

Master Thesis

---

**Correlation analysis of signals  
obtained by calcium imaging in the  
primary visual cortex of mice**

---

Xiuzhi Li

Supervisor: Prof. Dr. Stefan Rotter

M.Sc. Neuroscience  
University of Freiburg

October 4<sup>th</sup>, 2022

## Abstract

Calcium imaging throws light on studying the activities of a large neuronal population. However, analyzing fluorescence signals is challenging. Reconstructing spike trains from fluorescence signals by deconvolution is very sensitive to noise, not stable, and linked with high computational costs. In this thesis, we propose and demonstrate a new method to obtain spike train correlations by deconvolution of the correlations of fluorescence signals, avoiding the reconstruction of spike trains altogether. Working with surrogate data generated by various point processes, it is shown that the method captures the correlation structure faultlessly. The inferred cross-covariance is represented with experimental data obtained from calcium imaging of the mouse's primary visual cortex, with which a connectivity inference is attempted. In addition, a validation process by applying the procedure of handling experimental data on surrogate data demonstrated that the cross-covariance of underlying spike trains could be estimated with high reliability. Finally, we show that the methodology can be used to analyze the correlation of membrane potentials in a sub-threshold regime based on a shotnoise model of simple linear-modeled neurons.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Calcium Imaging . . . . .	2
1.1.1	Intracellular calcium dynamics and its modeling . . . . .	3
1.1.2	Calcium indicators . . . . .	5
1.1.3	Data analysis and its difficulties . . . . .	7
1.2	Neural coding: from firing rate to correlation . . . . .	8
1.2.1	Correlation quantification . . . . .	9
1.3	Question statement . . . . .	10
<b>2</b>	<b>Materials and Methods</b>	<b>11</b>
2.1	Software and hardware . . . . .	11
2.2	Fluorescence signal modeling . . . . .	11
2.2.1	Point process . . . . .	12
2.2.2	Kernel . . . . .	14
2.2.3	Exact integration as fast convolution . . . . .	16
2.3	Signal characterization and correlation inference . . . . .	17
2.3.1	One signal . . . . .	17
2.3.2	Two signals . . . . .	18
2.3.3	Multiple signals . . . . .	24
2.3.4	Modeling membrane potential . . . . .	24
2.4	Pre-processing of experimental data . . . . .	25
<b>3</b>	<b>Results</b>	<b>27</b>
3.1	Simulated fluorescence data . . . . .	27
3.1.1	Kernel selection . . . . .	27
3.1.2	Independent Poisson process . . . . .	28
3.1.3	Correlated Poisson process with double input . . . . .	29
3.1.4	Non-stationary Poisson process . . . . .	30
3.1.5	SIP . . . . .	31

3.1.6	MIP . . . . .	32
3.1.7	HOC . . . . .	34
3.2	Experimental data . . . . .	35
3.2.1	Data normalization . . . . .	36
3.2.2	Kernel fitting . . . . .	36
3.2.3	Validation of kernel fitting . . . . .	38
3.2.4	Correlation analysis . . . . .	39
3.2.5	Validation with simulated data . . . . .	43
3.3	Membrane potential analysis . . . . .	46
<b>4</b>	<b>Conclusion and discussion</b>	<b>49</b>
4.1	Baseline . . . . .	49
4.2	Pre-processing of experimental data . . . . .	50
4.3	Estimation of unknown kernel . . . . .	51
4.4	Outlook . . . . .	52
<b>Acknowledgements</b>		<b>54</b>
<b>References</b>		<b>58</b>

# List of Figures

1	Imaging and non-imaging techniques . . . . .	2
2	Simultaneous electrophysiology and fluorescence measuring . . . . .	3
3	Comparison between action potential and fluorescence trace . . . . .	4
4	Intracellular calcium dynamics . . . . .	5
5	GCaMP working mechanism . . . . .	5
6	Calcium indicator . . . . .	7
7	Correlation quantification . . . . .	10
8	Response kernel . . . . .	15
9	Exact integration . . . . .	17
10	Method overview . . . . .	22
11	Method illustration . . . . .	23
12	Data pre-processing . . . . .	26
13	Simulation with different kernels . . . . .	28
14	Simulated data via stationary Poisson process . . . . .	29
15	Simulated data via correlated stationary Poisson process with doubled input . . . . .	30
16	Simulated data via non-stationary Poisson process . . . . .	31
17	Simulated data via SIP process . . . . .	32
18	Simulated data via MIP . . . . .	33
19	Simulated data with HOC spike trains . . . . .	34
20	Neurons under FOV . . . . .	35
21	Data normalization . . . . .	36
22	Kernel fitting of experimental data . . . . .	37
23	Time constants . . . . .	38
24	Validation of kernel fitting . . . . .	39
25	Cross-covariance estimation of multiple neurons . . . . .	40
26	Estimated cross-covariance of seed neuron . . . . .	41

27	Estimated connectivity with signal covariance . . . . .	42
28	Estimated connectivity with inferred covariance . . . . .	43
29	Time constant comparison . . . . .	44
30	Estimation and ground truth comparison . . . . .	45
31	Estimated cross-covariance of simulated neurons . . . . .	46
32	Simulation design . . . . .	48
33	Membrane potential analysis . . . . .	48
34	Shotnoise as baseline . . . . .	50

## List of Tables

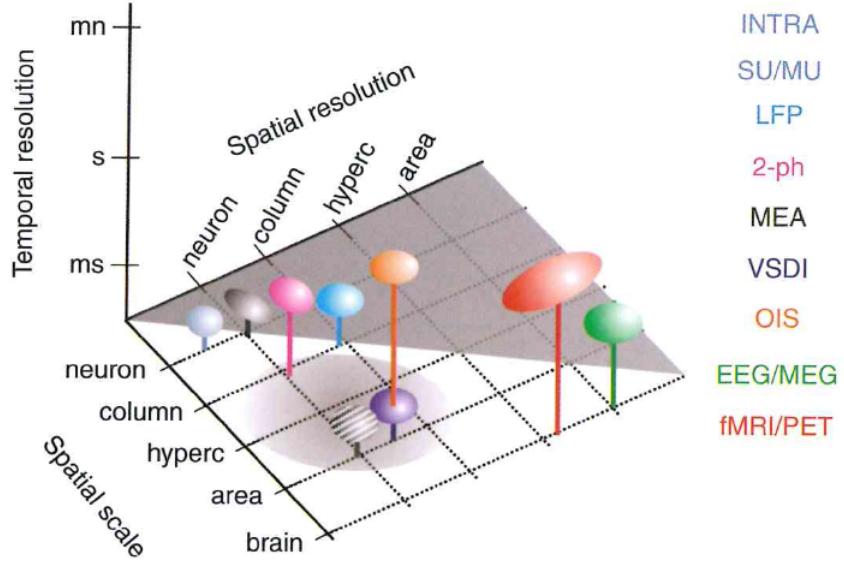
1	Simulation parameters of LIF_PSC_alpha neuron . . . . .	47
---	---	----

## List of Algorithms

1	Simulation of stationary Poisson process . . . . .	12
2	Simulation of non-stationary Poisson process . . . . .	13
3	Simulation of SIP . . . . .	13
4	Simulation of MIP . . . . .	14
5	Construction of HOC via component rates . . . . .	15

# 1 Introduction

Recording neural activity is critical for neuroscience research. Neural activity sculpts how neurons transfer information and interact with each other. Different markers or indicators can mediate it. Nowadays, a variety of techniques have been developed to measure brain activity, such as single-channel recording, multi-channel recording, local field potentials (LFPs), electroencephalogram (EEG), magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), positron emission tomography (PET), calcium imaging, optical imaging of intrinsic signals (OIS), voltage-sensitive dye imaging (VSDI). These techniques can be classified into two main categories: electrophysiology and imaging techniques. Electrophysiology (patch clamp, multi-channel recording, LFPs, EEG, MEG) measures and manipulates neuronal dynamics such as spike trains, EPSC, EPSP, and LFP. Imaging techniques (fMRI, PET, calcium imaging, OIS, VSDI) can capture neuronal activities and identify structures and connections at different scales, from cortex areas and neurons to dendrites and spines. Each technology has advantages and disadvantages regarding invasiveness and temporal and spatial resolution. Therefore, combining two or more technologies to reveal the complex spatiotemporal patterns of neural activity from a single neuron to the population level is common. Figure 1 gives an overview of temporal resolution and spatial scale.

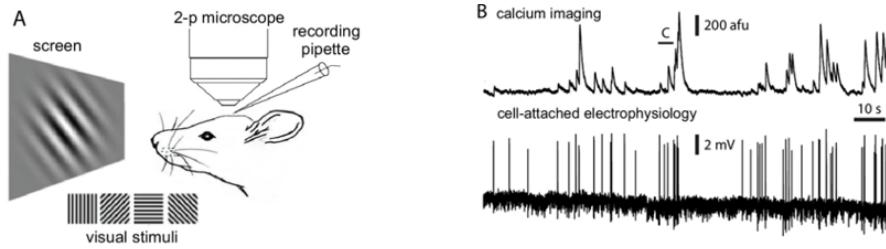


**Figure 1:** Temporal resolution and spatial scale of imaging and electrophysiology techniques. The mesoscopic scale is represented by the oval shaded area [1]. INTRA: intracellular recording, SU/MU: single-unit/multi-unit recording, LFP: local field potential, MEA: multi-electrode array, VSDI: voltage-sensitive dye imaging, OIS: optical imaging of intrinsic signals, EEG/MEG: electroencephalography/magnetoencephalography, fMRI/PET: functional magnetic resonance imaging/poistron emission tomography.

## 1.1 Calcium Imaging

As one of the imaging techniques, calcium imaging has become more popular in the several decades. Neuronal spiking is assessed by measuring changes in fluorescence. Calcium imaging together with recombinant DNA technologies and microscopy is a powerful tool in modulating and visualizing cellular and molecular interactions. It can be used not only to observe neuronal activity but also to associate behavioral characterization with physiological states. Under the field of view (FOV), it is possible to record the activities of hundreds of neurons. It also enables recording free-moving animals in in-vivo experiments. Two-photon microscopy has excellent spatial resolution, but because the calcium signal is slow, two-photon calcium imaging has very good spatial resolution but low temporal resolution. Although experiments' setup and execution could be demanding, calcium imaging can be combined with

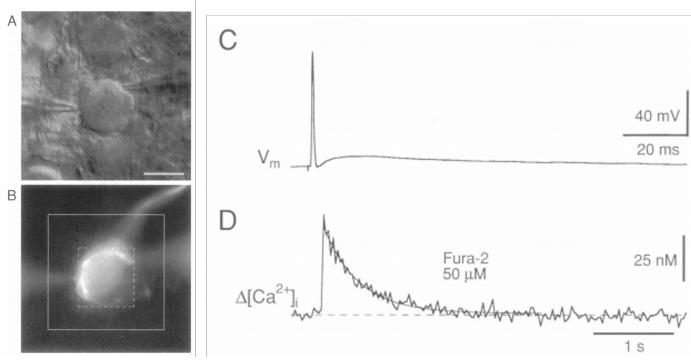
electrophysiology to compensate the low temporal resolution. Figure 2 is an example. Intra-cellular recording directly measures neuronal activities, usually taken as ground truth. Calcium imaging records fluorescence signals as an indicator of neuronal spiking activity, which is indirect. Fluorescence signals and neuronal spikes differ from each other. Figure 3 shows single action potential and its evoked calcium signal, obtained from recording rats' calyx of Held - a large synaptic terminal in brainstem slice. We see action potential is quick (in ms) with a short decay, while fluorescence signal (in this case, Fura-2) is slow in several seconds with a longer decay. The fluorescence signal is much noisier than the action potential. The calcium indicator Section 1.1.2 is one key point to facilitate functional imaging with alive animals. For example, with a cranial window over visual cortex of a transgenic mouse, it is able to image large scale of neuronal activities on layer 2/3 and layer 4.



**Figure 2:** In-vivo calcium imaging accompanied with simultaneous electrophysiology recording on genetically modified mice to produce a calcium indicator (GCaMP6s, GCaMP6f) in neurons of the primary visual cortex, adapted from [2]. Left: Experimental design that recruits two-photon imaging and cell-attached recording simultaneously. Right: recorded fluorescence trace (top) and membrane potential trace (bottom). afu: arbitrary fluorescence units.

### 1.1.1 Intracellular calcium dynamics and its modeling

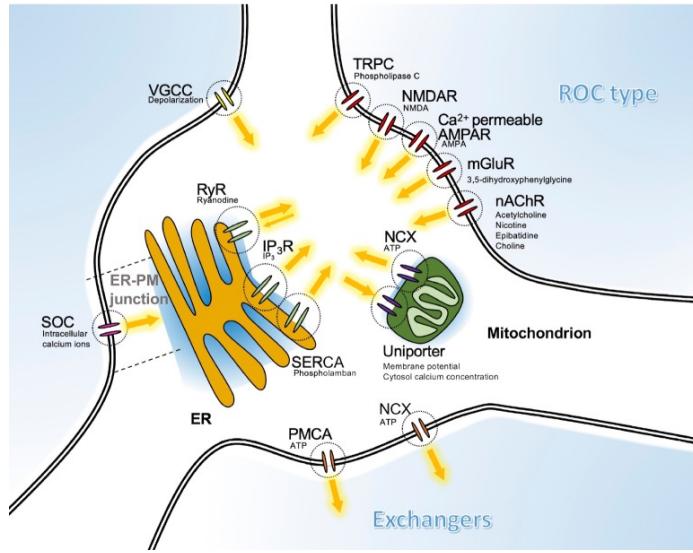
Calcium imaging visualizes calcium signaling within neurons based on intracellular calcium dynamics. The interaction of calcium ions between the extracellular milieu and intracellular space generates the dynamics of calcium activity. There are mainly three types of calcium channels on the membrane. See Figure 4. The first is voltage-gated calcium channels (VGCC); the second is receptor-operated calcium channels (ROC); the third is store-operated calcium channels (SOC). VGCC responds to electrical signals induced by calcium binding. It helps to propagate electrical impulses.



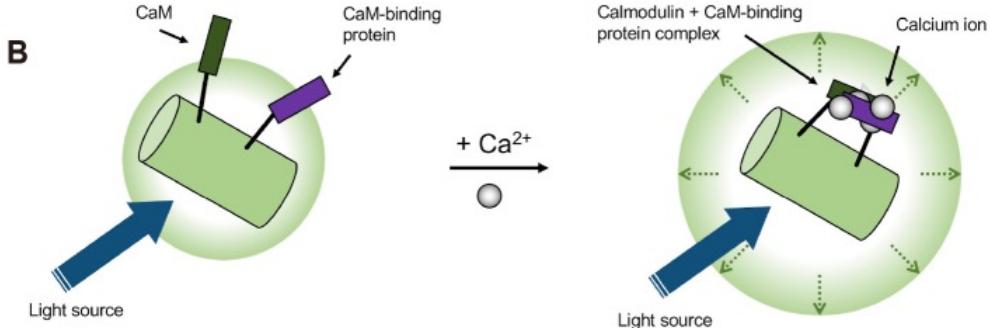
**Figure 3:** Fluorescence signal evoked by a single action potential, adapted from [3]. A: intracellular recording. B: Calcium imaging of A. Squares indicate regions selected in different types of fluorescence measurements. C: Presynaptic action potential evoked by orthodromic stimulation. D: Presynaptic fluorescence signals evoked by a single action potential in cells loaded with Fura-2.

ROC is a calcium-permeable receptor channel. They are open when binding with excitatory neurotransmitters, e.g., glutamate, AmPAR, NMDA. Besides, the endoplasmic reticulum (ER) and mitochondria can release calcium intracellularly. SOC is activated by stromal interaction molecule (STIM) proteins accumulated at ER and plasma membrane (PM) junctions [4]. In excitatory neurons, when membrane is depolarized, VGCC open and calcium enormously flux in. Therefore calcium concentration is associated with action potential events. Helmchen 2012 [5] summarized calcium dynamics as four processes: calcium binding, influx, extrusion, and diffusion. Calcium influx will increase the intracellular concentration. Extrusion is, therefore, important to keep a low intracellular calcium concentration at a balanced level. In addition, free calcium ions, as well as calcium-binding molecules, e.g., most calcium indicators, can diffuse. Mathematical equations were used to describe the four processes. There are various models for them. For example calcium microdomain models high calcium concentration in small spatial scale beneath the membrane due to opening of calcium-permeable channel pores (diffusion process). An 1-dimensional diffusion model to depict the radial distance dependent calcium redistribution processes. A leaky "chemical" cable analogue can model calcium diffusion, extrusion and buffering. In a simplified single-compartment model, an exponential function can be used to describe the kinetics of the binding process. The calcium signal from a neuron is a convolution of the spikes with this exponential function as proposed [5]. In biological calcium imaging experiments, fluorescent calcium indicators are needed

to serve as a proxy for the change of calcium concentration during binding process.



