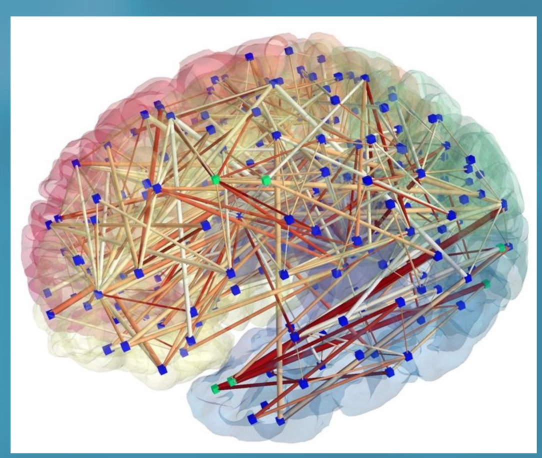
Behavioural data, such as paw movement, whisking and licking activity, moments of first texture touch and reward (if applicable) are recorded.

Sych et al, Nature Methods 2019



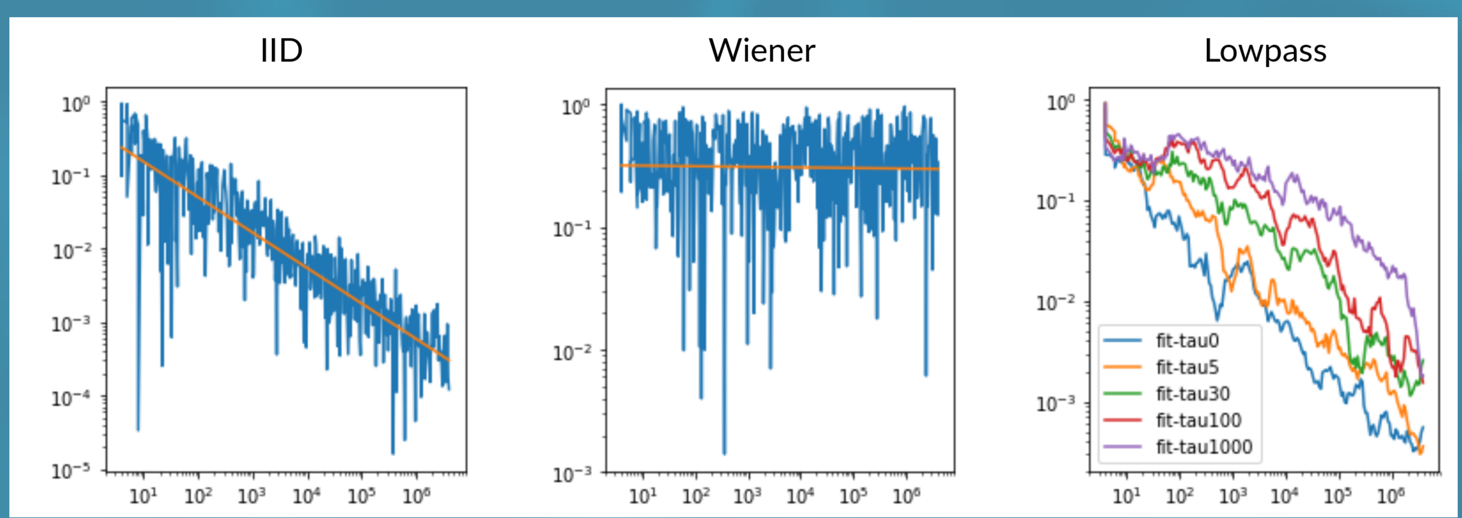
Functional Connectivity (FC) quantifies the capacity of one variable to be predictive of another given simultaneous measurements of their values. In neuroscience both neuronal and behavioural data can be used to infer neuronal and neuro-behavioural FC. It resembles Anatomical Connectivity (AC), however, the strength of the connections is determined by their activity (or lack of it) during the recording, and thus is task-specific.

FC estimators may use time-delay to infer the effect of past values of one variable on present values of another. In presence of multiple variables, bivariate estimators seek only use a pair of variables at a time, while multivariate estimators consider dataset as a whole, taking account of potential proxies, synergies and common ancestry.



FC has been used extensively in fMRI and EEG, and is recently being applied to calcium imaging data.