**Figure 4:** Intracellular calcium dynamics [4].

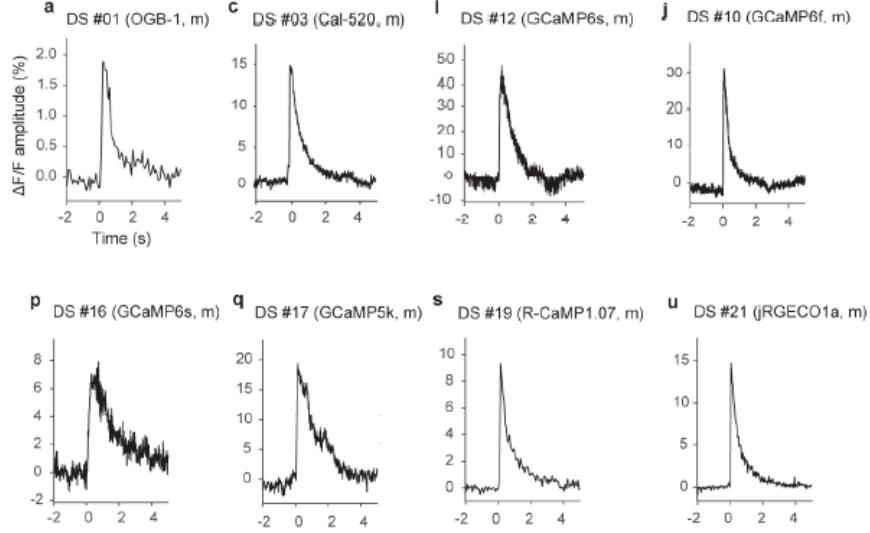


**Figure 5:** GCaMP working mechanism [4].

### 1.1.2 Calcium indicators

The development of optical reporters promotes the imaging techniques. The first calcium indicator is Aequorin found in *Aequorea victoria* in 1962. Nowadays, the indicators used in calcium imaging can be classified into two categories: organic dyes and genetically-encoded calcium indicators(GECIs). The former are small organic

molecule indicators,e.g. Oregon Green BAPTA-1. The latter are calcium-sensitive fluorescent proteins, which further have two major classes: FRET-based indicators and single protein indicator. Fluorescence resonance energy transfer (FRET) happens on the distance change between a donor fluorescent protein and an acceptor fluorescent protein, which can be induced by a conformation change via calcium binding to the linker. FRET leads to fluorescence emissions so that the intensity can be recorded. Single protein indicator e.g. the GCaMP family, are circularly permuted protein that can sense the conformational change made by calcium binding, mediated by the fluorescence intensity of chromophore inside [5]. GCaMP is a synthetic protein containing green fluorescent protein (GFP), calmodulin and a peptide sequence M13 (Figure 5). When binding to calcium, it emits green luminance with a peak excitation wavelength of 480 nm and a peak emission wavelength of 510 nm. Comparing with organic dyes, GECIs has an advantage of being able to target on location in the brain, specific neuron types and subcellular compartments such as soma or axons. GECIs also has higher signal-to-noise ratio. To be noticed is, calcium indicator itself has an influence on the intracellular calcium dynamics, which can inference the readout information. Besides that, different indicators have different response times, which is an important factor of data modeling. In a recent work [6], with a large dataset of curated ground truth recordings on zebrafish and mouse, covering different calcium indicators and induction methods, excitatory and inhibitory neurons, different sampling frequency and different brain regions, using regularized deconvolution to generate the linear kernels of fluorescence transient. The kernels are optimized such that the trace  $\Delta F/F$  is maximal approximated by convolving the ground truth spike train with the kernel. It was found that depending on animal types, brain regions, neuron types and calcium indicators, the area under the kernel curve varied. Figure 6 shows different peak amplitudes and decay times with different calcium indicators. Even for neurons within same dataset, the kernels appear in diversity as well. This suggests that single neuron responds differently when underlying action potential goes on.



**Figure 6:** Kernels with different calcium indicator, adapted from [6]. Across datasets as well as within datasets, kernels vary in terms of amplitude and shape. m: mouse, zf: zebrafish.

### 1.1.3 Data analysis and its difficulties

Calcium imaging throws light on studying the activities of a large neuronal population. Together with genetic encoding techniques, it enables in-vivo experiments with head-fixed or even free-moving awake animals, which is a big advance in scientific research. In the meantime, analyzing fluorescence signals is more complex than analyzing signals from electrophysiology. The raw data of calcium imaging is movie or image stacks. A common benchmark of data analysis involves motion correction, source extraction, cell registration, spike inference, and further analysis of spike trains. Spike inference is very challenging among these steps. Assuming fluorescence signal is the result of action potentials convolving a response function. To infer the underlying spike train is a so-called "inverse problem." Various methods have been developed to solve the inverse problem: deconvolution techniques [7, 8], maximum a posteriori (MAP) principle [9], template-matching-based approach [10], approximate Bayesian inference [11]. Applications and packages are available online: CellSort [12], CaImAn [13], Suite2P [14], MLspike [15], CASCADE [6] and so on. It has been great progress with these works. However, general difficulties in terms of spike inference still exist. The main aspects are outlined below.

- **Overexpression.** Fluorescence overexpression potentially leads to toxicity. Neuron nuclear is bright due to the accumulation of calcium indicators.
- **Photobleaching.** The fluorescence signal fades during the experiment due to the photochemical destruction of the fluorophore.
- **Nonlinearity of fluorescence.** The saturation curve of the calcium-binding ratio can be linearly approximated only within a limited range. Therefore the amplitude of observed fluorescence signals does not always linearly represent the number of underlying action potentials.
- **Sample resolution.** With high microscopy resolution, the single action potential is more visible. The low resolution of imaging could lead to missing action potentials.
- **Noise.** Deconvolution is very sensitive to noise. Fluorescence signals contain much noise from experimental settings, background activity, interference of the calcium-binding domain, and calcium buffering with the cellular spiking process.
- **Neuron type.** Different neuron type demonstrates different response kernels. E.g., inhibitory neurons tend to have shorter decay.
- **High computation cost.** Deconvolution is a computation intensive technique. Applying it to a population of neurons needs high computation resources.

Deconvolution has to be performed to extract spike trains from fluorescence signals. This method, however, is prone to be numerically unstable and noise sensitive. Focusing on correlation analysis and to avoid the deconvolution problem, in this thesis, a new approach is proposed to indirectly infer the cross-correlation of underlying spike trains.

## 1.2 Neural coding: from firing rate to correlation

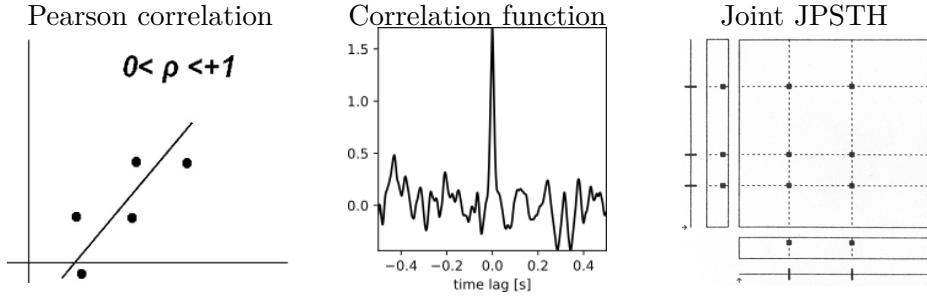
In characterizing single neuron activity, the firing rate is the main indicator. It gives information on whether the neuron is sensitive or not involved in a certain

activity. Especially when the subject is experiencing some events on behavior level, by observing single neuron activity or population activity, researchers can attribute single neuron event to that even in macro-world, for example, "Grandmother" neuron. Besides single neurons, neuronal population activity is also of interest. To describe a population of neurons, mean firing rate, coefficient of variation (CV) of inter-spike-interval can be applied. From single neuron to neuronal population, on the one hand, precise spiking time is important for neural coding. On the other hand, the neuron does not fire alone. The coordinated activities of neurons reflect a certain response pattern. There are assumptions about the role of correlated neuronal activities, such as correlations as a side-effect of synaptic interaction, as an additional dimension of coding in the brain, or as it regulates the flow of information [16]. No doubt, correlation is the essential information and modulates the firing rate. Neurons in the network are sensitive to this correlated activity as input [17]. Only synchronized input can trigger spikes in feed-forward inhibition networks, resulting in sparse and precise response [18]. Sources of correlations can be shared input to neurons or direct synaptic connections. Correlation itself does not mean causation. Correlated activities of neurons do not indicate which neuron drives which neuron. Correlations have something to do with time scale. On different time scales, there might be different correlations.

### 1.2.1 Correlation quantification

In this thesis, we mainly discuss the correlation between paired neurons. There are different ways to quantify correlations of neuronal activity, among which the most well-known are Pearson correlation coefficient, correlation function, and joint peri-stimulus time scatter diagram (JPSTH) [19]. The Pearson correlation coefficient is a linear correlation coefficient with normalized value scope [-1, 1]. Negative values mean negative correlations, while positive values indicate positive correlations. Absolute value towards 1 indicates strong relations, while around zero means two variables are hardly linearly correlated. The Pearson correlation coefficient cannot capture a nonlinear relationship between two variables. With JPSTH it is to create a two-dimensional scatter diagram of the firings of the two neurons relative to each stimulus onset. Each dot in Figure 7 right panel is a coincidence of the two spike trains during one stimulus period. Repeating this for each trial, dot density is built on the diagonal. With subsequent dynamic correction, including subtraction and scaling, reintegration along the diagonal generates the ultimate JPSTH. The cross-

correlation function is commonly used to measure the similarity of two signals. A cross-correlation function is generated by shifting one signal from the other and calculating the expectation value of the dot product of two signals within the overlap. Prominence peaks in the curve demonstrate correlations between the two signals with certain time lags (Figure 7, middle panel).



**Figure 7:** Correlation quantification. Left: Pearson correlation with a positive correlation coefficient. Middle: correlation function. Right: JPSTH [19].

### 1.3 Question statement

Calcium imaging with high-resolution microscopy is a powerful tool in neuroscience research. A reliable analysis of fluorescence signals is important to correctly reveal the biological meaning on a neuronal level. Among all analyses, inferring spike trains from fluorescence signals is challenging but a prerequisite of other steps. On the other hand, neuronal information lies in firing rates and is delivered by correlations. A common methodology for correlation analysis is first to extract discrete spikes or spike rates via deconvolution and then calculate neuronal correlations. However, the reconstruction of spike trains by deconvolution is very sensitive to noise, not stable, and linked with high computational costs. In this thesis, we propose and demonstrate a new method to obtain spike train correlations by deconvolution of the correlations of fluorescence signals, avoiding the reconstruction of spike trains altogether.

## 2 Materials and Methods

### 2.1 Software and hardware

This thesis proposes a new method to solve the "inverse problem" - get spike train correlation from recorded fluorescence signals. Implementation of the method, surrogate data modeling, and experimental data analysis were completed using Python (version 3.8) in Spyder (version 4.0.1) on a desktop with Intel Core i5-6500 CPU @ 3.2GHz and 16GB RAM. Experimental data (\*.tiff) was pre-processed in Suite2P (version 0.10.1) [14], exported as a python data file (\*.npy), and then loaded into Spyder for further processing. Simulation of neurons was carried out with NEST 3.0 [20].

### 2.2 Fluorescence signal modeling

With calcium imaging, fluorescence signal increases every time an action potential happens. The fluorescence signal can be taken as a function of action potential. We build a function  $f : x \rightarrow f(x)$  in which  $x$  is action potential of this neuron in time series and  $f(x)$  is fluorescence signal that can be observed. The function is to convolve a response kernel. In other words, fluorescence signal is modeled as a point process convolving a response kernel, represented below.

$$s = P * h, \quad (1)$$

where  $s$  is fluorescence signal;  $P$  stands for point process. A spike train generated by Poisson processes is one typical example, but it is not limited to it.  $h$  is the response kernel function. Two types of kernel functions were investigated: exponential decay

and beta function. Details are in Section 2.2.2. The numerical convolution simulation was implemented by the method of exact integration [21], which improves the computation speed enormously. Equation (1) and convolution theorem give in frequency domain

$$S(\phi) = \hat{P}(\phi) \hat{h}(\phi). \quad (2)$$

### 2.2.1 Point process

#### Poisson process

One realization of spike train can be described as

$$x(t) = \sum_i \delta(t - t_i),$$

where  $\delta$  is Dirac function and  $t_i$  is the spike timing. One of the most prevalent point processes to simulate stochastic spike trains is Poisson process, in which the spike count within a fixed time interval follows a Poisson distribution. Poisson distribution is a limit of a binomial distribution. Each spike in a Poissonian spike train is independent of the other. The Poisson process can be expressed as

$$P[N = k] = \frac{\Lambda^k}{k!} e^{-\Lambda}, \quad \Lambda = \lambda T,$$

where  $\lambda$  is the rate of spike train and  $T$  is the simulation time;  $N$  is total number of spikes;  $k$  is an integer [22]. Whether the rate of generated spike train is a constant or a function of time, there are stationary Poisson processes and non-stationary Poisson processes. Listed below (Algorithm 1 and 2) are the two simulation algorithms which were used in this thesis.

---

#### Algorithm 1 Simulation of stationary Poisson process [22]

---

**Require:**  $T \geq 0$

- 1: sample random number  $K$  from  $P(\lambda T)$
  - 2: distribute  $K$  spikes  $t$  uniformly and independently on  $[0, T]$
  - 3: spike train = collection of all points on time line  $[0, T]$
- 

Repeat simulation processes multiple times (Algorithm 1 and 2), we get parallel spike trains, which means spike trains are independent of each other. The cross-correlation and cross-covariance between spike trains are foreseen as zero. In addition, correlated

---

**Algorithm 2** Simulation of non-stationary Poisson process [22]**Require:**  $T \geq 0$ 

- 1: fix  $R$  with  $\lambda(t) \leq R$ , and sample random number  $K$  from  $P(RT)$ , where  $P(RT)$  is a Poisson distribution with parameter  $RT$
  - 2: distribute  $K$  points  $(t, r)$  uniformly and independently on  $[0, T] \times [0, R]$
  - 3: select points with  $r \leq \lambda(t)$ , discard all the others
  - 4: spike train = t-coordinates from the selected points
- 

homogeneous Poisson processes were also simulated. Single interaction process (SIP) and multiple interaction process (MIP) [23] are two examples of compound Poisson processes.

**SIP**

Given a realization of stationary Poisson process  $z(t)$  with rate  $\nu$  as a shared component, and another parallel realization of stationary Poisson processes  $x(t)$  with rate  $\lambda$  as a private component, single spike train  $m(t)$  is defined as

$$m(t) = x(t) + z(t).$$

As can be seen  $m(t)$  is a superposition of two Poissonian spike trains, its rate is therefore  $\lambda + \nu$ . The cross-covariance function of two spike trains simulated in the same way is  $\gamma(\Delta) = \nu\delta(\Delta)$  [24]. We see the cross-covariance is a delta peak with rate of shared component  $\nu$  as the amplitude.

---

**Algorithm 3** Simulation of SIP**Require:**  $T \geq 0$ 

- 1:  $x(t) \leftarrow P(\lambda T)$  ▷ Algorithm 1
  - 2:  $y(t) \leftarrow P(\mu T)$  ▷ Algorithm 1
  - 3:  $z(t) \leftarrow P(\nu T)$  ▷ Algorithm 1
  - 4:  $m(t) \leftarrow x(t) + z(t)$
  - 5:  $n(t) \leftarrow y(t) + z(t)$
- ▷  $m(t)$  and  $n(t)$  are two correlated Poissonian spike trains with common component  $z(t)$ . Rates of  $m(t)$  and  $n(t)$  are  $\lambda + \nu$  and  $\mu + \nu$  respectively. Cross-covariance  $\gamma_{mn}(\Delta) = \nu\delta(\Delta)$ .
-

## MIP

Given a realization of stationary Poisson process  $w(t)$  with rate  $\alpha$  as a mother process, individual spike train is produced by randomly and independently copying spikes from  $w(t)$  with a ratio  $\beta$ , which is formulated as

$$x_i(t) = \mathbf{Copy}(\beta)w(t).$$

The rate of each spike train is  $\alpha\beta$ . Cross-correlation function of two simulated spike trains is  $\rho(\Delta) = \alpha\beta^2\delta(\Delta) + (\alpha\beta)^2$ . Cross-covariance function of two simulated spike trains is  $\gamma(\Delta) = \alpha\beta^2\delta(\Delta)$  [24]. We see, both cross-correlation and cross-covariance of the simulated spike trains are delta peaks. The peak amplitude of can be quantified by the rate of mother process and the copy ratio.

---

### Algorithm 4 Simulation of MIP

---

**Require:**  $N$ : Integer, number of neurons

**Require:**  $T \geq 0$

**Require:**  $0 \leq \beta_i \leq 1$

- ```

1:  $w(t) \leftarrow P(\alpha T)$                                 ▷ Algorithm 1
2: for  $0 \leq i \leq N$  do
3:    $x_i(t) \leftarrow \mathbf{Copy}(\beta_i)w(t)$ 
4: end for

```
- 

## HOC

Spike train with a higher-order correlation structure was also tried in simulated data. Spike trains of neurons are compound Poisson processes. For  $n$  neurons, there are, in total,  $2^n - 1$  components. Each component is a stationary Poisson process. The realization of spike trains follow Algorithm 5. Details can be found in Staude et al. 2010 [25].

### 2.2.2 Kernel

Two kernel functions were investigated in order to model fluorescence signals accurately: exponential kernel function and beta kernel function, see Figure 8.

Exponential function is defined as:

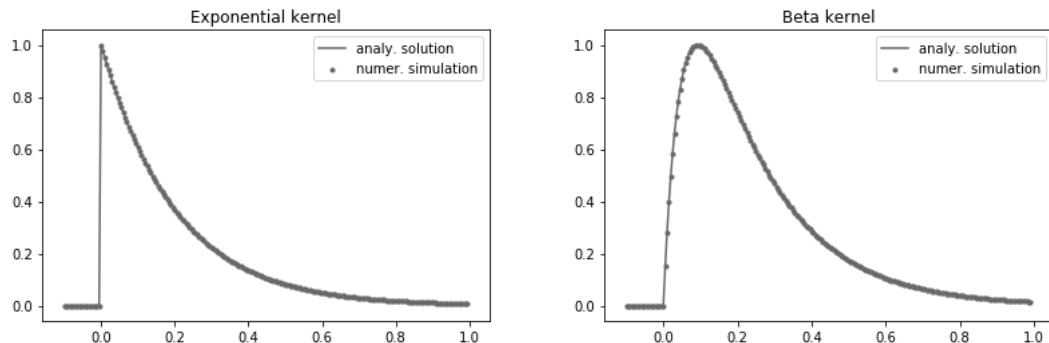
$$h(t) = \theta(t) C e^{-at}$$

**Algorithm 5** Construction of HOC via component rates**Input:** list of component rates  $r$ , simulation time  $T$ **Output:** correlated spike trains

```

1: if length of  $r \neq 2^N - 1$  then
2:   print "The number of spike trains should be  $2^N - 1$ .  $N$  is the number of
   neurons."
3: else
4:   Stimulate Poissonian spike trains for each rate independently.      ▷
   Algorithm 1
5:   Merge all spike trains into one carrier process.
6:   For each spike in the carrier process, creates an integer mark: index of the
   component + 1.
7:   Generate marked process as a list of two arrays: [carrier, integer mark].
8:   Change integer marks into binary marks.
9:   For each neuron, find where the binary mark is 1, and assign the correspond-
   ing spikes to that neuron.
10: end if

```



**Figure 8:** Response kernels. Left: exponential decay with  $a = 5$ . Analytical form is  $\theta(t)e^{(-5t)}$ . Right: beta function with  $a = 5, b = 20$ . Analytical form is  $(e^{-5t} - e^{-20t})/(e^{(-1/5)\ln 4} - e^{(-4/3)\ln 4})$ .

where  $\theta(t)$  is the Heaviside function that sets the value to zero for negative time arguments.  $C$  controls the amplitude of the kernel function. In this thesis,  $C = a$ , which normalizes the area under the curve as 1.  $1/a$  is the time constant of fluorescence decay. Smaller  $a$ , longer decay.

While the exponential function has only one parameter  $a$ , which controls the signal decay time, another kernel - beta kernel function has two parameters that describe rising and decay time, respectively. Beta function is given as

$$h(t) = \rho(e^{-at} - e^{-bt}),$$

in which  $a$  and  $b$  determine the shape of the function. Of the two values,  $1/a$  and  $1/b$ , the smaller is rising time constant, while the larger is decay time constant.  $\rho$  is a coefficient to normalize the amplitude of the kernel into 1, which can be represented through  $a$  and  $b$ :

$$\rho = \frac{1}{e^{\frac{a}{b-a} \ln \frac{a}{b}} - e^{\frac{b}{b-a} \ln \frac{a}{b}}}.$$

In the implementation of exact integration, the pulse train input was scaled with  $\rho(b - a)$  so that the amplitude of the kernel function is 1.

### 2.2.3 Exact integration as fast convolution

Instead of convolution, exact integration [21] was applied to simulate surrogate fluorescence signals. The solution of a first-order linear differential equation  $\dot{y} = Ay + x$  is  $y(t) = e^{A(t-s)} y(s) + \int_{s+}^t e^{A(t-\tau)} x(\tau) d\tau$ . In discrete time, with a pulse train as an input, it can be transformed as

$$y_{k+1} = e^{A\Delta} y_k + x_{k+1} \quad (3)$$

where  $\Delta$  denotes time step size, and matrix exponential  $e^{A\Delta}$  is called propagator, a fixed numerical matrix to transform data from one step to the next step. Propagator of exponential kernel is a single value  $e^{A\Delta} = e^{-a\Delta}$ . Propagator of the beta kernel is given

$$e^{A\Delta} = \begin{bmatrix} e^{-a\Delta} & 0 \\ \frac{1}{b-a}(e^{-a\Delta} - e^{-b\Delta}) & e^{-b\Delta} \end{bmatrix}.$$

From Equation (3), we see the output of the current step depends on the input of the current step and the output of the last step. An original paper graph (Figure 9)

clearly illustrates the algorithm. Exact integration is proved as a very efficient and accurate method to numerically simulate convolution operation.

$$\begin{array}{ccccccc} & x_1 & & x_2 & & x_3 & \dots \\ & \downarrow & & \downarrow & & \downarrow & \\ y_0 & \rightarrow & y_1 & \rightarrow & y_2 & \rightarrow & y_3 \rightarrow \dots \end{array}$$

**Figure 9:** Exact integration workflow adapted from [21]. Input  $x$  and output  $y$  of each simulation step on a discrete-time grid, with step size  $\Delta$ .  $y_0$  is the initial value. In each step, the propagator  $e^{A\Delta}$  times the input of the current step plus the output from the last step yields the output of the current step. The propagator is a fixed numerical matrix, which is calculated only once.

## 2.3 Signal characterization and correlation inference

### 2.3.1 One signal

Several statistical measurements can be used to characterize a signal, including mean value, variance, auto-correlation, auto-covariance in the time domain, and power spectral density in the frequency domain. Provided stationary signal  $x(t)$ , *mean*  $\mu$  and *variance*  $\sigma^2$  are defined as

$$\mu = \mathbb{E}[x(t)], \quad \sigma^2 = \mathbb{E}[x(t)^2] - \mathbb{E}[x(t)]^2,$$

through which both  $\mu$  and  $\sigma$  are generated as numbers. The mean and the variance do not change over time.

*auto-correlation*  $\rho_{xx}$  and *auto-covariance*  $\gamma_{xx}$  are defined as

$$\rho_{xx} = \mathbb{E}[x(t + \Delta)x(t)], \quad \gamma_{xx} = \mathbb{E}[x(t + \Delta)x(t)] - \mathbb{E}[x(t + \Delta)]\mathbb{E}[x(t)].$$

In the case of stationary signal, both auto-correlation and auto-covariance are independent of time. They are functions of lag time  $\Delta$  and symmetric.

$$\rho_{xx}(\Delta) = \rho_{xx}(-\Delta), \quad \gamma_{xx}(\Delta) = \gamma_{xx}(-\Delta)$$

Transforming auto-correlation into the frequency domain, we get *power spectral density* (PSD) of the signal. Bringing Equation (2) to the formula, PSD can be derived

as follow:

$$\pi_{xx}(\phi) = \mathbb{E}[S_x(\phi) S_x(\phi)^*] = \mathbb{E}[(\hat{P}_x \hat{h})(\hat{P}_x \hat{h})^*] = \mathbb{E}[\hat{P}_x \hat{P}_x^*](\hat{h} \hat{h}^*) = \lambda (\hat{h} \hat{h}^*),$$

where  $\pi_{xx}(\phi)$  denotes PSD of signal  $x(t)$ ;  $S_x(\phi)$  corresponds to signal in the frequency domain; superscript  $*$  denotes the complex conjugate of corresponding variables;  $\hat{P}_x$  stands for the underlying spike train in the frequency domain;  $h$  is the kernel function and  $\hat{h} \hat{h}^*$  is power spectral density (PSD) of the kernel;  $\lambda$  is a scaling factor. Because the point process can be expressed as the Dirac function in time series and in the frequency domain Dirac function is transformed as a constant, the power spectral density of a signal equals the scaled power spectral density of the corresponding kernel. This enables, in the case of an unknown kernel, the PSD of kernels can be represented by the PSD of the signal. What's more, according to Wiener–Khinchin theorem, the Fourier transform of auto-correlation of signal in time series is the PSD of the signal. By performing inverse Fourier transform of the PSD, it yields signal auto-correlation in the time domain:

$$\rho_{xx}(\Delta) = \check{\pi}_{xx}(\phi),$$

where  $\check{\pi}$  denotes inverse Fourier transform.

### 2.3.2 Two signals

#### Correlation of signals

Two signals are jointly stationary, if both of them are stationary individually, and their joint moments, in particular their pair-wise correlation, does not depend on (wall clock) time. Throughout this thesis, we always take joint stationarity for granted, without further discussion. Some results, however, could be reformulated to hold also for more general cases. Given two jointly stationary signals  $x(t)$  and  $y(t)$ , the *cross-correlation* between times  $t$  and  $t + \Delta$  ( $\Delta$  is a time shift) is defined as

$$\rho_{xy}(\Delta) = \mathbb{E}[x(t + \Delta) y(t)].$$

The *cross-covariance* of  $x(t)$  and  $y(t)$  between time  $t$  and  $t + \Delta$  is defined as

$$\gamma_{xy}(\Delta) = \mathbb{E}[x(t + \Delta) y(t)] - \mathbb{E}[x(t + \Delta)] \mathbb{E}[y(t)].$$

Joint stationarity readily implies an important symmetry of correlations

$$\rho_{xy}(\Delta) = \rho_{yx}(-\Delta) \quad \text{and} \quad \gamma_{xy}(\Delta) = \gamma_{yx}(-\Delta).$$

The expectation  $\mathbb{E}[\cdot]$  in the definition of the correlation function stands for an ensemble expectation. Although it could be estimated from repeated trials, it is common to replace it by time averaging. For jointly stationary signals this can be safely done, as the value of the product at different times  $t$  represent different samples with the same mean. For signals of finite duration  $T$ , the time averaging must take the effective signal overlap  $T - \Delta$  into account. So we practically estimate correlations from signals of length  $T$  as

$$\begin{aligned}\rho_{xy}(\Delta) &= \frac{1}{T - \Delta} \int_0^{T-\Delta} x(t + \Delta)y(t) dt, \\ \rho_{yx}(\Delta) &= \frac{1}{T - \Delta} \int_{\Delta}^T x(t - \Delta)y(t) dt.\end{aligned}$$

A discrete version is

$$\begin{aligned}\rho_{xy}(k) &= \frac{1}{N - k} \sum_{n=1}^{N-k} x(n + k)y(n), \\ \rho_{yx}(k) &= \frac{1}{N - k} \sum_{n=k+1}^N x(n - k)y(n).\end{aligned}$$

With recorded calcium signals, we target the correlations of underlying spike trains. Unlike conventional methods to do deconvolution, we solve the inverse problem from the model mathematically to infer the cross-correlation of spike trains directly from calcium signals. Here we consider two situations: neurons have the same kernel function, and neurons have their separate kernel functions.

### Same kernel for neurons

We start with calculating *cross-spectral density* (CSD) of two signals  $x(t)$  and  $y(t)$  in frequency domain. According to the definition, CSD equals the expectation of

the multiplication of Fourier transform of  $x(t)$  and complex conjugate of  $y(t)$  in the frequency domain. Bringing Equation (2) into it, we get

$$\pi_{xy}(\phi) = \mathbb{E}\left[S_x(\phi) S_y(\phi)^*\right] = \mathbb{E}[(\hat{P}_x \hat{h})(\hat{P}_y \hat{h})^*] = \mathbb{E}[\hat{P}_x \hat{P}_y^*](\hat{h} \hat{h}^*) \quad (4)$$

where  $\pi_{xy}(\phi)$  denotes CSD of two signals  $x(t)$  and  $y(t)$ ;  $S_x(\phi)$  and  $S_y(\phi)$  correspond to signals  $x(t)$  and  $y(t)$  in frequency domain; superscript \* denotes complex conjugate of corresponding variables;  $\hat{P}_x$  and  $\hat{P}_y$  stand for the underlying spike trains of  $x(t)$  and  $y(t)$  respectively;  $h$  is the common kernel function and  $\hat{h} \hat{h}^*$  is power spectral density (PSD) of the kernel.

Inverse Fourier transformation of CSD yields in the time domain:

$$\rho_{xy} = \rho_{P_x P_y} * \rho_{hh}, \quad (5)$$

where  $\rho_{xy}$  is cross-correlation of  $x(t)$  and  $y(t)$  in time domain;  $\rho_{P_x P_y}$  is cross-correlation of spike train underlying in  $x(t)$  and  $y(t)$  and  $\rho_{hh}$  is auto-correlation of the kernel function, defined as

$$\rho_{hh}(\Delta) = \int k(t + \Delta)k(t)dt.$$

Same as above, we can get cross-covariance:

$$\gamma_{xy} = \gamma_{P_x P_y} * \rho_{hh} \quad (6)$$

where  $\gamma_{xy}$  is cross-covariance of  $x(t)$  and  $y(t)$  in time domain;  $\gamma_{P_x P_y}$  is cross-covariance of underlying spike trains and  $\rho_{hh}$  is auto-correlation of the kernel function.

With calcium imaging, fluctuations of fluorescence signals for each neuron can be directly measured, with which we can calculate the cross-correlation of the signals. Cross-correlation of the underlying spike trains is of interest. The traditional method is to first extract each spike train by deconvolving the fluorescence signal. Here instead of performing deconvolution, we bring components in Equation (5) into the frequency domain. With known CSD of the signals and calculated PSD of the kernel (Section 2.2.2), we get CSD of underlying spike trains. Transforming it into the time domain, we get the cross-correlation of the spike trains.

### Different kernels for neurons

In Section 2.3.2, we inferred the cross-correlation of spike trains by assuming all neurons share the same response function. In experimental data, it is noticed that the decay time varies from long (10 seconds or more) to short (2-3 seconds). Therefore it makes sense to customize different kernels for each neuron. Although different kernels, the inference of cross correlation is similar. Taking two signals  $x(t)$  and  $y(t)$ , with their own kernels  $h_x$  and  $h_y$  and Point processes  $P_x$  and  $P_y$ . According to Equation (2), we get  $x(t)$  and  $y(t)$  in frequency domain:

$$S_x = \hat{P}_x \hat{h}_x \quad S_y = \hat{P}_y \hat{h}_y.$$

The CSD of two fluorescence signals is then:

$$\pi_{xy}(\phi) = \mathbb{E}[S_x(\phi) S_y(\phi)^*] = \mathbb{E}[(\hat{P}_x \hat{h}_x)(\hat{P}_y \hat{h}_y)^*] = \mathbb{E}[\hat{P}_x \hat{P}_y^*](\hat{h}_x \hat{h}_y^*)$$

Inverse Fourier transformation yields in the time domain:

$$\rho_{xy} = \rho_{P_x P_y} * \rho_{h_x h_y}, \quad \gamma_{xy} = \gamma_{P_x P_y} * \rho_{h_x h_y},$$

where  $\rho$  means cross-correlation,  $\gamma$  means cross-covariance,  $P$  is the point process, standing for underlying spike trains and  $h$  is a kernel function,  $*$  means convolution operation.

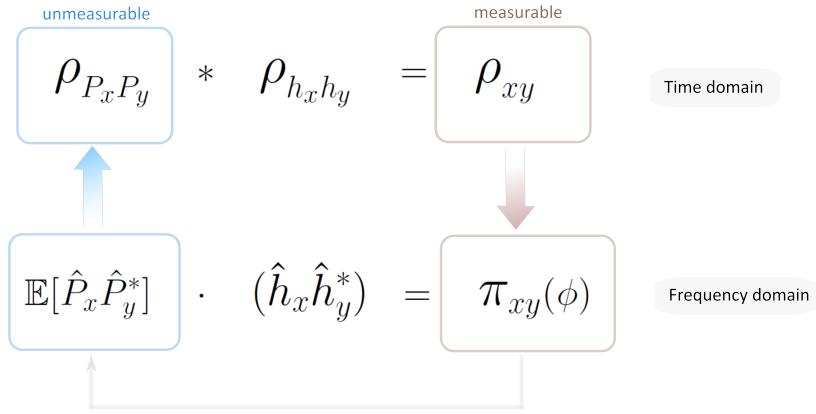
To summarize, in the time domain we have

$$\rho_{P_x P_y} * \rho_{h_x h_y} = \rho_{xy},$$

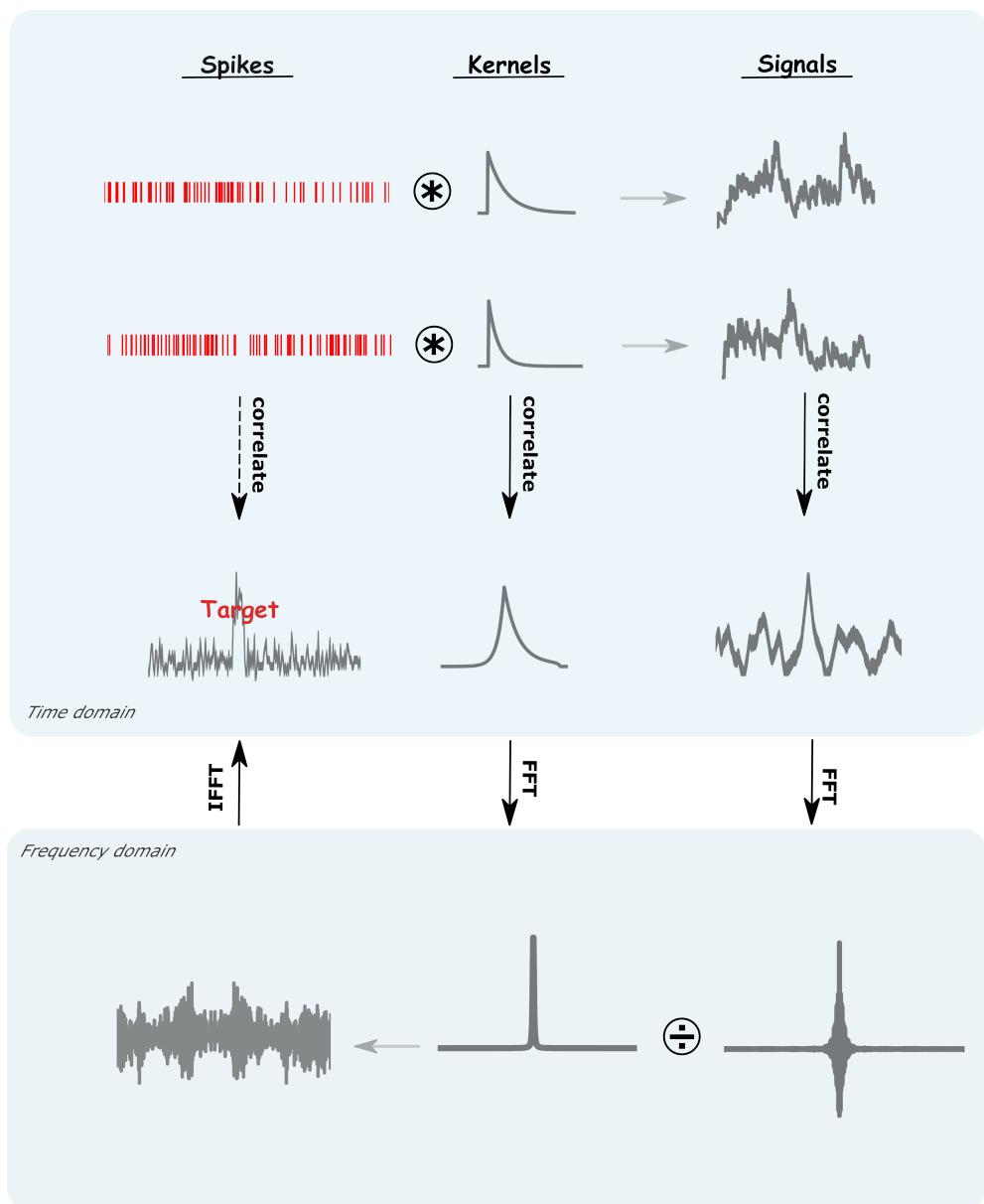
and correspondingly in frequency domain

$$\mathbb{E}[\hat{P}_x \hat{P}_y^*] \cdot (\hat{h}_x \hat{h}_y^*) = \pi_{xy}(\phi),$$

in which,  $\rho_{S_x S_y}$  can be calculated.  $\rho_{P_x P_x}$  is our goal. Dividing CSD of signals by CSD of kernels and transforming it into time domain is the cross-correlation of underlying spike trains. The process is visualized in Figure 10. By performing calculations in the frequency domain, direct deconvolution can be avoided. Deconvolution in the time domain can be implemented by spectrum division in the frequency domain. A graphic demonstration is shown in Figure 11.



**Figure 10:** Method overview in case that single neuron has its own kernel function. In time domain,  $\rho_{xy}$ : cross-correlation of  $x(t)$  and  $y(t)$ ;  $\rho_{P_x P_y}$ : cross-correlation of underlying spike trains;  $\rho_{h_x h_y}$ : cross-correlation of the kernel function;  $*$ : convolution operation. In frequency domain,  $\pi_{xy}(\phi)$ : CSD of two signals  $x(t)$  and  $y(t)$ ;  $\hat{h}_x \hat{h}_y^*$ : cross-spectral density of kernel function;  $\hat{P}_x \hat{P}_y$ : CSD of underlying spike trains. In the time domain, cross-correlation of underlying spike trains is of our interest; however not measurable with calcium imaging techniques. What can be measured is fluorescence signals. The cross-correlation of two signals can be, therefore, directly calculated. By transforming all components into the frequency domain, the CSD of signals can be obtained. By dividing the CSD of signals by the PSD of the kernel, we get the CSD of spike trains. Through inverse Fourier transform of CSD of spike trains, we get the cross-correlation of spike trains in the time domain.