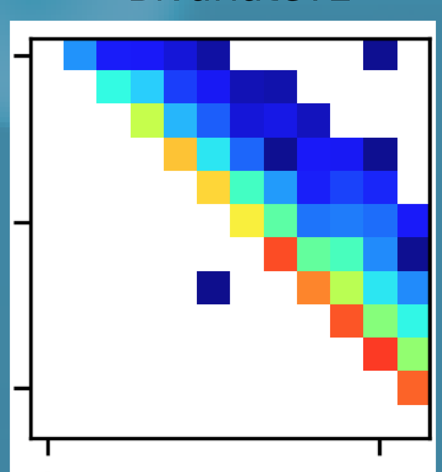
We have studied several model-free estimators: Cross-correlation and Spearmann rank are fast to estimate, but suffer from large amounts of false positives in case of non-iid data (which is the case for neuronal and behavioural data), as well as false negatives due to inability of finding non-linear relationships.



Cross-correlation of white noise as function of data size

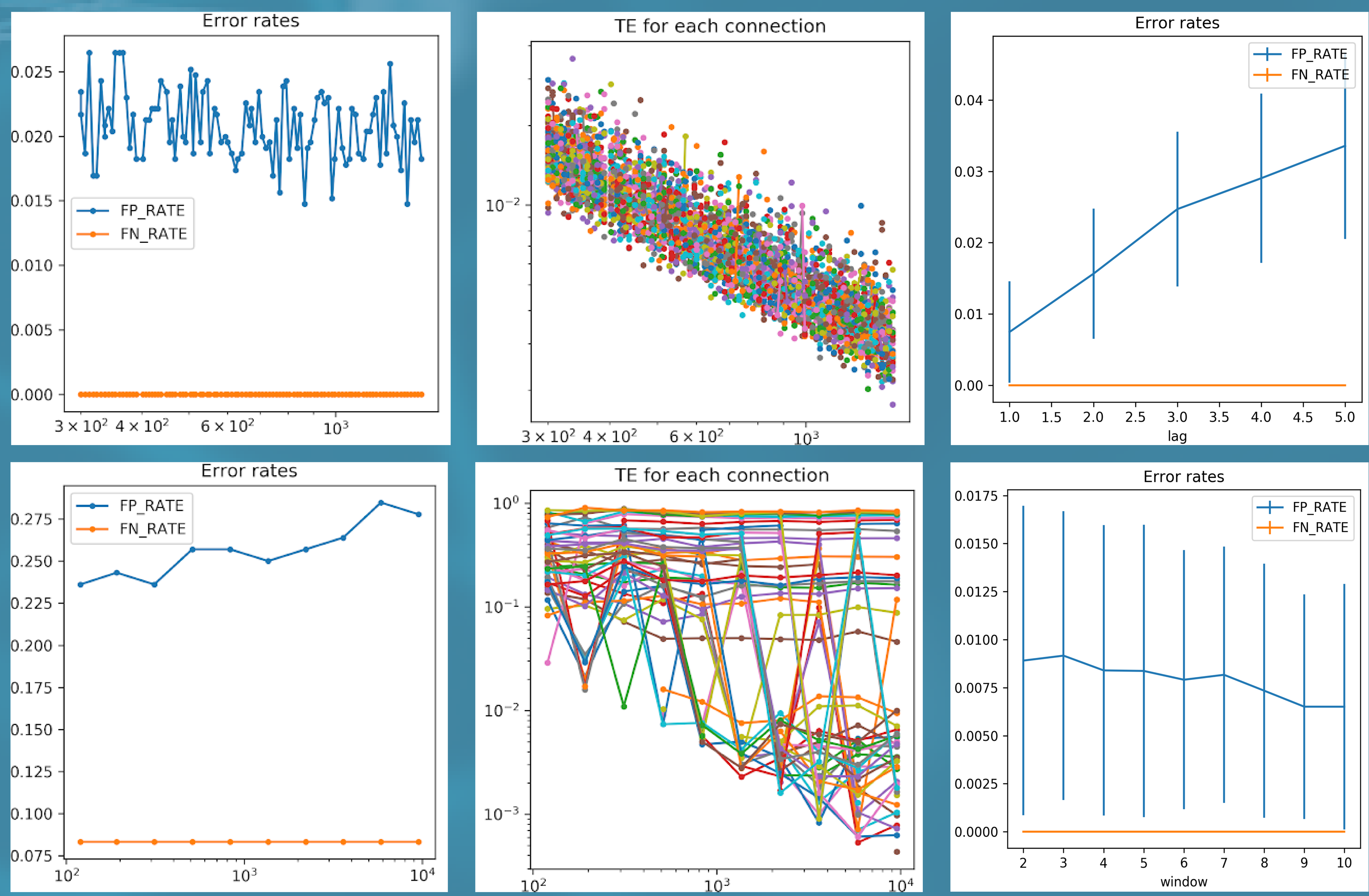
Mutual Information estimators can handle arbitrary non-linear relationships, but is slower and more data-hungry. A further extension - Transfer Entropy estimators - also solve the non-iid problem, as they specifically exclude communication of information already contained within the target variable.