**Figure 11:** Deconvolution in the time domain as division in the frequency domain. Fluorescence signals are modeled as a spike train convolving with a kernel function. The correlation of underlying spike trains is the target. Instead of performing deconvolution, correlations of kernels and signals are transformed into the frequency domain. The power spectral density of signals divides the power spectral density of kernels, yielding the power spectral density of underlying spike trains. Bringing it into the time domain, we get the target cross-correlation.

### 2.3.3 Multiple signals

Let  $\mathbf{P}$  denote a vector of point processes of  $n$  neurons. Each neuron has a length of  $m$  numerical data points.  $\mathbf{h}$  denote a column vector of kernels with  $m$  data points for  $n$  neuron. The  $n$  signals in time series can be expressed as

$$\mathbf{s} = \mathbf{P} * \mathbf{h}.$$

Correspondingly in frequency domain:

$$\mathbf{S}(\phi) = \hat{\mathbf{P}}\hat{\mathbf{h}}.$$

The cross-spectral density of multiple fluorescence signals can be expressed in matrix [26]:

$$\pi_{1..n}(\phi) = \begin{pmatrix} \pi_{11} & \pi_{12} & \dots & \pi_{1n} \\ \pi_{21} & \pi_{22} & \dots & \pi_{2n} \\ \dots & \dots & \dots & \dots \\ \pi_{n1} & \pi_{n2} & \dots & \pi_{nn} \end{pmatrix},$$

in which elements on the diagonal are PSDs of each signal. Off-diagonal elements are CSDs of all paired signals.

### 2.3.4 Modeling membrane potential

Besides handling fluorescence data, we consider the possibility of applying the method in analyzing sub-threshold membrane potential. The neuronal membrane can be taken as a low-pass filter that filters pre-synaptic activity. It generates a shot noise process: convolving post-synaptic current (PSC) with an impulse response of membrane yields membrane potential activity in the sub-threshold regime. Given neurons  $x$  and  $y$ , we then have:

$$U = M * I, \quad \gamma_{U_x U_y} = \gamma_{I_x I_y} * \rho_{MM},$$

where  $U$  is membrane potential;  $M$  is impulse response of membrane;  $I$  is PSC;  $\gamma_{U_x U_y}$  is cross-covariance of the membrane potential;  $\gamma_{I_x I_y}$  is cross-covariance of PSC of the two neurons, and  $\rho_{MM}$  is the auto-correlation of impulse response of

membrane.

Same as in Section 2.3.2, it yields in frequency domain:

$$\mathbb{E}[\hat{I}_x \hat{I}_y^*] (\hat{M} \hat{M}^*) = \pi_{U_x U_y}(\phi) \Rightarrow \mathbb{E}[\hat{I}_x \hat{I}_y^*] = \frac{\pi_{U_x U_y}(\phi)}{(\hat{M} \hat{M}^*)}.$$

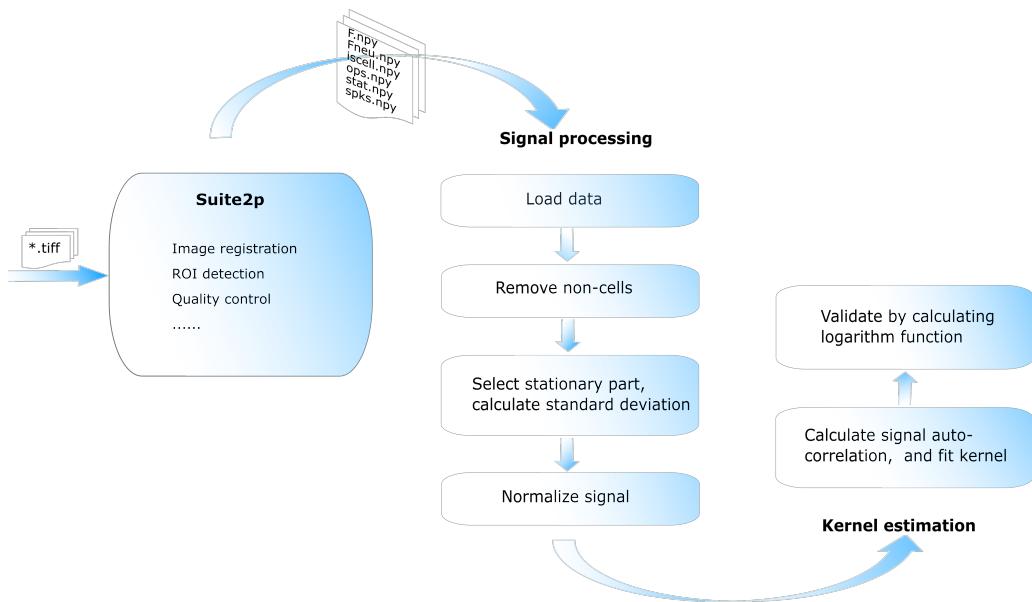
Bring  $\mathbb{E}[\hat{I}_x \hat{I}_y^*]$  to time domain, we can get the cross-covariance of PSC ( $\gamma_{I_x I_y}$ ).

To be emphasized is that this methodology applies only to linear neuron model, such as leaky integrate-and-fire model, not for other neuron models such as conductance-based neuron model or adaptive exponential integrate-and-fire neuron model. Besides, the methodology applies to membrane potential exceptionally in the sub-threshold regime, meaning the neurons do not fire.

## 2.4 Pre-processing of experimental data

Apart from simulated data, the method was also tried on experimental data. Experimental data was provided by Prof. Björn Kampa's group from RWTH Aachen University. Experiments were performed with 10-20 weeks old transgenic mice with cranial windows implanted on V1. Images were obtained using two-photon calcium imaging of spontaneous activity recorded in GCaMP6s-labeled neurons in V1 layer 2/3 of while mice were head fixed on a light-weight running wheel. No stimulus was presented. The microscope had a frame rate of 30 Hz. Pre-processing procedures were performed beforehand as summarized in Figure 12. Image stacks were loaded into Suite2p. Several operations were performed by Suite2p: image registration via phase correlation which corrects for the effects of brain movement with a rigid alignment algorithm; ROI detection, which generates identified ROIs with positive weights for each pixel; ROI labeling and quality control. Several \*.npy files are generated by Suite2p and are loaded into Python for further handling. In Python, ROIs that are not identified as cells were eliminated. Signals are normalized by subtracting the global mean and dividing by the standard deviation of a stationary part in the signal. In the case of experimental data, the kernel is unknown. As the auto-correlation of one signal is a bi-exponential function with the same time constant as the kernel, it can extract the kernel function from the auto-correlation of the signal. Therefore next step is to calculate the auto-correlation of signals and fit an exponential function into the data of half-signal auto-correlation. The fitting process is performed by method `scipy.optimize.curve_fit`, which applies non-linear

least squares to fit an exponential function and uses the Levenberg-Marquardt algorithm to solve the minimization problem. It yields estimated values of parameters and variances of the estimate as the goodness of fitting. To validate fitting process further, the natural logarithm of fluorescence auto-covariance was calculated, which is expected as a linear function.



**Figure 12:** Pre-processing of experimental data. \*.tiff are image stacks recorded from experiments. Suite2p outputs six python data files. F.npy: fluorescence traces. Fneu.npy: neuropil fluorescence traces. iscell: mark each ROIs whether it is a valid cell and the probability of that. This can be calibrated manually from Suite2p GUI. ops.npy: options and intermediate outputs. stat.npy: list of statistics computed for each cell. spks.npy: results of spike inference in Suite2p. Files used in this thesis are F.npy and iscell.npy.

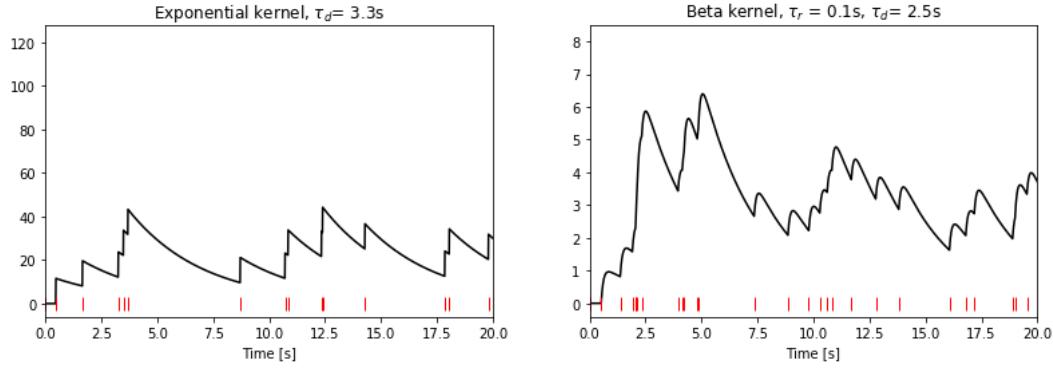
# 3 Results

## 3.1 Simulated fluorescence data

The fluorescence signal is modeled as a low pass filter of the underlying spike train. Synthetic fluorescence data was accordingly simulated, of which the underlying spike trains are known. Simulated data enables possibilities to pre-define the properties of underlying spike trains. Different correlation structures were attempted, and the estimated spike train was compared with ground truth.

### 3.1.1 Kernel selection

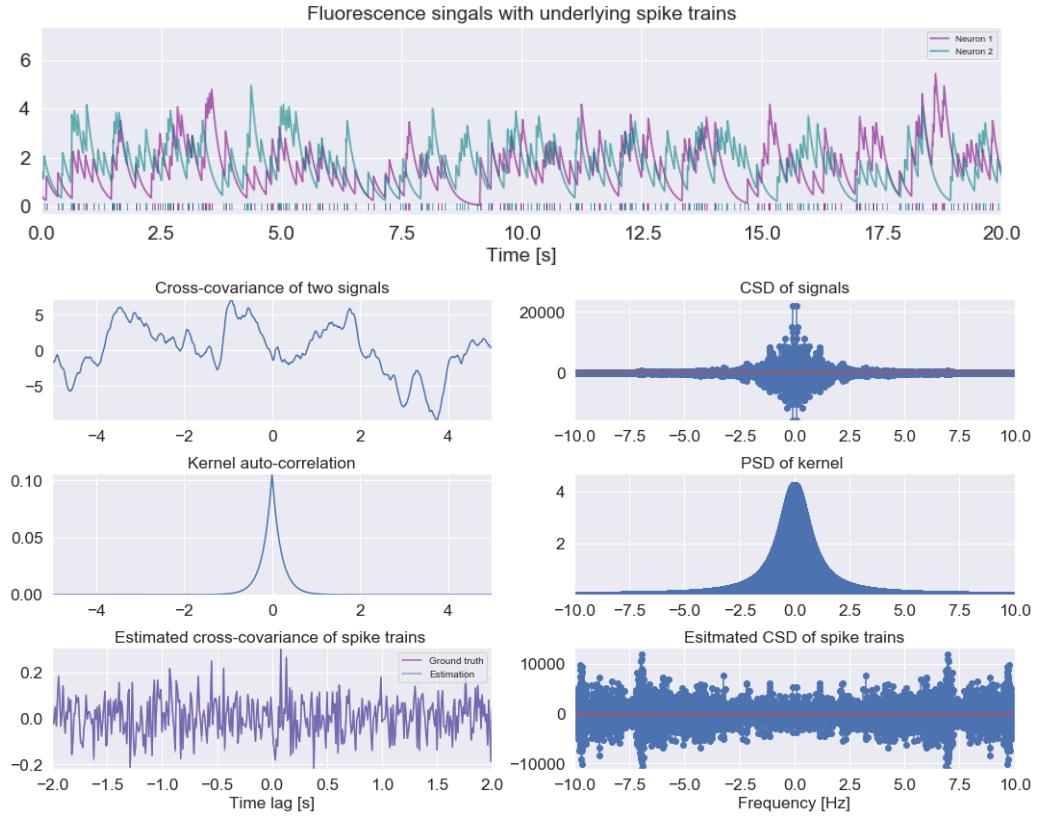
As mentioned in Section 2.2, exponential decay and beta function were considered to simulate fluorescence signals. Examples are shown in Figure 13. Referring to [6], exponential kernel is more akin to experimental data. All listed simulations and analyses of fluorescence signals throughout this thesis were therefore based on exponential kernels.



**Figure 13:** Synthetic fluorescence data with different kernel functions. Red markers at the bottom are underlying spike trains. Left: simulated with an exponential kernel function. The decay time constant is 3.33 s. The underlying spike train is generated via the Poisson process with a rate of 1 Hz. The simulation duration is 100 s, and the time step is 0.01 s. Right: simulated with beta function. The rising time constant is 0.1 s. The decay time constant is 2.5 s. A Poisson process yields the underlying spike train with a rate of 1 Hz. Simulation duration is 1000 s with time step 0.01 s.

### 3.1.2 Independent Poisson process

The first tried is the independent stationary Poisson process Section 2.2.1. Spike trains simulated with stationary Poisson process independently have no correlations, which is indicated by the fact that no significant peaks in the signal cross-covariance and true cross-covariance of spikes trains in Figure 14. By analyzing a pair of surrogate fluorescence signals generated by convolving spike trains with the exponential kernel, the estimated cross-covariance of spike trains matches the curve of ground truth completely around lag zero.

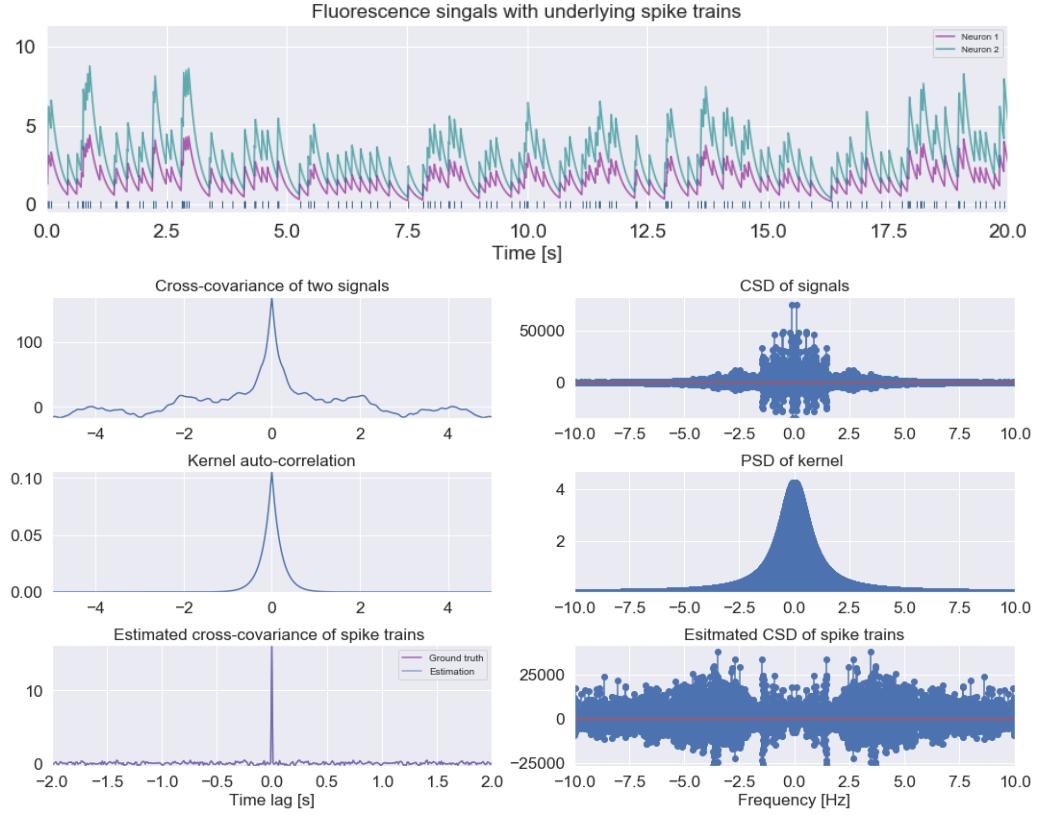


**Figure 14:** Fluorescence signals of two neurons with underlying parallel spike trains simulated via stationary Poisson process. The rate of input spike train is 8 Hz. Simulation is done by exact integration with exponential kernel [21]. The decay time constant of the kernel is 0.2 s. Simulation time is 100 s. The simulation time step is 0.01 s. The top panel are raw simulated signals of two neurons with underlying spike trains marked as small bars with same color at the bottom. Panels in the left column are signals in time domain. Panels in the right column are the corresponding power/cross spectrum of the signal in the same row.

### 3.1.3 Correlated Poisson process with double input

In contrast to the independent stationary Poisson process, cross-covariance of signals simulated with correlated stationary Poisson process Section 2.2.1 have an obvious peak at lag zero. One simplest way to create correlated spike trains is to simulate one Poissonian spike train and double the spike counts within the same time bin for the other. Consequently, one fluorescence signal has twice the amplitude of the other. In Figure Figure 15, the spiking rate of neuron 1 is 8 Hz, and of neuron 2

is 16 Hz. The estimated cross-covariance has a peak amplitude of 16.15, while the ground truth has a peak amplitude of 16.16. They are in line with each other.

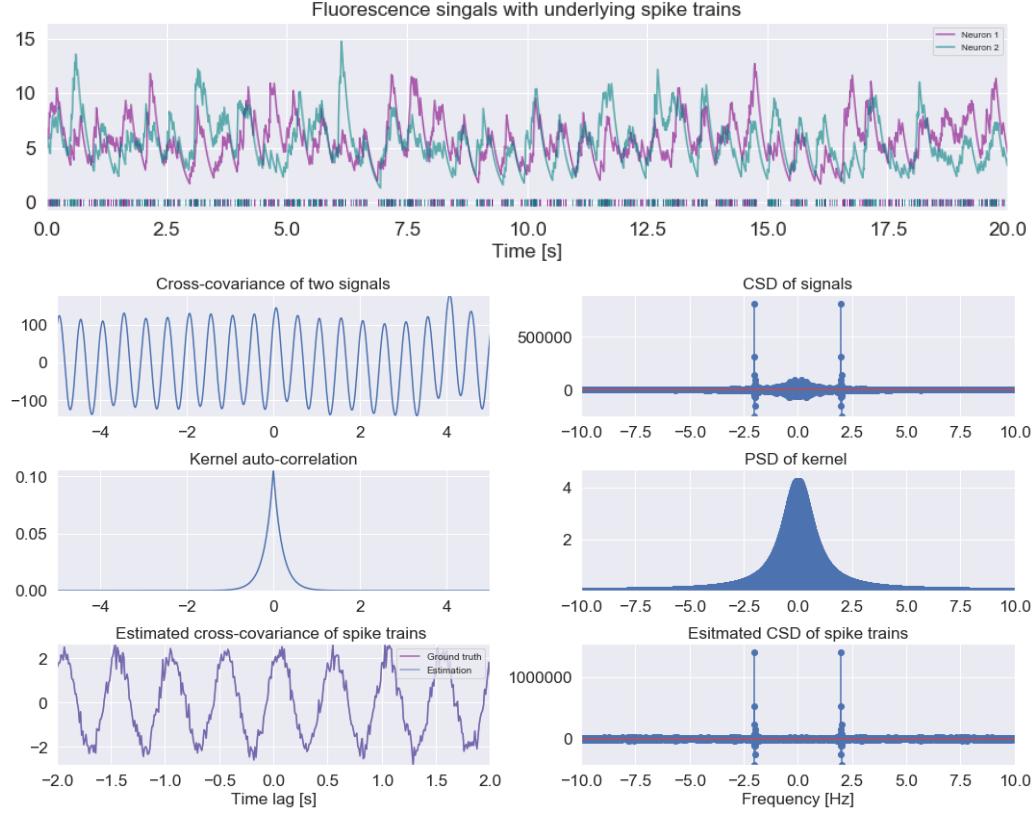


**Figure 15:** Fluorescence signals of two neurons with underlying spike trains simulated via homogeneous Poisson process. The rate of one input spike train is 8 Hz. Rate of another spike train is 16 Hz with doubled spike counts. Simulation is done by exact integration with exponential kernel [21]. The decay time constant of the kernel is 0.2 s. The simulation time is 100 s. The simulation time step is 0.01 s. The top panel are raw simulated signals of two neurons with underlying spike trains marked as small bars with same color at the bottom. Panels in the left column are signals in time domain. Panels in the right column are the corresponding power/cross spectrum of the signal in the same row.

### 3.1.4 Non-stationary Poisson process

With the non-stationary Poisson process, the spiking rate changes over time. One typical example is oscillation. In Figure 16, the cross-covariance of two simulated

fluorescence signals appears to oscillate, as well as the cross-covariance of spike trains. Prominent peaks in frequency domain correspond to the oscillation frequency in time domain. The estimated cross-covariance and ground truth have no deviations.

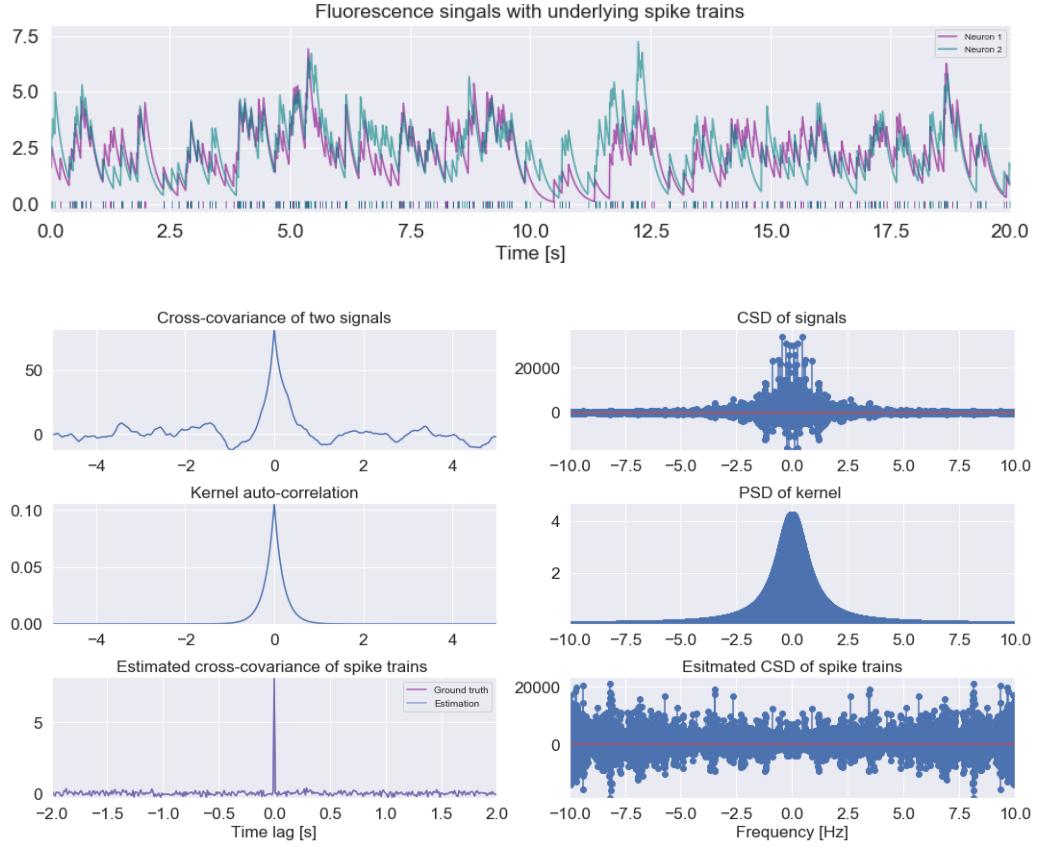


**Figure 16:** A: fluorescence signals of two neurons with underlying spike trains simulated via non-stationary Poisson process (Algorithm 2). Input spike rate follows  $27(1 + 0.8\sin(4\pi t))$ . Simulation is done by exact integration with exponential kernel [21]. The decay time constant of the kernel is 0.2 s. Simulation time is 100 s. The simulation time step is 0.01 s. The top panel are raw simulated signals of two neurons with underlying spike trains marked as small bars with same color at the bottom. Panels in the left column are signals in time domain. Panels in the right column are the corresponding power/cross spectrum of the signal in the same row.

### 3.1.5 SIP

Analyzing the fluorescence signals generated by SIP Section 2.2.1, the estimated cross-covariance of underlying spike trains and the ground truth are totally in agree-

ment around lag zero. The peak amplitude is 8.10 for estimation, and 8.15 for ground truth for the case in Figure 17.

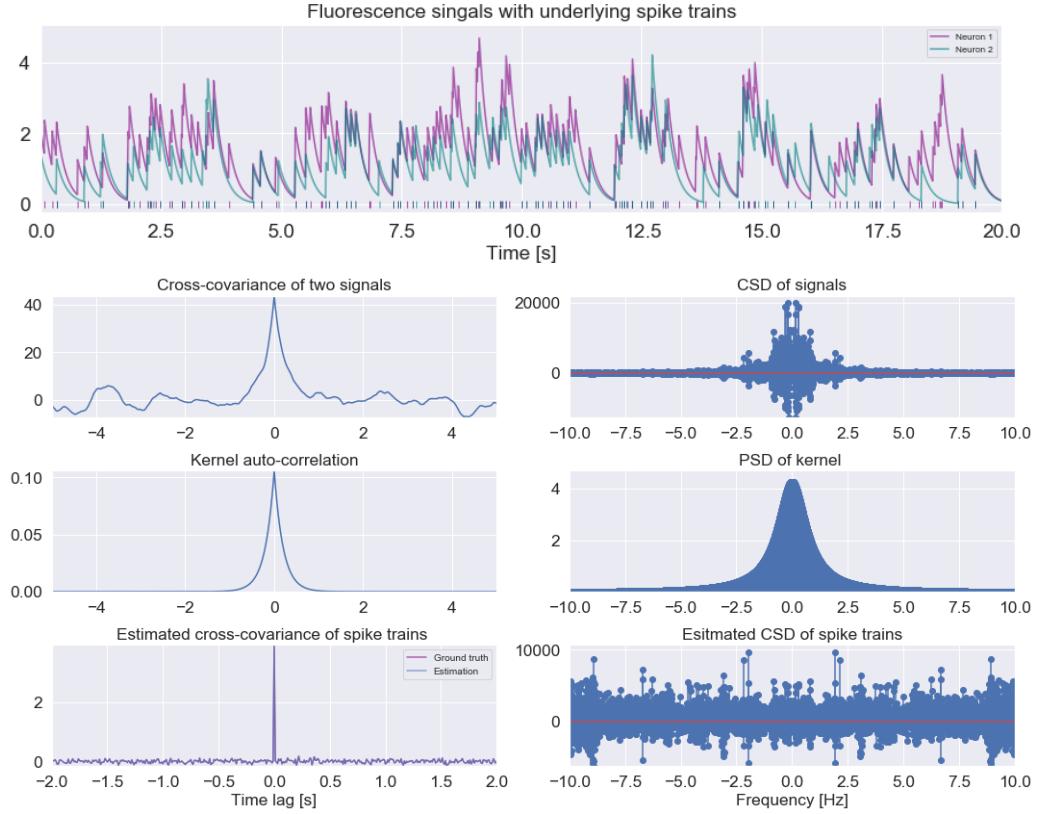


**Figure 17:** Fluorescence signals of two neurons with underlying spike trains are simulated via SIP process (Algorithm 3). The rate of a common spike train is 8 Hz. Rates of private rates are 3 Hz and 5 Hz. Simulation is done by exact integration with exponential kernel [21]. The decay time constant of the kernel is 0.2 s. Simulation time is 100 s. The simulation time step is 0.01 s. Panels from top to down are raw simulated signals, cross-covariance of two signals, auto-correlation of the exponential kernel, the estimated cross-covariance of underlying spike trains, and its ground truth.

### 3.1.6 MIP

Analyzing the fluorescence signals generated by MIP Section 2.2.1, the estimated cross-covariance of underlying spike trains and the ground truth are identical around

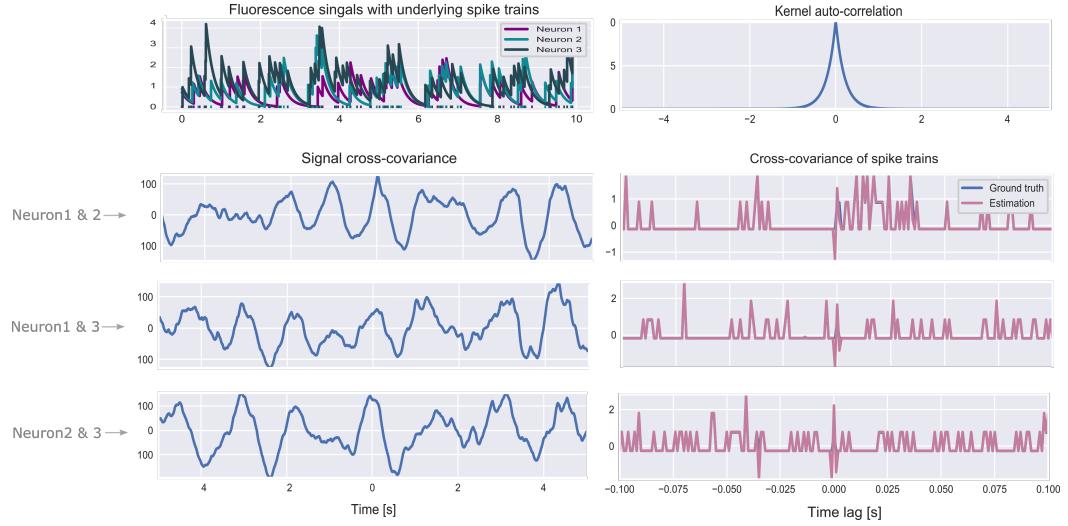
lag zero. The peak amplitude is 3.92 for estimation, and 3.93 for ground truth Figure 18. To be noticed, the cross-covariance of underlying spike trains is not deterministic a delta peak. Due to the random copy process, the shared spikes between two signals are stochastic.



**Figure 18:** A: fluorescence signals of two neurons with underlying spike trains simulated via MIP process (Algorithm 4). The rate of the mother spike train is 8 Hz. Copy ratios of simulated spike trains are 0.8 and 0.6. Simulation is done by exact integration with exponential kernel [21]. The decay time constant of the kernel is 0.2 s. Simulation time is 100 s. The simulation time step is 0.01 s. The top panel are raw simulated signals of two neurons with underlying spike trains marked as small bars with same color at the bottom. Panels in the left column are signals in time domain. Panels in the right column are the corresponding power/cross spectrum of the signal in the same row.

### 3.1.7 HOC

With HOC Section 2.2.1, three traces of fluorescence signals were generated. Cross-covariance estimation between each two signals are displayed in Figure 19. There is almost no deviation of estimated cross-covariance from the ground truth around lag zero.



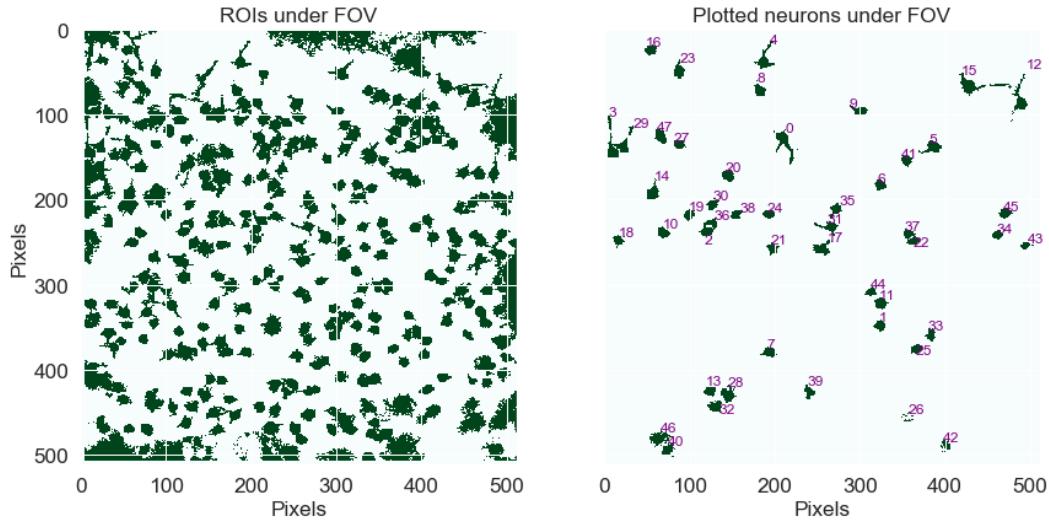
**Figure 19:** Fluorescence signals of three neurons with underlying spike trains in higher-order correlations. Component rates are [5, 1, 0.5, 0.2, 1, 2, 0.5] Hz respectively. Simulation is done by exact integration with exponential kernel [21]. The decay time constant of the kernel is 0.2 s. The simulation time is 10 seconds. The simulation time step is 0.001 seconds. The top left panel is raw simulated fluorescence signals of three neurons with underlying spike trains marked as small bars with the same color at the bottom. The top right panel is kernel auto-correlations. Each row below is an analysis of a pair of neurons. In the left column are cross-covariance of fluorescence signals. In the right column are the estimated cross-covariance and its ground truth.

Seen from above, by applying the method of cross-correlation inference on surrogate data, estimated cross-covariance is highly in agreement with ground truth. Whether it is an independent Poisson process or correlated Poisson process, simple correlation or complex correlation structure, second-order correlation or higher-order correlations, the method can always capture the correlation structure of underlying spike trains accurately. It is worth mentioning that the represented plots above in this section are about the accurate cross-covariance inference of spike trains from

surrogate data. In fact, the same fidelity goes for cross-correlation inference. This is because cross-correlation and cross-covariance have the same structure, only differ in an offset component that is generated by the multiplication of the global mean. This convinces us of the reliability of the method and attempts it on the experimental data.

### 3.2 Experimental data

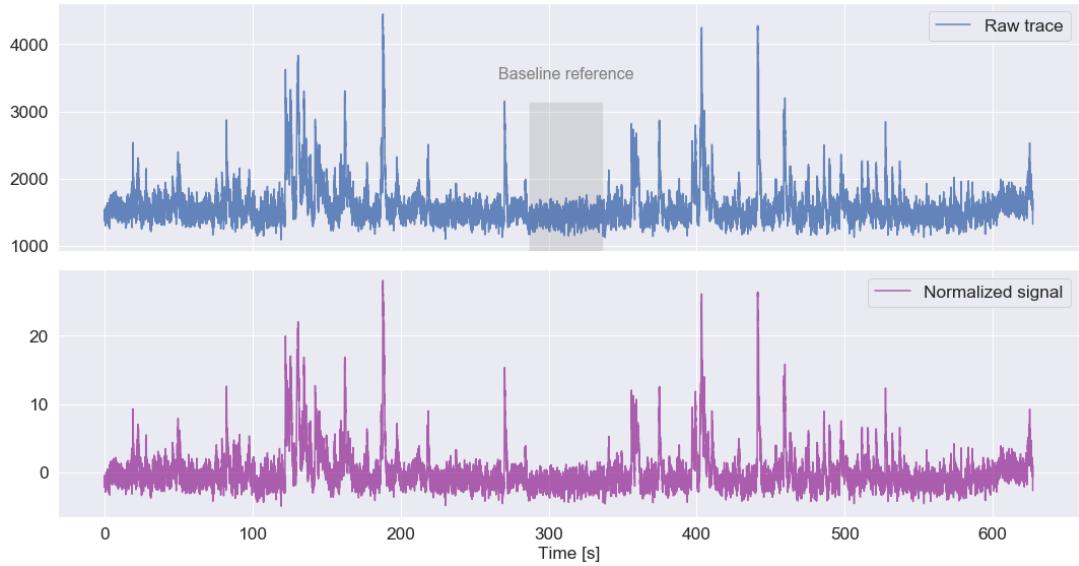
Analysis in this section is based on one dataset of spontaneous activity recorded by two-photon calcium imaging of GCaMP6s-labeled neurons in the mouse primary visual cortex. Data acquisition and pre-processing procedure are described in Section 2.4. For this dataset, 453 ROIs were defined by Suite2p, among which 303 regions were identified by valid cells, among which 48 were randomly selected into analysis. Figure 20 depicts the spatial distribution of the neurons under FOV.



**Figure 20:** Spatial information of neurons. Left: ROIs that identified as neurons under FOV. Dark green areas are where neuron(s) are located. Right: location of analyzed neurons marked with ids, which are corresponding to neuron ids in Figure 23.