CrossCorrelation BivariateTE MultivariateTE



FC estimated for sequentially connected neurons

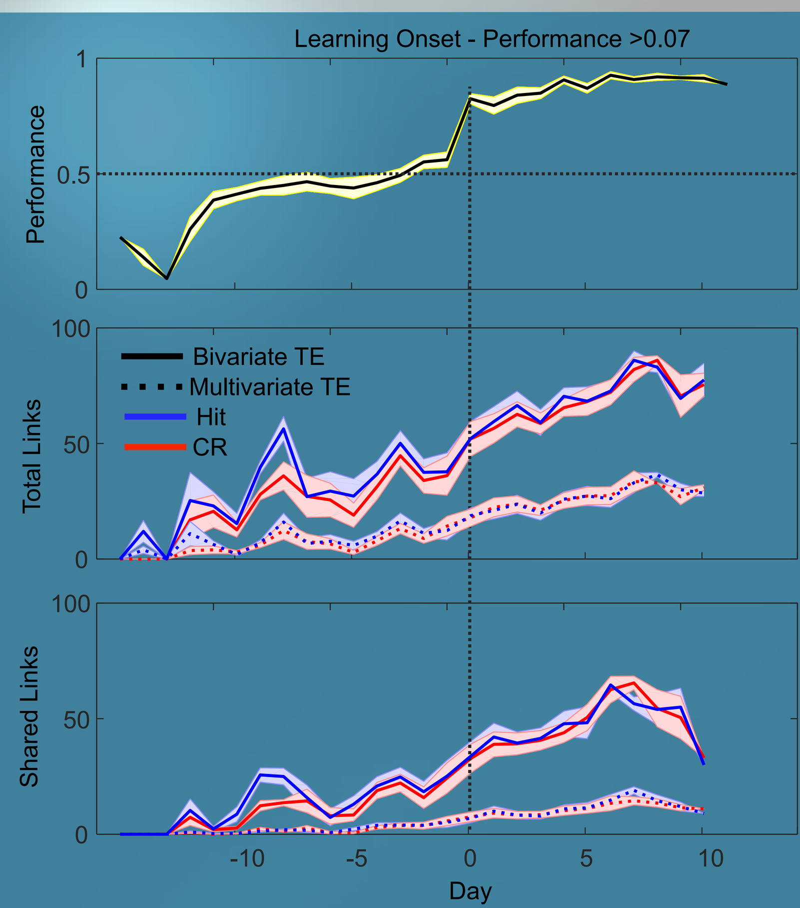
Another approach to FC is to hypothesize several competing models with known dynamics and connectivity, and compare them using bayesian model selection criteria. Multivariate Autoregression (aka Granger Causality) fits a linear function of several past values of the network to predict its current values. Dynamical Causal Model uses a forwards model, which describes how neuronal activity is converted into observables (e.g. calcium activity), as well as a neuronal model, which describes how neuronal populations interact. Both models are fitted simultaneously, obtaining FC for underlying neuronal populations.



We have studied performance of TE connectivity estimators for simulated data with known connectivity. We demonstrate that the number of false positives remains constant with data number (top left), and number of false negatives remains at zero, meaning that TE is very good at finding true links (bottom left). The absolute magnitude of false links decreases with data size (top middle), while the magnitude of true links stays high, resulting in magnutude separation with data size (bottom middle). We must note however that with real data we have not been able to obtain sufficient data to observe this separation, so we rely on the constant error rate when making predictions about robustness of our estimation

We also study the effects of maximal lag and window size parameters on the estimated error rates. Increasing maximal lag increases false positive rate, as it effectively increases the number of tests to be performed. Increasing the window counter-intuitively decreases the number of false positives, as they become less likely to pass the threshold given more data.

For most naive mice the task performance improves gradually, reaching expert levels after several days of training. FC in their brains also changes during this period. We have computed TE using 200ms timestep sampling rate, using up to 5 timestep history, resulting in a sliding window of 1s. We show that on average, the total number of significant links present in such a window increases with learning. This is true for both more conservative Multivariate estimator, and the less conservative Bivariate estimator. The total number links shared between neighboring timesteps increases as well.



In the future we will proceed to Implement a DCM to simulate our experiment, and compare its performance to that of MAR and TE for simulated and real data. In addition, we will compare the known metrics of connectomics to find those most informative for our experiment. Finally, we will report detailed statistics on involvement of neuro-neuronal and neuro-behavioural links for different behavioural regimes, presenting significant changes in distribution and freqnency of appearance, as well as identify any emerging sub-circuits.