### 3.2.1 Data normalization

After loading data generated from Suite2p, fluorescence traces of valid cells were normalized. A relatively stationary part where no obvious peaks were selected. Raw fluorescence trace subtracting its global mean and divided by the standard deviation of the baseline reference part yields the normalized signal. The amplitude of the normalized signal is decreased, and the baseline is around zero. Figure 21 gives an example.

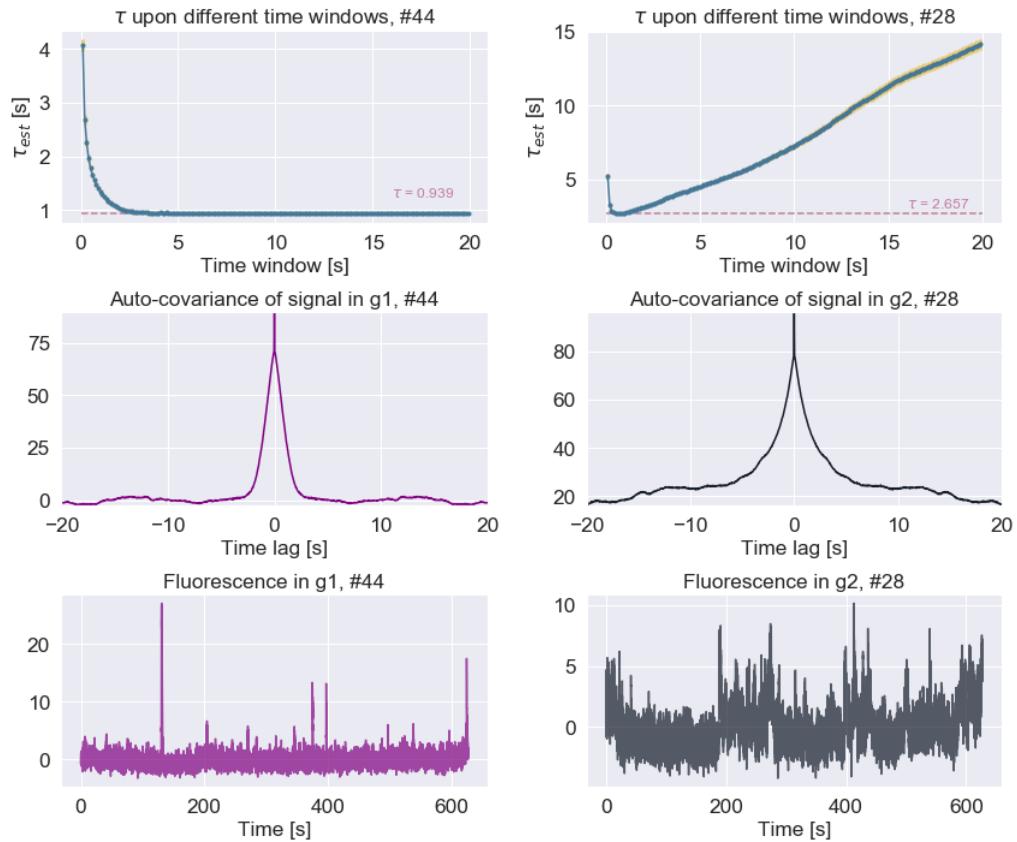


**Figure 21:** Raw fluorescence traces and normalized signal.

### 3.2.2 Kernel fitting

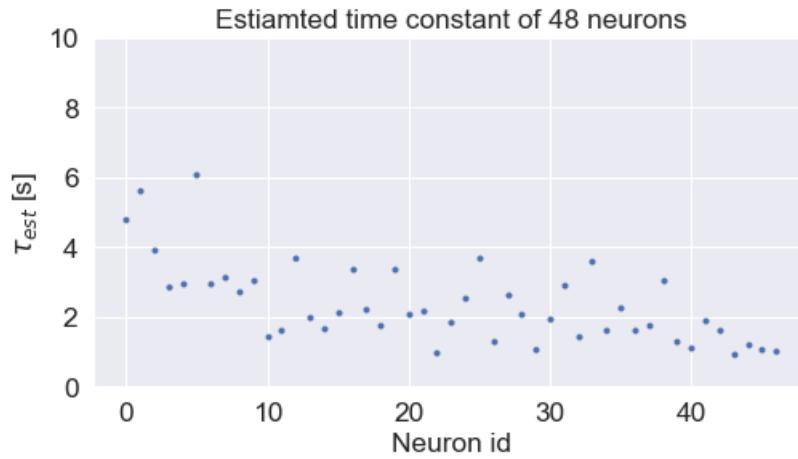
For experimental data, the kernel is unknown. It was first assumed that all neurons share the same response kernel-induced calcium change and use auto-correlation of either signal as a substitute for auto-correlation of the kernel. The generated results were not consistent and explainable. Inspecting data, decay time and amplitudes of peaks vary between neurons. It is, therefore, reasonable to estimate kernel function individually. Auto-correlation of normalized fluorescence trace was calculated. Based on this, an exponential function was fitted into half of the data upon a specific time window. As the fitting results are sensitive to window size, time windows size was varied from 0.1 to 20 seconds with step 0.1 seconds to see the trend of time

constant. The minimum value of the time constant was picked as the final value of the parameter in the kernel. It was noticed that the time constant curve upon different time windows tends to have different trends. In Figure 22 left side, the value of the estimated time constant stabilizes with increasing time window. In contrast, on the right side, it declines and then increases with a larger standard deviation error, which seems to have something to do with whether the auto-correlation of the signal has zero-crossing within the time window. Looking into more neurons, as seen from Figure 23, it is true that neurons' time constants are different, which aligns with the initial speculation.



**Figure 22:** Kernel fitting of experimental data.

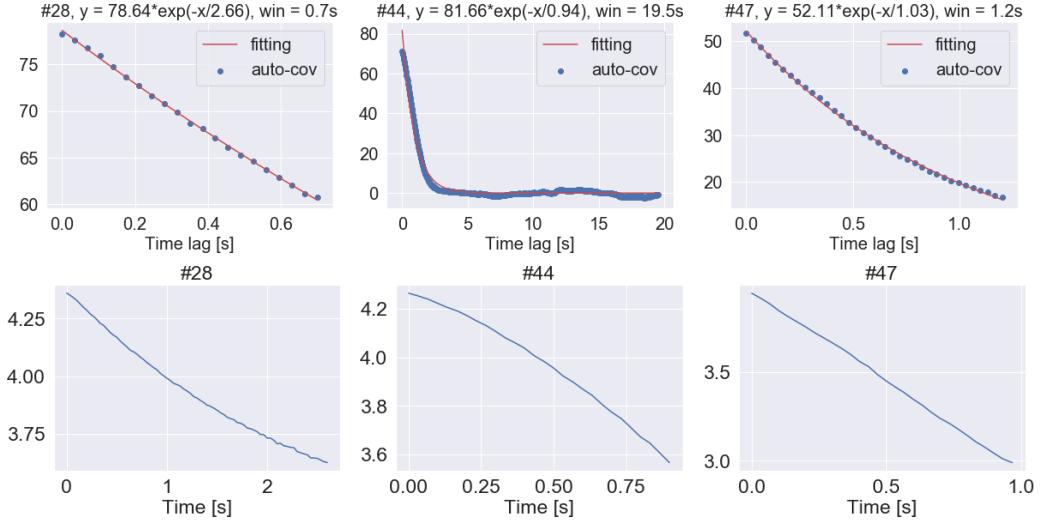
The final optimized time constants for the 48 neurons are presented in Figure 23. The shortest estimated time constant is 0.94 s, and the largest 6.06 s.



**Figure 23:** Time constants of different neurons

### 3.2.3 Validation of kernel fitting

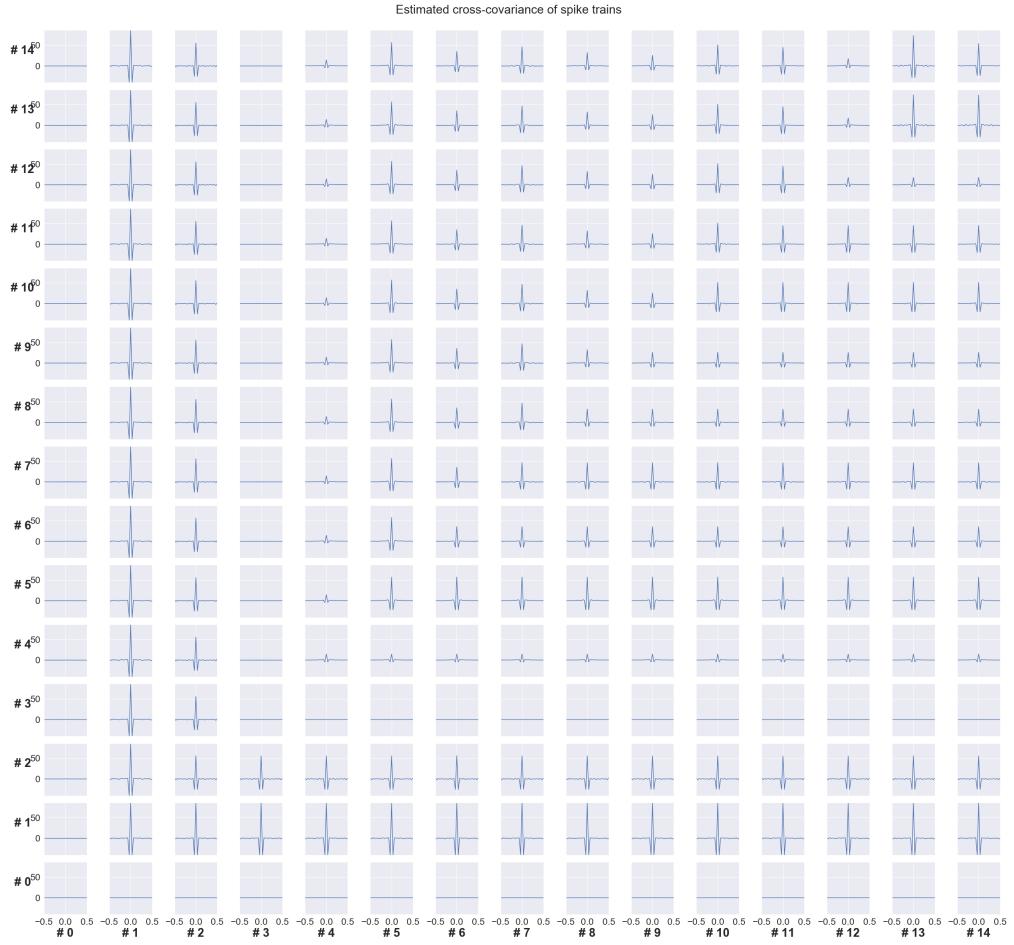
As ground truth is absent, we take advantage of the characteristics of the exponential kernel to validate how well the data points of auto-covariance follow an exponential function. Three examples are shown in lower panels in Figure 24, where we see the curve of neuron #47 is closest to a straight line. The lines of neurons #28 and #44 appear slightly curvy but not dramatic. These align with the actual fittings shown in the upper panels of Figure 24, where the optimized time constants were obtained. Neuron #28 has the smallest time window and least data points fitted in. Afterward the estimated time constant increases instead of stabilizing, as can be seen in Figure 23.



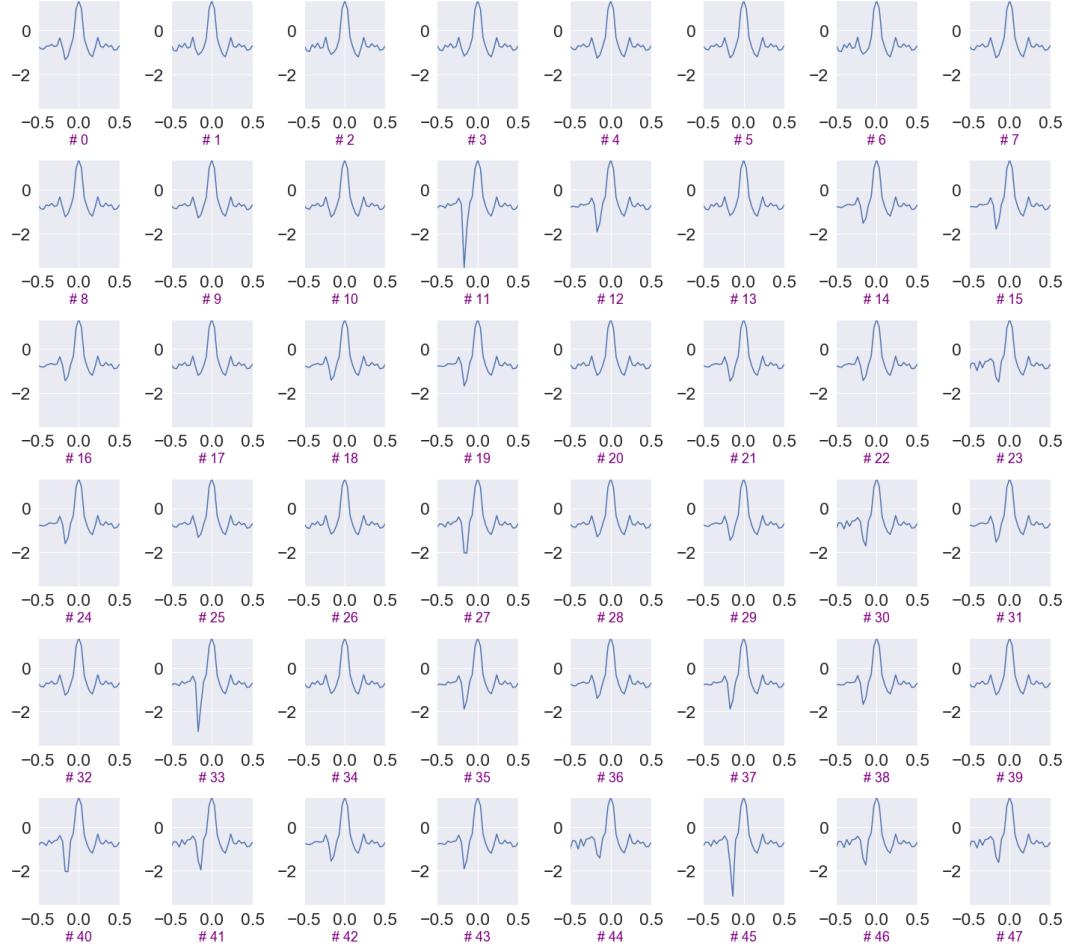
**Figure 24:** Upper row of panels: data fitting results where the optimized time constant was obtained. In the titles, the fitted functions were presented respectively. win: time window of the right side of the fluorescence auto-covariance, where the fitting happened. It corresponds to the range on the x-axis. Lower row of panels: results of taking natural logarithm on the fractions of adjacent points from the fluorescence auto-covariance. The time on the x-axis corresponds to the optimized time constant from each neuron.

### 3.2.4 Correlation analysis

Cross-covariance was estimated for multiple neurons, as seen in Figure 25 and Figure 26. The scale of the y-axis was unified to make the amplitude comparable between different neurons. Some neurons' auto-covariances are small, while some are quite large. High amplitudes have a dominant effect. Suspectly it could be due to noise components from experiments that make all the covariance look homogeneous.



**Figure 25:** Estimated covariance between multiple neurons. To ensure the clarity of reading, an analysis of 15 neurons is plotted on the grid.

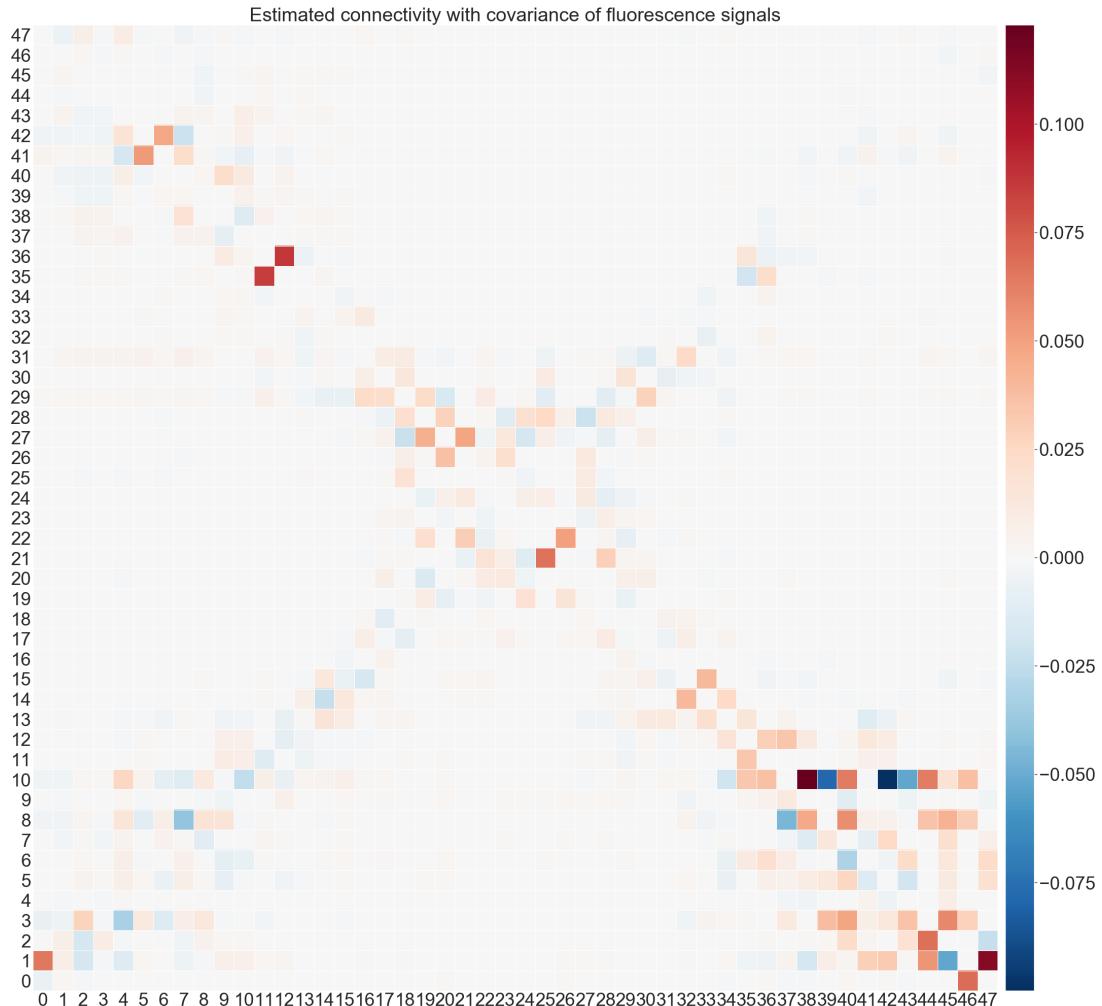


**Figure 26:** Estimated covariance of neuron #17 with all neurons. Amplitude on the y-axis is in logarithmic scale.

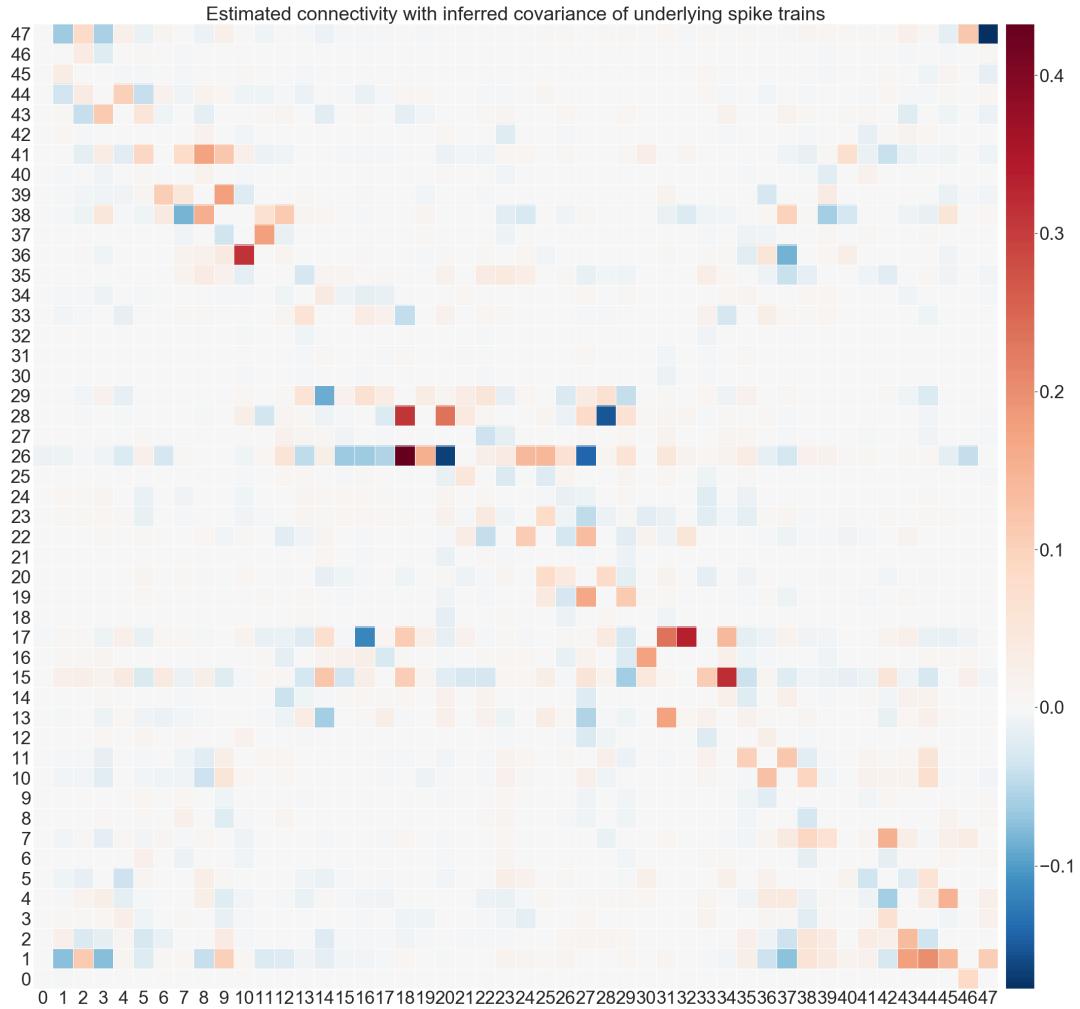
### Connectivity estimation

Schiefer et al. have developed a method to infer the directed connectivity of neurons based on the covariance matrix at time lag zero. An effective connectivity is estimated by exploiting the covariance matrix in a set of data based on a linear model. A sparse solution can be given via L1-minimization [27]. We attempted this method with experimental data. Connectivity matrix in Figure 27 was generated using covariance matrix from the fluorescence signals, in which along the diagonals connectives are relative visible. Figure 28 was estimated by using inferred cross-covariance of underlying spike trains. We see same information on the diagonals.

Beside that, additional connectives are visible. Comparing Figure 27 and Figure 28, the latter gives more details in the connectivity matrix. To be pointed out, connectivity estimation is a preliminary study as listed in this thesis. It needs further investigation and validation to make it a reliable application in the future.



**Figure 27:** Estimated connectivity for 48 neurons with covariance of fluorescence signals at lag zero.

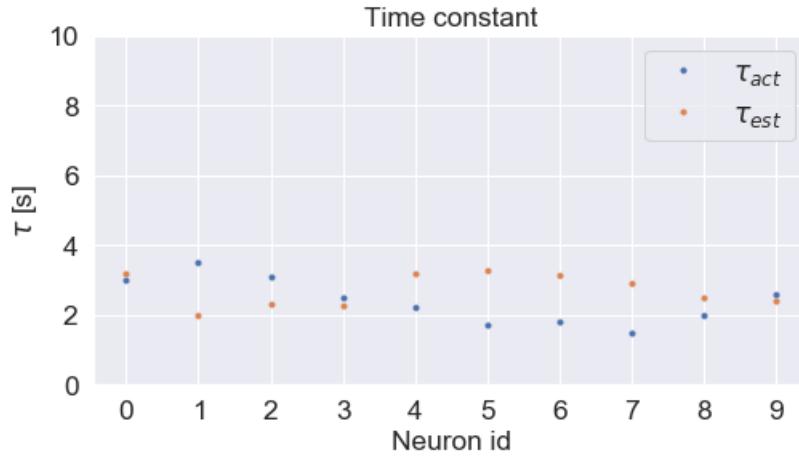


**Figure 28:** Estimated connectivity with inferred covariance of underlying spike trains.

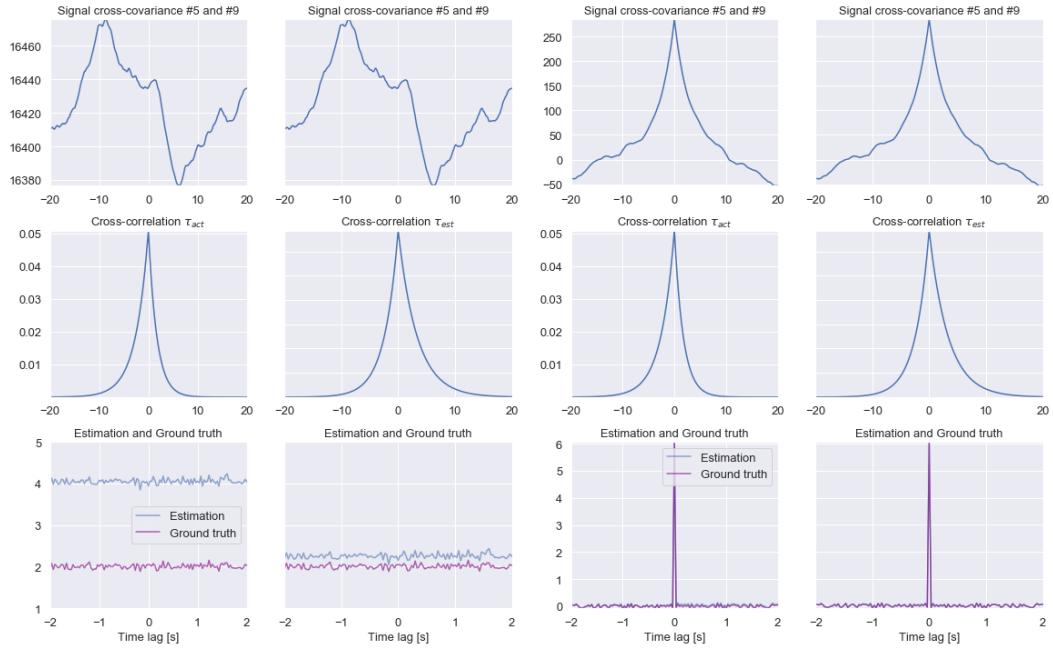
### 3.2.5 Validation with simulated data

Without ground truth, sustaining and demonstrating fidelity of the results is difficult. Therefore we applied experimental data processing flow on as set of simulated fluorescence based on MIP process. Then compare the inferred results with ground truth. MIP was simulated with parameters of experimental data. Number of neurons: 10. Time constants of the 10 neurons are [3, 3.5, 3.1, 2.5, 2.2, 1.7, 1.8, 1.5, 2.0, 2.6]. Rate of mother Poisson process is 10 Hz. Copy ratio for the 10 neurons are

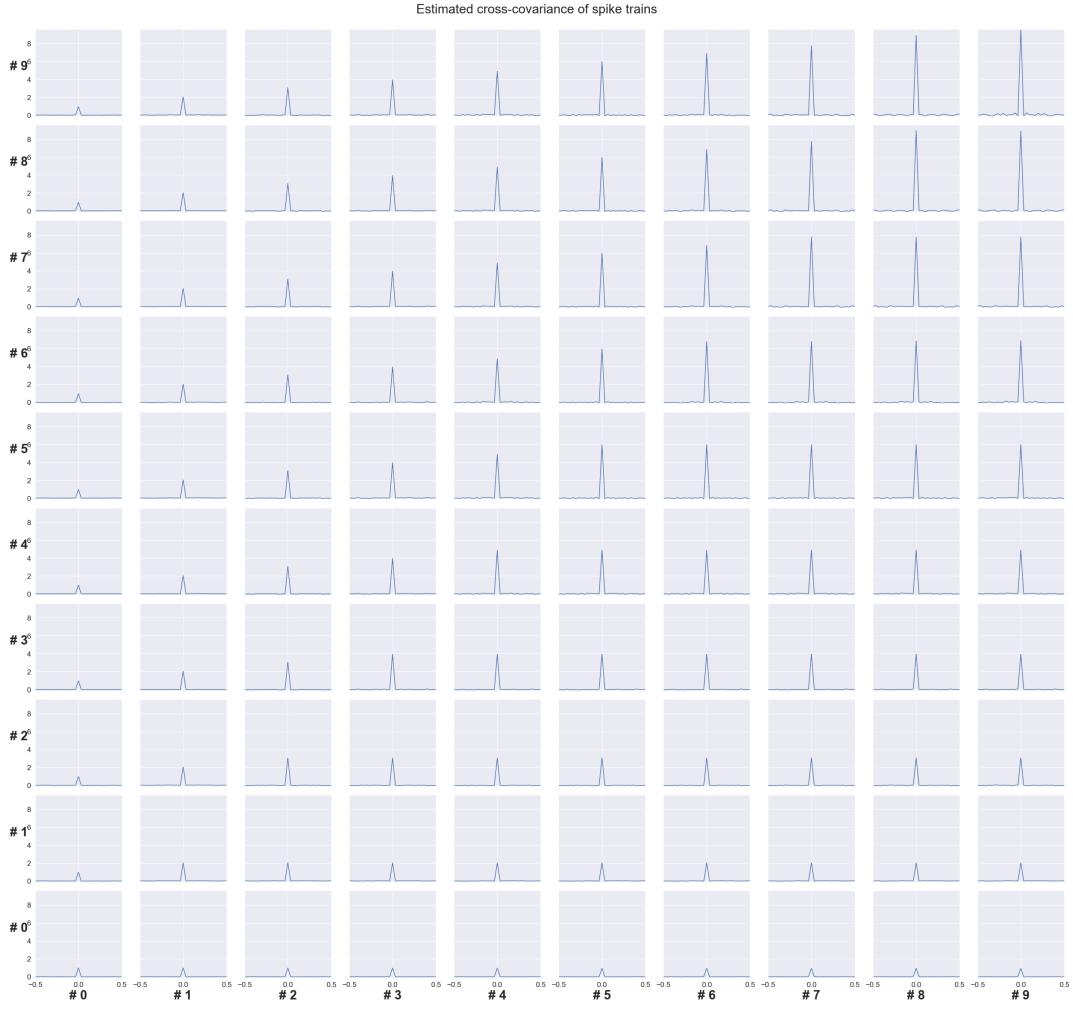
from 0.1 to 1 with step 0.1 respectively. Simulation duration 627 seconds. The first 10 seconds was cut off to eliminate the effect from rising period. Figure 29 illustrates the estimated times constants and the actual time constant for each neuron. Estimation for some neurons are very good, such as neuron #0, #3 and #9. For other neurons, there are obvious deviations. Maximum deviation is less than one second. Using the estimated time constant, cross-covariance are inferred. Results can be seen in Figure 30. Cross-covariance estimation has a high fidelity, whatever using estimated time constant or actual time constant in the kernel. It has a tolerance for errors in the time constant estimation. In contrast, the cross-correlation has an offset difference between estimation and ground truth, both in using estimated time constants and actual time constants, although the cross-correlation structure doesn't differ. Interesting is, in some cases cross-correlation estimation using estimated time constant has better accuracy than using actual time constant. To be noticed is in Section 3.1, simulated neurons have same kernel. Both cross-covariance and cross-correlation and ground truths are in high agreement. Here single neuron have different kernels. This could be a reason for the offset difference in the cross-correlations. Figure 31 demonstrates the estimation results by using estimated time constants. Copy ratio increases from neuron #0 to #9. We see neurons with high copy ratios have high peak amplitudes.



**Figure 29:** Compare estimated time constant with ground truth.  $\tau_{est}$ : estimated time constant of kernel.  $\tau_{act}$ : actual time constant of the kernel.



**Figure 30:** Compare inferred cross-covariance and cross-correlation with ground truths. Two columns on the left: cross-correlation estimation with actual and estimated kernels. Two columns on the right: cross-covariance estimation with actual and estimated kernels.



**Figure 31:** Estimated cross-covariance of 10 simulated neurons.

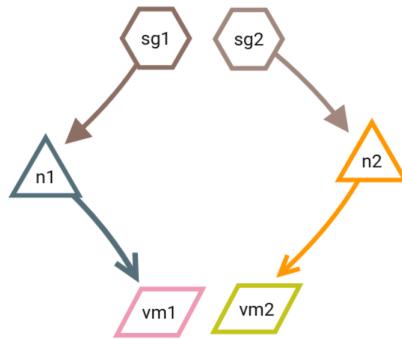
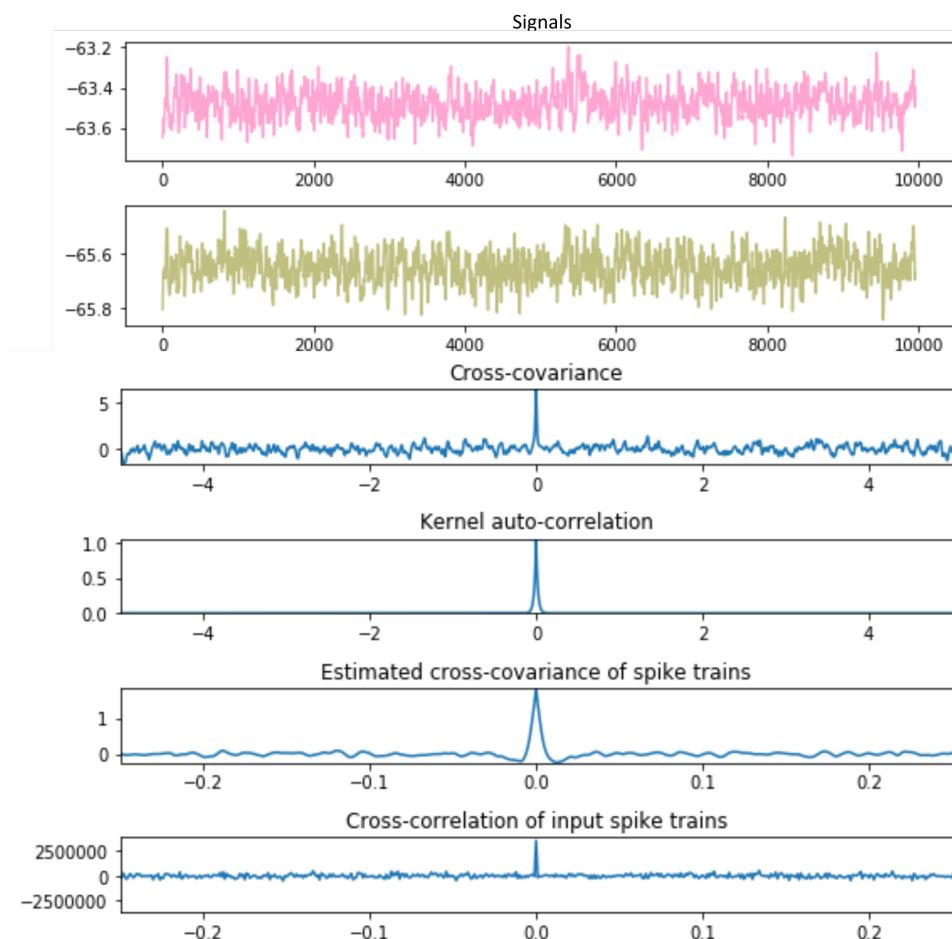
### 3.3 Membrane potential analysis

As described in Section 2.3.4, the method can also be used to analyze membrane potential. As for implementation, Nest 3.0 [20] provides a convenient and reliable way instead of simulating surrogate data in Python. Two single iaf\_PSC\_alpha neurons were created in NEST. Each is connected to a spike generator (sg1 and sg2) and multimeter (vm1 and vm2), as can be seen in Figure 32. Spike timing of sg1 and sg2 are injected with pre-simulated spike trains using SIP. Both of them are compounded Poisson process with a common spike train and a private spike

train (SIP). The simulation aims to record membrane potential in the sub-threshold regime, meaning neither neuron fires. Neuron and simulation parameters are shown in Table 1, among which the membrane time constant was intentionally set as 200 ms, which is not close to the real biological neuron but to simulate a very slow response of the membrane. The synaptic time constant is then very small in comparison. In this sense, the cross-covariance of PSC is analog to incoming spike trains' cross-covariance. Analysis results can be found in Figure 33, where we see the shape of cross-covariance and cross-correlations are well re-produced. The amplitude has a difference.

| Parameter                   | Value | Unit |
|-----------------------------|-------|------|
| Membrane capacity           | 250   | pF   |
| Resting membrane potential  | -70   | mV   |
| External input current      | 0     | pA   |
| Membrane potential          | 0     | mV   |
| Reset membrane potential    | 0     | mV   |
| External input current      | 0     | pA   |
| Spike threshold             | -55   | mV   |
| Refractory period           | 2.0   | ms   |
| Membrane time constant      | 200   | ms   |
| Excitatory synaptic weight  | 0.1   | mV   |
| Synaptic transmission delay | 1.5   | ms   |
| Simulation time             | 20000 | ms   |
| Simulation resolution       | 1     | ms   |

**Table 1:** Neuron-, synapse- and simulation-parameters in NEST.

**Figure 32:** Simulation design**Figure 33:** Simulation in NEST and analysis result. Panels from top to down are: membrane potential of neuron 1 and 2, cross-covariance of the membrane potential, kernel auto-correlation, estimated cross-covariance of input spike trains, true cross-covariance of spike trains.

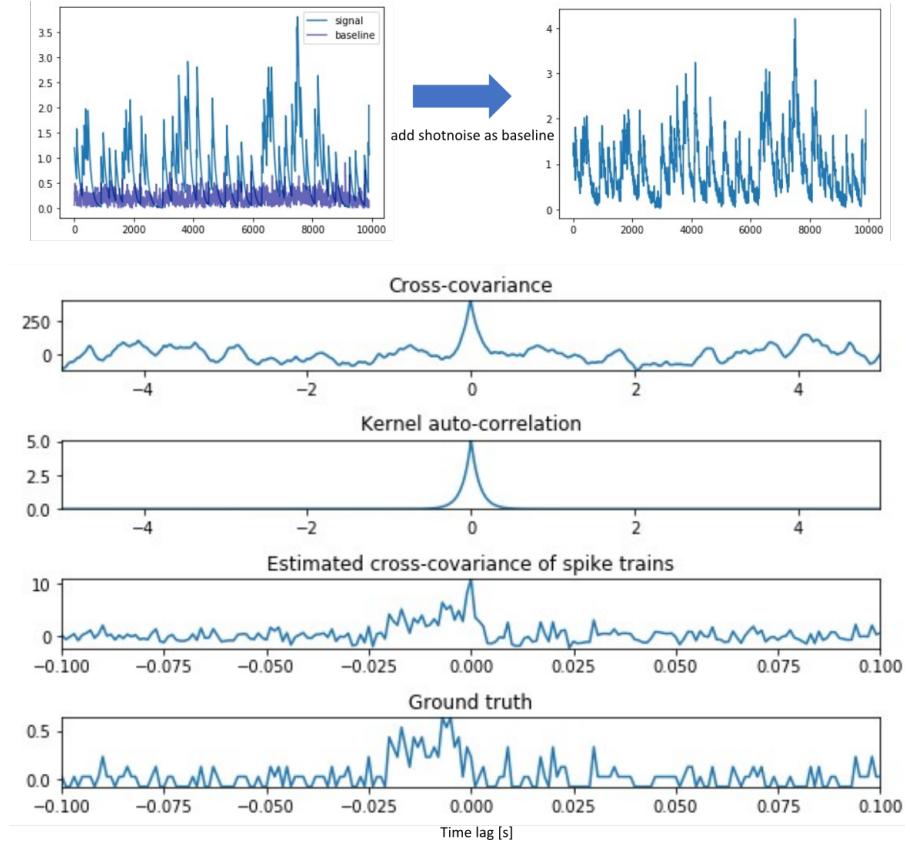
## 4 Conclusion and discussion

This thesis presents a new method to infer the cross-correlation of underlying spike trains from fluorescence signals, without performing deconvolution. Mathematical foundations of the method were introduced. The method has experimented with surrogate data, simulated with the same kernel but different correlation structures, such as the independent homogeneous Poisson process, independent in-homogeneous Poisson process, correlated spike trains, and spike trains with higher-order correlations. It turned out the correlation structure can be accurately re-represented. The method was also tried on experimental data obtained from calcium imaging of the mouse's primary visual cortex. With pre-processed data, neuronal correlations can be revealed. Furthermore, to validate the procedure of handling experimental data, the same procedure was tried with surrogate data simulated by MIP with different kernels. Results exhibited that the cross-covariance of underlying spike trains could be faultlessly captured. Furthermore, the cross-correlation structure strongly agreed with the ground truth, although there were offset differences. Finally, the same mythology can be used to analyze correlation of pres-synaptic spike trains, more specifically post-synaptic current, from membrane potentials in a sub-threshold regime based on a shotnoise model of leaky integrate-and-fire neurons. Overall, the method has a wide application range and is reliable. It saves computational costs by transferring deconvolution in the time domain into division operation in the frequency domain. There are some noteworthy aspects to mention when developing and applying the method.

### 4.1 Baseline

The general model to simulate fluorescence signal is a point process convolving with a kernel function. It is a concise model focusing on the most significant characteristics of the signal without considering baseline and noise. It was assumed that additional baseline fluctuations and noise would not significantly influence the method itself.

This assumption was verified by adding component shotnoise to the signal. As it turns out, the inferred results do not deviate much. In Figure 34, adding shortnoise makes the signal more noisy. Baseline fluctuations can lead the signal into a negative area around zero. However, the main correlation structure can still be deciphered.



**Figure 34:** Add shotnoise as fluorescence baseline. Comparing with ground truth, the estimated cross-covariance do not deviate much.

## 4.2 Pre-processing of experimental data

There are some pre-processing procedures for experimental data. Common procedures are subtracting background luminance such as neuropil from raw signals. Then, the noise would be removed by bandpass filtering. After the smoothing procedure, the baseline  $F_0$  would be calculated, and calcium transient  $(F - F_0)/F_0$  would

be obtained. In our case, none of them were performed because a different perspective was focused. Firstly, neuropil itself conceives biological meaning. Slow wave activity in the mouse visual cortex before eye-opening was present both in the neuropil and in neurons [28]. Furthermore, neuropil is a reliable reporter for Up-Down state activity in motor, and somatosensory cortex [29]. From the data exported from Suite2p, subtracting neuropil multiplied by certain factor yields negative signals. There is also no convincing way to set the value of that factor. Thus neuropil information generated from Suite2p was not handled. Besides, starting from the model, the fluorescence signal results from the spike train convolving with a kernel function. Peaks in fluorescence signal are assumed to be evoked by spiking activities of the neuron, which means raw fluorescence signal is of interest. The baseline could have some influence but would not deviate from the result. Therefore no baseline and calcium transient were calculated. Besides, the filtering procedure can introduce additional peaks, contaminating the raw signal. Correlation inferred by the method is sensitive to peaks and the time that the peaks happen. For this reason, the filtering procedure was intentionally avoided. Instead, we pick up a scrap of the signal that does not have obvious peaks and normalize the signal, which regularizes the amplitudes and avoids contaminating the signal.

### 4.3 Estimation of unknown kernel

With experimental data, the kernel is unknown. Various methods were adopted to derive the kernel. Firstly we assumed neurons have the same kernel. Investigating experimental data by manually selecting and examining peaks from neurons made it more meaningful to set each neuron's private kernel. To estimate the kernel, average spike-triggered fluorescence signal, fitting an exponential function into data, and calculating fractions of adjacent points as a function of unknown parameters were attempted. Averaging spike-triggered fluorescence signal relies on the available ground truth, a critical limitation. Given the fluorescence data model, analytically, fractions of adjacent points are a function of unknown parameters and time steps. As the time step is very small and the data is noisy, the method could not successfully get the kernel decay parameter. Above all these, the original preferred way was to use auto-correlation of fluorescence signal as a substitute for kernel auto-correlation, based on if the underlying spike train of the fluorescence signal is the Poisson process. The auto-correlation of the spike train is approximately a delta peak, which

becomes a scaling factor in the frequency domain. Therefore, the PSD of the fluorescence signal is scaled PSD of the kernel. Applying signal auto-correlation as kernel auto-correlation didn't generate meaningful results, which is attributed to noise and complex structure in the raw data. The oscillations and small structures in signal auto-correlation make the correlation structure invisible in the results. Therefore, a kernel fitting was eventually applied and worked well.

## 4.4 Outlook

Surrogate fluorescence data with ground truth is needed in scientific research. Corresponding tools were developed, such as NAOMi [30]. Data modeling method in this thesis can be potentially developed into a standalone toolkit to simulate fluorescence data. It has three core advantages. Firstly two kernel types were provided: a simple exponential kernel and a beta kernel. With beta kernel function, it is possible to customize rising time and decay time separately. As known, different calcium indicators have different response times. Beta function enables more accurate and flexible modeling. Second, exact integration was used to perform convolution numerically. It is a version of fast convolution that is much more efficient and saves computing costs, which has already been employed in large network simulation tools such as NEST. Thirdly, stochastic point process pre-defines correlation structure of underlying spike trains. A variety of point processes are available to simulating fluorescence data with diverse correlations involving customizing a neuronal population with higher-order correlations. Furthermore, compared with other spike inference algorithms, our method has no dependency on the non-negativity assumption. As long as optical reporters can capture the activity of inhibitory neurons such as GABAergic neurons, our approach can be applied to analyze the correlations. The time lag of cross-correlation peaks reveals correlation patterns. Combined with spatial information, neuronal connectivity, spiking propagation and correlation propagation can be explored further. In addition, the method is not limited to analyzing calcium imaging data. In principle, it is applicable in any linear convolution model, such as analyzing simplified model of LFP data. Last but not least, the method is based on covariance calculation, strongly associated with connectivity inference. Technically, in light of the available covariance of spiking data, connectivity can be inferred effectively. Functionally, neuronal correlation and connection have close and reciprocal relations. Combining these may facilitate a way of using calcium imaging to unravel

the myth of "salt and pepper" organized neurons in the mouse primary cortex.

## Acknowledgements

First and foremost, I would like to express my deep gratitude to Prof. Stefan Rotter, who initiated the idea and supervised the thesis. Without his guidance in every step throughout the process, it wouldn't have been possible for me to complete it. I benefit from his broad knowledge, patience as a supervisor, and creativity, focus, critical thinking, logical reasoning and stress resistance as a researcher.

I am sincerely thankful to Prof. Björn Kampa for providing the data and offering the opportunity to witness how the experiments were performed, which was very useful in my understanding of the data. I also want to thank him for the advice and discussions on the project and insightful literature recommendations. My thanks also go to Sophie Wetz from Prof. Kampa's group, who is very supportive and open, with whom I always had pleasant and helpful interactions.

I would like to specially thank Dr. Han Lu for her support personally and providing valuable and constructive suggestions to revise my writing.

## References

- [1] R. Brette and A. Destexhe, eds., *Plate section.* Cambridge University Press, 2012.
- [2] L. Huang, P. Ledochowitsch, U. Knoblich, J. Lecoq, G. J. Murphy, R. C. Reid, S. E. de Vries, C. Koch, H. Zeng, M. A. Buice, *et al.*, “Relationship between simultaneously recorded spiking activity and fluorescence signal in gcamp6 transgenic mice,” *Elife*, vol. 10, p. e51675, 2021.
- [3] F. Helmchen, J. Borst, and B. Sakmann, “Calcium dynamics associated with a single action potential in a cns presynaptic terminal,” *Biophysical journal*, vol. 72, no. 3, pp. 1458–1471, 1997.
- [4] J. Oh, C. Lee, and B.-K. Kaang, “Imaging and analysis of genetically encoded calcium indicators linking neural circuits and behaviors,” *The Korean Journal of Physiology & Pharmacology*, vol. 23, no. 4, pp. 237–249, 2019.
- [5] F. Helmchen, “Calcium imaging,” in *Handbook of Neural Activity Measurement* (R. Brette and A. Destexhe, eds.), p. 362–409, Cambridge University Press, 2012.
- [6] P. Rupprecht, S. Carta, A. Hoffmann, M. Echizen, A. Blot, A. C. Kwan, Y. Dan, S. B. Hofer, K. Kitamura, F. Helmchen, and R. W. Friedrich, “A database and deep learning toolbox for noise-optimized, generalized spike inference from calcium imaging,” *Nature neuroscience*, vol. 24, no. 9, pp. 1324–1337, 2021.
- [7] J. T. Vogelstein, A. M. Packer, T. A. Machado, T. Sippy, B. Babadi, R. Yuste, and L. Paninski, “Fast nonnegative deconvolution for spike train inference from population calcium imaging,” *Journal of neurophysiology*, vol. 104, no. 6, pp. 3691–3704, 2010.

- [8] E. Yaksi and R. W. Friedrich, “Reconstruction of firing rate changes across neuronal populations by temporally deconvolved ca<sup>2+</sup> imaging,” *Nature methods*, vol. 3, no. 5, pp. 377–383, 2006.
- [9] E. A. Pnevmatikakis, D. Soudry, Y. Gao, T. A. Machado, J. Merel, D. Pfau, T. Reardon, Y. Mu, C. Lacefield, W. Yang, *et al.*, “Simultaneous denoising, deconvolution, and demixing of calcium imaging data,” *Neuron*, vol. 89, no. 2, pp. 285–299, 2016.
- [10] B. F. Grewe, D. Langer, H. Kasper, B. M. Kampa, and F. Helmchen, “High-speed *in vivo* calcium imaging reveals neuronal network activity with near-millisecond precision,” *Nature methods*, vol. 7, no. 5, pp. 399–405, 2010.
- [11] J. T. Vogelstein, B. O. Watson, A. M. Packer, R. Yuste, B. Jedynak, and L. Paninski, “Spike inference from calcium imaging using sequential monte carlo methods,” *Biophysical journal*, vol. 97, no. 2, pp. 636–655, 2009.
- [12] E. A. Mukamel, A. Nimmerjahn, and M. J. Schnitzer, “Automated analysis of cellular signals from large-scale calcium imaging data,” *Neuron*, vol. 63, no. 6, pp. 747–760, 2009.
- [13] A. Giovannucci, J. Friedrich, P. Gunn, J. Kalfon, B. L. Brown, S. A. Koay, J. Taxidis, F. Najafi, J. L. Gauthier, P. Zhou, *et al.*, “Caiman an open source tool for scalable calcium imaging data analysis,” *Elife*, vol. 8, p. e38173, 2019.
- [14] M. Pachitariu, C. Stringer, M. Dipoppa, S. Schröder, L. F. Rossi, H. Dalgleish, M. Carandini, and K. D. Harris, “Suite2p: beyond 10,000 neurons with standard two-photon microscopy,” *BioRxiv*, p. 061507, 2017.
- [15] T. Deneux, A. Kaszas, G. Szalay, G. Katona, T. Lakner, A. Grinvald, B. Rózsa, and I. Vanzetta, “Accurate spike estimation from noisy calcium signals for ultrafast three-dimensional imaging of large neuronal populations *in vivo*,” *Nature communications*, vol. 7, no. 1, pp. 1–17, 2016.
- [16] E. Salinas and T. J. Sejnowski, “Correlated neuronal activity and the flow of neural information,” *Nature reviews neuroscience*, vol. 2, no. 8, pp. 539–550, 2001.
- [17] M. Okun and I. Lampl, “Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities,” *Nature neuroscience*, vol. 11, pp. 535–7, 06 2008.

- [18] J. Kremkow, L. U. Perrinet, G. S. Masson, and A. Aertsen, "Functional consequences of correlated excitatory and inhibitory conductances in cortical networks," *Journal of computational neuroscience*, vol. 28, no. 3, pp. 579–594, 2010.
- [19] A. Aertsen, G. Gerstein, M. Habib, and G. Palm, "Dynamics of neuronal firing correlation: modulation of" effective connectivity", " *Journal of neurophysiology*, vol. 61, no. 5, pp. 900–917, 1989.
- [20] S. Spreizer, J. Senk, S. Rotter, M. Diesmann, and B. Weyers, "Nest desktop, an educational application for neuroscience," *Eneuro*, vol. 8, no. 6, 2021.
- [21] S. Rotter and M. Diesmann, "Exact digital simulation of time-invariant linear systems with applications to neuronal modeling," *Biological Cybernetics*, vol. 81, pp. 381–402, 1999.
- [22] S. Rotter, "Lecture 4: The poisson process." Lecture of Computational Neuroscience, M.Sc. Neuroscience, University of Freiburg, May 2020.
- [23] A. Kuhn, A. Aertsen, and S. Rotter, "Higher-order statistics of input ensembles and the response of simple model neurons," *Neural Comput*, vol. 15, pp. 67–101, Jan 2003.
- [24] S. Rotter, "Lecture 11: Correlated neuronal populations." Lecture of Computational Neuroscience, M.Sc. Neuroscience, University of Freiburg, May 2020.
- [25] S. G. Benjamin Staude and S. Rotter, "Higher-order correlations and cumulants in book analysis of parallel spike trains," in *Analysis of Parallel Spike Trains* (S. Grün and S. Rotter, eds.), pp. 253–280, Springer US, 2010.
- [26] S. L. Marple Jr, *Digital Spectral Analysis*. Courier Dover Publications, 2019.
- [27] J. Schiefer, A. Niederbühl, V. Pernice, C. Lennartz, J. Hennig, P. LeVan, and S. Rotter, "From correlation to causation: Estimating effective connectivity from zero-lag covariances of brain signals," *PLoS computational biology*, vol. 14, no. 3, p. e1006056, 2018.
- [28] N. L. Rochefort, O. Garaschuk, R.-I. Milos, M. Narushima, N. Marandi, B. Pichler, Y. Kovalchuk, and A. Konnerth, "Sparsification of neuronal activity in the visual cortex at eye-opening," *Proceedings of the National Academy of Sciences*, vol. 106, no. 35, pp. 15049–15054, 2009.

- [29] J. N. Kerr, D. Greenberg, and F. Helmchen, “Imaging input and output of neocortical networks *in vivo*,” *Proceedings of the National Academy of Sciences*, vol. 102, no. 39, pp. 14063–14068, 2005.
- [30] A. Song, J. L. Gauthier, J. W. Pillow, D. W. Tank, and A. S. Charles, “Neural anatomy and optical microscopy (naomi) simulation for evaluating calcium imaging methods,” *Journal of Neuroscience Methods*, vol. 358, p. 109173, 2021.

**Appendix A – Declaration on an extra page of the Master Thesis**

**Erklärung**

Hiermit versichere ich, dass ich die eingereichte Masterarbeit selbstständig verfasst habe, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, alle wörtlich oder sinngemäß aus anderen Werken übernommenen Inhalte als solche kenntlich gemacht habe und die eingereichte Masterarbeit weder vollständig noch in wesentlichen Teilen Gegenstand eines anderen Prüfungsverfahrens war oder ist.

Ort, Datum

Unterschrift

**Declaration**

I hereby certify that I have written the Master Thesis independently, have not used any sources and aids other than those indicated, have clearly identified all contents taken over verbatim or analogously adopted from other works, and that the Master Thesis has not been or is not the subject of any other examination procedure, neither in full nor in substantial parts.

Place, Date   Freiburg, 04.10.2022

Signature   